Quality & Quantity
https://doi.org/10.1007/s11135-024-02048-9

Changing landscape of fake news research on social media: a bibliometric analysis

Abdelkebir Sahid[1] · Yassine Maleh[2] · Karim Ouazzane[3]

Accepted: 24 December 2024

Abstract
Today, social networks represent an indispensable means of communicating and disseminating customer-focused information in various fields, including news, advertising, and public relations. Social media's ease of access and ability to spread quickly allows it to attract a large audience while promoting misinformation, including "fake news," whether intentional or not. Misinformation on social media is increasing, and violent societies, individuals, organizations, and even states are growing. This study aims to identify the shifting paradigm of fake news research on social media through bibliometric analysis. The authors quantitatively examined the selected articles through co-citation and bibliographic linkage analysis from Scopus and WOS. The authors quantitatively evaluated the selected papers using co-citation and bibliographic linkage analysis. Building on the results of the bibliometric studies, the researchers developed an analysis of the evolution of social media-related research on fake news to highlight the issues and progress made by researchers in the fight against this phenomenon by presenting the key and emerging actions of its concealment, especially on social networks. This study also discusses the weak surveillance approach and its use to detect fake news with limited labeled data. Finally, we address essential proposed solutions in the field of fake news.

Keywords Disinformation · Misinformation · Deepfake · Fake news · Weak social surveillance

Yassine Maleh
y.maleh@usms.ma; yassine.maleh@ieee.org

Abdelkebir Sahid
ab.sahid@uhp.ac.ma

Karim Ouazzane
k.ouazzane@londonmet.ac.uk

1 Hassan 1st University, Settat, Morocco

2 University Sultan Moulay University, Beni Mellal, Morocco

3 London Metropolitan University, London, England

1 Introduction

During a crisis or a significant event, it is vital to identify a situation and have a transparent information space; historical political, diplomatic, civic, social, and vendetta conflicts have all involved various degrees of misleading, propagandistic, and fictitious pamphlets and inventions. In fact, since antiquity, humans have attempted to affect each other. Such influence constitutes a permanent component of interactions within communities, between societies, and, finally, between states. The purpose of this influence is to alter the behavior of one party for the benefit of the other, such as territorial concession, submission to authority, renunciation, or establishment of beliefs. Historically, power was the answer to this objective. However, early strategic theorists had already recognized the importance of influencing behavioral change through non-military methods, such as the strategies deployed by Sun Tzu in the 5th century BC and translated by Griffith et al. (Griffith & Hart, 1971). AQ1

Indeed, the above strategies have been around for a long time, although today's informational context has provided them with a more favorable space than before. Also, society's information consumption habits are changing considerably thanks to the emerging digital stream. Technological progress, the democratization of information sources, and informational echo chambers have re-engineered a permanent flow of individual information, resulting in constant consumption. This high information velocity induces consumers to devote only a few seconds to a piece of Information (2015). Most people are satisfied with the title of the article and the associated visuals without reading the whole text and ignoring the information source. Therefore, the emotional interpretation of articles has significantly increased, leading to a rise of "emotional" articles (Sługocki and Sowa 2021). This form of news relies upon cognitive biases, giving them a wider audience than fact-based, balanced, and accurate information. According to (Vosoughi et al. 2018), in a study that examined the differential spread on Twitter of authentic and verified fake news between 2006 and 2017, Fake news was significantly more widely, deeply, quickly, and extensively spread than factual news across all news categories. This impact was more about fake political news than terrorism, science, disaster, urban myths, or the state of finance. This shift influences social discourse and global social conversation, creating an environment in which conspiracy theories and unproven ideas are more exposed to the detriment of expert viewpoints and current arguments. This phenomenon, driven mainly by social media, has a more significant impact on traditional media. Faced with challenges or disruptions to their traditional economic models, many people seek answers by using social media and focusing on emotional content rather than balanced and precise content. This informational framework provides an ideal environment for disinformation campaigns.

Misinformation is a rising problem exacerbated by the rapid dissemination of fake news online, mainly via social media. Social media platforms are the preferred method for acquiring and consuming knowledge. Due to users' ease of creating and disseminating content, social media platforms are a perfect breeding ground for misleading information, including "fake news." Economic and ideological motives inspire the manufacture of false news. Fake news can significantly affect individuals, businesses, and society, impacting readers' trust in the information ecosystem.

In 2013, the stock market lost $130 billion in value minutes after an Associated Press tweet alleging a bomb that injured Barack Obama due to the uncertainty and solid emotional charge produced by the circumstances (Derouiche and Frunza 2020). The COVID-19 problem increased the population's need for knowledge in 2020 due to the uncertainty and substantial emotional charge caused by the circumstances. Since the beginning of this health crisis, an avalanche of scandals and fake news concerning coronavirus has varied from false remedies to false reporting, and conspiracy theories have followed it; this content has impacted people to at least some degree (Hasebrink and Hölig 2020). Around half of US adults (48%) claim they've seen at least some invented news or information regarding the COVID-19 epidemic. Also, over 41% of French people report being victims of fake news, illustrating how the epidemic has affected media consumption in France (Hasebrink and Hölig 2020).

Fake News, as a more contemporary term, encompasses misinformation and disinformation but specifically relates to fabricated information that mimics legitimate news media content. Following fake news can be defined as 'fabricated information imitating news media content in form but not in organizational process or intent.' This definition encompasses various types identified by Tandoc Jr et al. (2018), including:

- Satire and parody (content using humor for commentary).
- Fabrication (entirely false content created to deceive).
- Manipulation (genuine content altered to deceive).
- Propaganda (false content designed to promote particular political causes or points of view).

Understanding these distinctions is crucial as each type requires different detection approaches and mitigation strategies. While misinformation might be addressed through fact-checking and education, disinformation often requires more complex interventions involving source tracking and network analysis. This conceptual framework guides our bibliometric analysis of how researchers have approached these interconnected challenges in the digital age.

Faced with the worsening problem of misinformation, accentuated by the invasion of "fake news," the world's institutions have decided to take action, especially on social networks Lazer et al. (2018),. News organizations have banded together to combat false news. Some media groups started and funded fact-checking projects. Technology companies, criticized for their involvement in spreading fake news, have also taken action by deactivating accounts that spread fake news. Governments have also implemented legislation and launched efforts, such as the Spanish National Police (2020), which announced publishing a "Guide against Fake News" on its website on March 27, 2020 (Montesi 2021). Furthermore, to combat the spread of myths and misleading information, the Moroccan Minister of Interior has developed a website, "http://www.covidmaroc.ma/," dedicated to reporting information regarding the COVID-19 epidemic.

Finally, the war against fake news may not deter actors with vested interests from developing new methods of disseminating false information and misinformation. As a result, researchers prefer to focus on understanding incorrect information, its ways of dissemination, and the factors that make people susceptible to being misled by fake news, which is frequently used by those responsible for its creation and dissemination. Propose holistic and practical solutions to these phenomena that affect individuals, organizations, and nations.

This paper aims to understand the evolution of research in fake news, its different forms, problems, and challenges, and to explore the latest solutions based on Weak Social Surveillance (WSS) to identify disinformation proactively. The paper is organized into distinct sections, ensuring a structured and comprehensive topic exploration. Following a systematic research methodology, the paper presents the results of a meticulous bibliometric analysis. Special attention is given to Bibliographic coupling, offering valuable insights. The Analysis of Fake News Content and Context follows, providing a thorough examination. The paper then delves into disinformation recognition, offering a detailed explanation of methodologies. A dedicated section is reserved for Discussion and future work, allowing for a comprehensive exploration of implications and potential avenues for future research. The paper culminates in a succinct yet impactful Conclusion, summarizing key findings and contributions to the fields under study.

2 Research methodology

Bibliometrics, as a quantitative study of literature based on bibliographies, offers a valuable perspective for understanding the evolution of science, technology, and scholarship (White 1989) This interdisciplinary field encompasses diverse academic disciplines, incorporating mathematical, natural, and social sciences and engineering (Zupic and Čater 2015). In the context of fake news research, exploring bibliometric trends becomes essential for comprehending the landscape of scholarly contributions.

This systematic review adheres to the PRISMA (Sarkis-Onofre et al. 2021) guidelines, aiming to provide a transparent and comprehensive synthesis of the literature on bibliometrics in studying fake news. Our objectives include examining bibliometrics' various definitions and designations, evaluating its interdisciplinary applications, and focusing on its role in thematic regions associated with fake news research concerns. We establish clear eligibility criteria to ensure rigor, articulate a robust search strategy, and outline a systematic study selection and data extraction process.

The review employs a thematic analysis, emphasizing a mapping methodology to visualize and analyze the broader domain of fake news research. This approach aligns with Chen (2017) work, which advocates scientific mapping as an effective technique for understanding scientific groupings and thematic fields. The results of this review are anticipated to contribute insights into the bibliometric landscape of fake news research, providing a foundation for future scholarship in this dynamic and critical field.

2.1 Selection and justification of bibliometric data sources

The selection of Scopus and Web of Science (WoS) as primary data sources for this bibliometric analysis was based on several critical considerations. These databases are widely recognized as the most comprehensive and authoritative bibliometric data sources in academic research for several reasons:

- **Complementary Coverage**: WoS provides comprehensive coverage of established high-impact journals, particularly in Western academia (Wang & Waltman, 2016), while Scopus offers extensive coverage of emerging research venues and regional publications (Archambault et al. 2009). This complementarity is particularly valuable for studying fake news, as it represents a global phenomenon that manifests differently across cultural and geographical contexts.

- **Quality Assurance**: Both databases maintain rigorous indexing criteria and quality control in journal selection, ensuring that indexed publications meet high academic standards through peer review processes. Through its Core Collection, WoS provides access to what is generally considered the highest-impact scholarly literature.
- **Metadata Quality and Structure**: Scopus provides superior metadata quality and standardized institutional affiliations (Singh et al., 2021), including standardized keywords, author affiliations, and citation networks. Additionally, Scopus's integration with ORCID enhances author identification accuracy, which is essential for comprehensive bibliometric analysis.
- **Citation Tracking and Analysis Features**: Both platforms offer robust citation tracking capabilities and provide standardized data export formats compatible with bibliometric analysis tools, ensuring data integrity and analysis reliability. While Scopus often includes more recent publications and broader international coverage, WoS offers more detailed citation data and established impact metrics.

This combination of features makes these databases particularly suitable for mapping the evolution and current state of fake news research in social media.

2.2 Selection and justification of bibliometric analysis tool

The selection of VOSviewer (version 1.6.20) as the primary tool for bibliometric visualization and analysis was based on several methodological considerations. First, VOSviewer excels in handling large-scale bibliometric data and creates clear, interpretable visualizations of complex bibliometric networks (Van Eck and Waltman 2010). Its specialized clustering algorithms are particularly effective for identifying and visualizing research communities and thematic relationships within the scientific literature. Second, VOSviewer offers robust integration capabilities with both Scopus and Web of Science data formats, ensuring seamless data processing and analysis consistency. Third, its distance-based visualization approach, where the distance between nodes represents their relationship strength, provides an intuitive understanding of intellectual relationships within the research field. Additionally, VOSviewer's ability to generate multiple types of bibliometric networks - including co-citation, bibliographic coupling, and keyword co-occurrence networks - allows for a comprehensive analysis of the fake news research landscape from different perspectives. The software's built-in normalization techniques also help address size-dependency issues common in bibliometric analyses, ensuring more reliable comparisons between different entities in the network.

3 Study design

3.1 Optimizing bibliometric analysis: the case for integrating scopus and web of science

Bibliometric analysis has become a valuable tool for studying research trends and knowledge in various fields. It can classify and evaluate multiple aspects, such as the most prolific authors, journals, institutions, and countries. Although often performed separately on the Web of Science (WoS) and Scopus databases, many researchers emphasize the advantages of combining these two primary sources of scientific literature.

However, manually merging data from WoS and Scopus proves arduous due to differences in identification fields. Few studies clearly explain the procedure for carrying out this fusion appropriately. The present research aims to fill this gap by proposing a four-step methodology for efficiently combining the WoS and Scopus databases while eliminating duplicates.

The aim is to demonstrate the potential disparities between the results of bibliometric analyses carried out separately on Scopus, WoS, or a merged database. This approach will provide a complete and more accurate picture of the bibliometric landscape under study. Such a rigorous merging procedure is essential to guarantee the reliability of subsequent bibliometric analyses based on this combined data.

3.2 Comparison between WOS and scopus

Formerly known as Web of Knowledge, WOS encompasses databases managed by the Institute of Scientific Information (ISI) by Thomson Reuters, covering various fields such as science, social sciences, humanities (Core Collection), computer science (Inspec), and biology (Biosis). Initially dominant, WOS faced competition from Scopus, which was introduced by Elsevier and gained popularity due to its widespread accessibility in universities worldwide. The debate on which database to use for analysis has become prominent in the literature.

Researchers compare the two databases using indicators such as general characteristics, coverage, citation-based rankings, journal classification methods, keyword attribution, and the Meyer index. WOS has offered advantages in terms of coverage over the years, especially for Western English-language journals, which provide diverse information across multiple disciplines. However, some argue that despite its coverage limitations compared to Scopus, WOS includes reliable and high-impact scientific studies per Bradford's law (Bradford 1934).

On the other hand, Scopus excels in publication coverage, especially in fields like wine tourism literature. It encompasses more journals and articles than WOS, covers various document types (including conference proceedings and books), and provides better coverage of publications from developing countries and more non-English publications than WOS (Archambault et al. 2009).

3.3 Integration rationale and methodology

Our decision to integrate WOS and Scopus databases stems from a methodologically robust approach designed to comprehensively capture the evolution of fake news research (Pranckutė, 2021). This integration is particularly vital given that fake news represents an emerging, rapidly evolving global phenomenon that manifests differently across cultural and geographical contexts (Zhu et al., 2021). While WOS's strength lies in its coverage of established high-impact journals, particularly in Western academia (Wang & Waltman, 2016), Scopus provides extensive coverage of emerging research venues and regional publications (Baas et al., 2020). This complementarity offers unique advantages for studying fake news research, where insights from diverse geographical and cultural contexts are crucial for understanding the phenomenon's global impact (Martín-Martín et al., 2021).

To ensure methodological rigor, we implemented a systematic integration protocol addressing potential data consistency concerns (Visser et al., 2021). Our deduplication process employs a hierarchical verification system using DOIs, titles, authors, and publication dates as primary identifiers (van Eck & Waltman, 2010). In cases of duplicate entries, we retain Scopus data based on three key advantages: comprehensive metadata coverage, standardized institutional affiliations, and precise author identification through ORCID integration (Singh et al., 2021). This choice aligns with recent bibliometric studies highlighting Scopus's superior metadata quality and completeness in emerging research fields (Huang et al., 2020).

The challenge of differing citation counting methodologies between databases is addressed through a carefully designed normalization framework (Glänzel & Moed, 2013). We first analyze citation patterns separately within each database, then apply a cross-validation process to verify trend consistency before final integration (Zhang & Glänzel, 2017). This approach maintains the integrity of citation metrics while enabling comprehensive analysis of research impact patterns. Special attention is given to consistency in citation counts through careful normalization procedures that account for database-specific counting methods (Waltman, 2016).

The advantages of this integrated approach manifest in several critical areas for fake news research. First, it enables better identification of international research collaborations, crucial for understanding how different regions approach fake news detection and mitigation (Shu et al. 2020a, b, c). Second, it provides improved visibility of emerging research networks, particularly from regions that might be underrepresented in single-database studies (Wang et al., 2020). Third, it offers comprehensive coverage of interdisciplinary work, essential given that fake news research spans multiple domains including computer science, social sciences, and communication studies (Zhou et al., 2020).

While acknowledging the methodological challenges of combining databases, our approach incorporates specific measures to maintain data integrity (Li et al., 2021). These include detailed documentation of integration procedures, systematic validation of results across both databases and careful attention to maintaining data consistency throughout the analysis process (Chen et al., 2022). This methodological framework aligns with current trends in bibliometric research, which increasingly recognizes the value of multi-database approaches for studying emerging fields (Yang et al., 2021).

3.4 Correlation between WOS and scopus

Despite their distinct characteristics, WOS and Scopus exhibit significant correlations. Although differing in scope and coverage policy, they maintain a high level of correlation, as shown by Archambault et al. (2009). Sánchez et al. (2017) note that WOS includes 54% of Scopus journal titles, while Scopus covers 84% of WOS titles. Sánchez et al. (2017) discovered a strong correlation ($R2 = 0.78$) between the two databases concerning the number of articles, with high average correlations observed in various fields such as social sciences, health sciences, and physical sciences. While comparisons do not favor one or the other for bibliometric analyses, researchers suggest the essential nature of both databases. Sánchez et al. (2017) advocate for the complementary nature of WOS and Scopus, while Mongeon and Paul-Hus (2016) propose their combined use. Given their complementary nature and evident correlation in publication coverage, their merger into a single database becomes crucial.

3.5 Data collection

Table 1 showcases the Boolean queries used to gather scholarly literature addressing fake news detection systematically. Two distinct search strategies were implemented, one for Scopus and the other for Web of Science (WOS), each tailored to the database's search capabilities.

For Scopus, the query focused on articles published between 2016 and 2023 and was restricted to English-language publications categorized as either Articles or Reviews, ensuring relevance and quality. Terms synonymous with false news detection, such as "fake news," "clickbait detection," and "deep fake detection," were targeted, with additional constraints applied to document types, language, and publication stage to enhance precision.

Similarly, the WOS query encompassed a broad spectrum of terms related to false news detection, spanning publication years from 2012 to 2023. The search criteria were refined to include only Open Access articles indexed within the Social Sciences Citation Index (SSCI) or the Science Citation Index Expanded (SCI-EXPANDED) and written in English. Document types were restricted to Review Articles and Articles to ensure a focused retrieval of relevant scholarly contributions.

These meticulous search strategies, outlined in Table 1, were pivotal in facilitating a comprehensive and rigorous data collection process, ultimately enabling the acquisition of diverse scholarly insights into the multifaceted phenomenon of false news detection.

3.6 Information about data

Table 2 presents a comparative overview of the leading bibliometric indicators related to the publication of fake news articles covered by the Scopus and Web of Science (WOS) databases during 2016–2023 and 2017–2024, respectively. It highlights significant growth

Table 1 Boolean queries for data collection from scopus and web of science (WOS)

Scopus	WOS
ALL ("dataset" AND "false news detection" OR "fake news" OR "clickbait detection" OR "misinformation detection" OR "rumor detection" OR "satire detection" OR "deep fake detection" OR "deep fake detection" OR "fake news detection") AND PUBYEAR > 2015 AND PUBYEAR < 2024 AND (LIMIT-TO (DOCTYPE, "ar") OR LIMIT-TO (DOCTYPE, "re")) AND (LIMIT-TO (LANGUAGE, "English")) AND (LIMIT-TO (PUBSTAGE, "final"))	ALL=("dataset" AND "false news detection" OR "fake news detection" OR "clickbait detection" OR "misinformation detection" OR "rumor detection" OR "satire detection" OR "deep fake detection" OR "deep fake detection" OR "fakenew detection") and Open Access and 2022 or 2021 or 2020 or 2019 or 2018 or 2017 or 2016 or 2015 or 2014 or 2012 or 2023 (Publication Years) and Social Sciences Citation Index (SSCI) or Science Citation Index Expanded (SCI-EXPANDED) (Web of Science Index) and English (Languages) Review Article or Article (Document Types))

Table 2 Main information on fake news publications from scopus and WOS (2016–2023) source Rstudio 4.3.3

Scopus	Results	WOS	Results
MAIN INFORMATION ABOUT DATA		MAIN INFORMATION ABOUT DATA	
Timespan	2016–2023	Timespan	2017–2024
Sources (Journals, Books, etc.)	1149	Sources (Journals, Books, etc.)	140
Documents	4142	Documents	482
Annual Growth Rate (%)	161.76%	Annual Growth Rate (%)	42.62%
Document Average Age	2.09	Document Average Age	1.98
Average Citations per Document	18.93	Average Citations per Document	12.79
DOCUMENT CONTENTS		**DOCUMENT CONTENTS**	
Keywords Plus (ID)	12,970	Keywords Plus (ID)	280
DOCUMENT TYPES		**DOCUMENT TYPES**	
Article	3877	Article	440
Review	265	Review	28
		Article; Early Access	11
		Article; Proceedings Paper	1
		Article; Retracted Publication	2

in the number of sources, documents, and annual growth rates associated with fake news publications, reflecting increasing prevalence and research interest in this phenomenon.

The data reveals that Scopus indexed a more significant number of sources (1149) and documents (4142) related to fake news compared to WOS (140 sources and 482 papers) during the specified periods. However, WOS had more referenced publications (15929) than Scopus, which did not provide this information.

Interestingly, the annual growth rate of fake news publications was substantially higher in Scopus (161.76%) compared to WOS (42.62%), indicating a more rapid increase in the coverage of this topic within the former database. Additionally, the average age of documents in both databases was relatively recent, with Scopus having a slightly higher average of 2.09 years compared to 1.98 years for WOS.

Moving to document contents, Scopus predominantly housed articles (3877) followed by reviews (265), while WOS had a higher proportion of articles (440) and a smaller number of reviews (28). Notably, Scopus also indexed a few retracted publications (2), indicating instances where fake news articles may have been published but later retracted due to inaccuracies or ethical concerns.

Furthermore, Scopus recorded an average of 18.93 citations per document. In contrast, WOS had an average of 12.79 citations per document, suggesting differences in the impact and visibility of fake news publications across the two databases.

Additionally, Scopus indexed a significantly more significant number of "Keywords Plus" (12970) compared to WOS (280), suggesting a more comprehensive coverage of keyword terms related to fake news.

3.7 Integration of WOS and scopus for bibliometric analysis

Bibliometrix, an R package for bibliometric analysis, offers extensive tools for collecting, managing, and visualizing bibliographic data, making it a crucial resource for researchers studying publication patterns and trends. (Aria & Cuccurullo, 2017). Abbreviated Bibliometrix as "Bx" in the scholarly literature.

In the process of integrating Scopus and Web of Science datasets, the mergeDbSources function within Bibliometrix (version 4.0.0) takes center stage. Employing a nuanced approach, this function harnesses key columns, including UT (Unique Article Identifier), DI (Digital Object Identifier), Title, and Publication Year, to sift through and eliminate duplicates. Duplicates, characterized by identical values for UT, DI, Title, and Publication Year across both datasets, are promptly identified and addressed. The default protocol is to retain the Scopus entry, ensuring continuity while seamlessly integrating any missing data from the corresponding Web of Science record. This meticulous de-duplication strategy maintains data integrity and prioritizes completeness by preserving the most comprehensive record among identified duplicates.

The integration process exemplifies Bibliometrix's commitment to facilitating robust bibliometric analysis while maintaining harmony between efficiency, accuracy, and adaptability. Detailed code instructions for the integration process can be found in Table 3.

Table 4 summarizes the primary information for the merged database, including timespan, document counts, average citations per document, and more. While mergeDbSources embodies this specific de-duplication logic, researchers retain the flexibility to tailor their approach using standard R functions such as dplyr::distinct, allowing for custom de-duplication strategies tailored to the unique demands of their research.

Table 3 Code to integrate scopus and WOS files from R-studio 4.3.3

The integration process exemplifies Bibliometrix's commitment to facilitating robust bibliometric analysis while maintaining harmony between efficiency, accuracy, and adaptability. # Install required packages if necessary
install.packages("bibliometrix")
install.packages("readxl")
install.packages("writexl")
library(bibliometrix)
library(readxl)
library(writexl)
# Load Scopus and Web of Science data
scopusfd <- read_excel("scopusfd.xlsx")
WosD <- read_excel("WosD.xlsx")
# Merge data using the mergeDbSources function and remove duplicates
merged_data <- mergeDbSources(WosD, scopusfd, remove.duplicated = TRUE)
# Save merged data without duplicates in a new Excel file
write_xlsx(merged_data, "Merged.xlsx")

Table 4 Main information summary of merged database by R-studio 4.3.3 (2017–2024)

Merged Database	
Description	Results
MAIN INFORMATION ABOUT DATA	
Timespan	2017–2024
Sources (Journals, Books, etc.)	1211
Documents	4181
Annual Growth Rate (%)	42.62%
Document Average Age	1.98
Average Citations per Document	12.69
References	230,393
DOCUMENT CONTENTS	
Keywords Plus (ID)	12,119
Author's Keywords (DE)	9141
AUTHORS	
Authors	1532
Authors of Single-Authored Documents	7
AUTHORS COLLABORATION	
Single-Authored Documents	9
Co-Authors per Document	2.26
International Co-Authorships (%)	4.138%
DOCUMENT TYPES	
Article	3777
Article; Early Access	11
Article; Proceedings Paper	1
Article; Retracted Publication	2
Book Chapter	134
Review	256

3.8 Data cleaning and quality assurance

Per the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) methodology, Fig. 1 illustrates this study's data collection, cleaning, and enhancement process. The steps undertaken are rigorously outlined to ensure transparency and replicability of the methodology. The data cleaning process involved merging and deduplicating datasets from Scopus and Web of Science (WoS) using the bibliometrix package (4.3.3) in RStudio. Specifically, the "mergeDbSources" function facilitated this process, followed by an automatic deletion procedure that successfully removed 313 duplicated documents, thereby enhancing the quality and reliability of the collected data. Gorraiz's (2016) study on adopting DOIs in the Web of Science Core Collection and Scopus, spanning 2005 to 2014, underscores the importance of permanent identification for improved visibility and citation practices. Conversely, the analysis of Franceschini et al. (2016) highlights the challenges associated with managing database errors, particularly in Scopus and WoS, emphasizing the criticality of data accuracy for dependable bibliometric analyses. By adhering to PRISMA guidelines and employing robust data cleaning techniques, this study ensures the integrity and validity of the dataset, facilitating accurate analysis and interpretation of the results.

Ensuring the quality and reliability of data collected from academic databases like Scopus and Web of Science (WoS) is crucial in research. This involves rigorous steps to clean and refine the data, addressing errors and inconsistencies. Once this cleaning process is

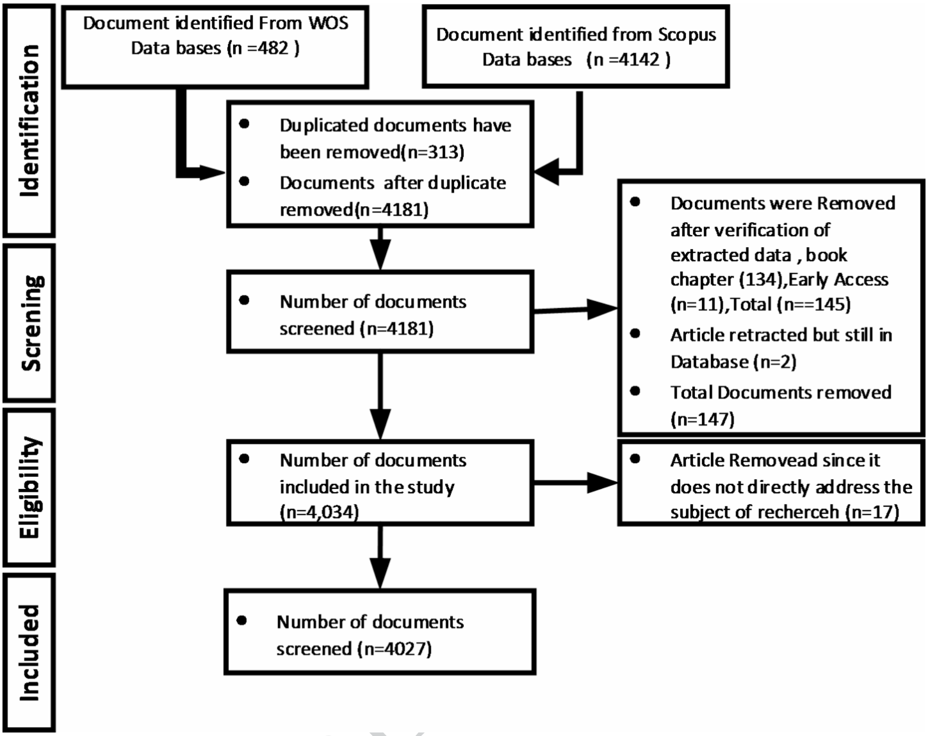

Fig. 1 Data extraction procedures comply with the PRISMA framework

complete, it's essential to analyze the results accurately. In this context, we'll discuss the data cleaning process undertaken to enhance the quality of our dataset derived from Scopus and WoS and integrate mixed data from these sources. Gorraiz et al. (2016) examine the adoption of DOIs in the Web of Science Core Collection and Scopus between 2005 and 2014, highlighting the importance of this permanent identification for improved visibility and citation practices. On the other hand, the analysis by Franceschini et al. (2016) sheds light on the challenges encountered in managing database errors, particularly in Scopus and WoS, thus underscoring the significance of data accuracy for reliable bibliometric analyses.

Subsequently, we utilized the "mergeDbSources" method in Rstudio, and the automatic deletion process resulted in the removal of 313 duplicated documents. The "Merged Database" table provides an overview of the combined dataset obtained from Scopus and WoS, covering the time-span of 2017 to 2024. It reveals 1211 sources (journals, books, etc.) and 4181 documents. The annual growth rate of publications was 42.62%, with an average document age of 1.98 years. The average citations per document stood at 12.69, and the merged dataset contained 230,393 references. Table 4 reports 12,119 "Keywords Plus" and 9141 author-supplied keywords regarding document contents. The dataset encompasses 1532 authors, with only 7 authors contributing to single-authored documents. The level of international co-authorship was 4.138%, and the average number of co-authors per document was 2.26. In terms of document types, the majority were articles (3777), followed by book Chaps. (134) and reviews (256). Additionally, there were 11 early access articles, 1 proceedings paper, and 2 retracted publications. Notably, the table highlights that 2.77% of the documents lacked DOI information, which is considered a “Good” status according to the missing data percentage thresholds. This missing DOI rate, while not ideal, is relatively low compared to other metadata fields.

After the initial automated de-duplication process, we conducted a thorough manual check to verify the accuracy of the DOI.Apart from the reported missing DOIs, no other errors were found at the DOI level. Noted in the manual stage, we proceeded with the following additional data cleaning steps: (1) Deleting 2 retracted articles from the dataset to ensure the integrity and reliability of the included publications. (2) Remove the 11 early access articles, as they represent preliminary versions that may undergo modifications before final publication, making them less suitable for solid bibliometric analysis. This step was taken to avoid potential duplication or inconsistencies in the data. (3) Elimination of the 134 book chapters, as this study was limited to indexed articles and journals with a previewing process due to the focus on analyzing the impact and influence of research published in the field of “fake news.”

3.9 Citation analysis and database integration process

Our database integration process employs a systematic approach utilizing the mergeDbSources function from Bibliometrix 4.0.0. This process analyzes four key identifiers: unique article identifiers (UT), DOIs, titles, and publication years. The analysis of citation patterns across databases revealed distinct characteristics:

Initial citation metrics showed:

- WOS: 12.79 citations per document.
- Scopus: 18.93 citations per document.
- Merged database: 13.11 citations per document.

The moderation in the final average demonstrates our deduplication effectiveness in preventing citation count inflation while maintaining comprehensive coverage. This outcome aligns with Sánchez et al. (2017)’s findings of strong correlation ($R2=0.78$) between the databases’ citation patterns.

The integration process revealed several key insights specific to fake news research:

- Highly cited papers maintained consistent relative rankings across both databases.
- Regional research impact became more visible in the merged dataset.
- Citation patterns reflected the rapid evolution of fake news research, particularly during critical events such as elections and the COVID-19 pandemic (Zhou et al. 2019).

These metrics validate our methodological choice, showing that the merged database provides a more balanced and comprehensive view of citation impact in fake news research while avoiding duplicate counting issues highlighted by Mongeon and Paul-Hus (2016).

3.10 Primary information of merged database

The “Cleaned Merged Database” Table 5 presents a comprehensive overview of the refined dataset obtained after combining and processing data from Scopus and Web of Science

Table 5 Summary of cleaned merged database (2017–2024) by R-studio 4.3.3

Cleaned merged database	
Description	Results
MAIN INFORMATION ABOUT DATA	
Timespan	2017:2024
Sources (Journals, Books, etc.)	1092
Documents	4029
Annual Growth Rate %	42.62
Document Average Age	1.98
Average citations per doc	13.11
References	222,045
DOCUMENT CONTENTS	
Keywords Plus (ID)	12,097
Author's Keywords (DE)	8970
AUTHORS	
Authors	13,845
Authors of single-authored docs	214
AUTHORS COLLABORATION	
Single-authored docs	223
Co-Authors per Doc	4.05
International co-authorships %	4.195
DOCUMENT TYPES	
Article	3774
Review	255

(WoS) databases, spanning 2017 to 2024. It reveals 1092 sources, encompassing journals, books, and other scholarly publications, contributing to 4029 documents within the dataset. The annual growth rate of publications is 42.62%, with an average document age of 1.98 years. The average number of citations per document is 13.11, and the consolidated dataset contains 222,045 referenced publications. Delving into the document contents, the table reports 12,097 "Keywords Plus" identifiers and 8970 author-supplied keywords, providing insight into the topical coverage of the included works. The dataset encompasses contributions from 13,845 authors, with 214 authors contributing to single-authored documents. Collaborative efforts are highlighted, with an average of 4.05 co-authors per document and an international co-authorship rate of 4.195%.

Regarding document types, the majority are articles (3774), followed by 255 reviews, reflecting the emphasis on peer-reviewed scholarly publications within the dataset. Notably, the table highlights that 2.18% of the documents lacked DOI (Digital Object Identifier) information, which is considered a "Good" status according to the missing data percentage thresholds. Although not ideal, this missing DOI rate is relatively low compared to other metadata fields. After the initial automated de-duplication process, a thorough manual check was conducted to verify the accuracy of the DOIs present in the dataset. Apart from the reported 2.18% missing DOIs, no other errors were found at the DOI level, ensuring high reliability for this crucial identification element in the cleaned dataset.

This methodology highlighted a comprehensive approach to building a reliable and targeted dataset for bibliometric analysis of fake news research. The data collection process involved implementing robust search strategies tailored to the Scopus and Web of Science databases, targeting relevant literature on fake news detection through carefully crafted

Boolean queries. Subsequently, data integration from these two leading academic databases was facilitated by the bibliometrix package's mergeDbSources function, which used a rigorous de-duplication process to identify and remove duplicate entries while preserving the complete information.

A crucial data cleaning step was undertaken, involving manual DOI verification, removal of retracted articles and early access publications, and elimination of book chapters to align the dataset with the scope and objectives of the study. These meticulous steps resulted in a refined dataset comprising 4029 documents from 1092 sources, covering a period from 2017 to 2024 and encompassing articles and journals related to fake news research. This comprehensive dataset, with its well-defined bibliometric indicators and rigorous data processing, provides a solid basis for conducting in-depth bibliometric analyses and mapping the evolution of this critical area of research.

3.11 Results of bibliometric analysis

3.11.1 Research evolution

Figure 2 illustrates a remarkable surge in research publications across the merged database, Scopus, and WOS from 2019 to 2020, followed by a gradual decline until 2023. This pattern consistently manifests across all three databases, suggesting a broader shift or increased focus in research activity during this period. Interestingly, while article volumes decreased after the 2020 peak, citation counts continued to rise steadily, implying a potential lag in the recognition and impact of these publications. This phenomenon underscores the importance of considering quantitative metrics like article counts and qualitative indicators such as citations when evaluating research trends.

The reasons behind the rapid growth in 2019–2020 and the subsequent decline are multifaceted and warrant further investigation. Potential contributing factors could include changes in research funding allocation, shifts in institutional or national research priorities,

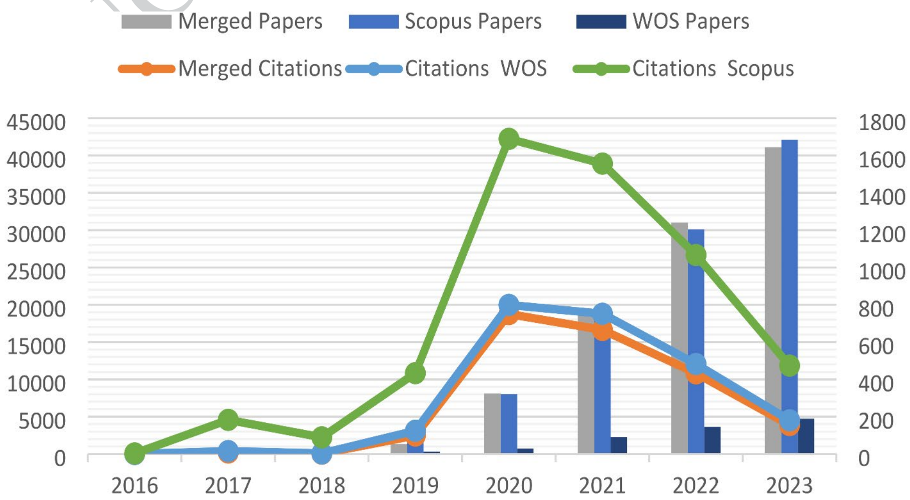

Fig. 2 Fake news research evolution in the 3 databases scopus WOS se (2016–2023)

collaborative efforts within the global scientific community, or significant global events or disruptions during that time frame. For instance, the COVID-19 pandemic, which emerged in late 2019 and persisted through 2020, could have significantly influenced research priorities and funding allocations, potentially driving the observed surge in publications related to public health, medicine, and associated fields. Additionally, exploring the specific subject areas or disciplines that experienced the most significant fluctuations could provide valuable insights into emerging research trends or areas of heightened interest.

These observed patterns may affect future research directions, resource allocation strategies, and policy decisions within academic and research institutions. The surge in publications during 2019–2020 could signal the emergence of new, promising fields or the convergence of interdisciplinary efforts addressing complex global challenges. Consequently, sustained investment and support for these areas might be warranted to capitalize on this momentum and foster continued innovation.

However, it is crucial to acknowledge the potential limitations and biases inherent in the databases used for this analysis. Factors such as indexing criteria, geographic representation, and disciplinary coverage may influence the accuracy and completeness of the data. Furthermore, citation metrics should be interpreted cautiously, as various factors can affect them, including self-citation practices and citation biases.

Overall, while the observed trends provide valuable insights into the dynamics of research activity, a nuanced approach considering both quantitative and qualitative indicators and contextual factors, such as major global events, is essential for a comprehensive understanding of the research landscape and its implications for the future.

3.12 Selection and justification of bibliometric analysis tool

The bibliometric network analysis of fake news encompassed all 4027 papers in the merged database, visualized through VOSviewer software (1.6.20) following the methodology previously described. To ensure map clarity, we incorporated the first author's name following Van Eck and Waltman's approach (Van Eck and Waltman 2010). Two complementary analytical techniques were employed: co-citation and bibliographic coupling. Co-citation analysis gauged the frequency of joint citations, while bibliographic coupling quantified shared citations between two documents in their respective bibliographies (Mas-Tur et al. 2021). These methods, widely applied across diverse research domains, facilitate identifying related works, assessing intellectual field structures, and exploring emerging trends (Phan Tan 2021). However, it is crucial to consider each technique's strengths and limitations when interpreting results, given their varied perspectives on document relationships and addressed subjects (Surwase et al. 2011).

4 Results

4.1 Data and procedure

The bibliometric network analysis of fake news encompassed all 4027 papers in the merged database, visualized through VOSviewer software (1.6.20) per the methodology described in the descriptive statistics section, with a focus on maintaining map clarity by incorporating the first author's name, following Van Eck and Waltman's approach (Van Eck and Waltman 2010). Co-citation and bibliographic coupling, established techniques in bibliometrics, were employed to scrutinize document relationships based on citation data. Co-citation analysis gauged the frequency of joint citations, while bibliographic coupling quantified shared citations between two documents in their respective bibliographies (Mas-Tur et al. 2021). Widely applied across diverse research domains, these methods facilitate identifying related works, assessing intellectual field structures, and exploring emerging trends (Phan Tan 2021). However, it is crucial to consider each technique's strengths and limitations when interpreting results, given their varied perspectives on document relationships and addressed subjects (Surwase et al. 2011).

Our comprehensive bibliometric study on "Fake News" integrates co-citation and bibliographic coupling analyses. Co-citation analysis maps knowledge structures, identifies evolving research trends, and assesses the author's impact on fake news. Simultaneously, bibliographic coupling analysis reveals content-relatedness among works, highlighting journal interconnectedness and identifying pivotal works bridging different fake news research subfields. This combined approach enriches our investigation, providing a nuanced understanding of the scholarly landscape around fake news by uncovering knowledge dynamics, thematic similarities, and key contributors in the field.

Given the topic's specificity and relative novelty, co-citation analysis using cited sources spanned a decade. However, only publications with a minimum of 20 citations were included, considering the advice of (Černe et al. 2016), who emphasized that co-citation analysis is less sensitive to the starting year, relying more on secondary articles for its insights.

4.2 Co-citation cluster analysis

4.2.1 Co-citation by authors

Author co-citation analysis is an essential bibliometric technique for mapping intellectual structures and delineating key research fronts within specific disciplines or fields (White 1989) This approach is based on the conceptual link observed when two authors are jointly cited by other scientific works, suggesting a shared perspective or contribution (McCain 1990). By quantifying the frequency with which distinct pairs of authors are co-cited in the literary corpus, researchers can discover emerging clusters and trends indicative of dominant intellectual currents (Culnan 1986; Gmür 2003). This analytical insight is invaluable for highlighting seminal authorities and pioneers who have profoundly influenced the trajectory of a field (Boyack and Klavans 2010). Furthermore, longitudinal examinations of author co-citation networks can reveal the dynamic evolution of knowledge domains, illuminating theoretical paradigms' increasing or declining role over time (Ding et al. 2001; Hota et al. 2020). Consequently, author co-citation analysis is an indispensable tool for information specialists, research strategists, and domain experts, facilitating the systematic elucidation and graphical representation of a discipline's conceptual and social fabric(Leydesdorff 1998) (Leydesdorff 1998).

Based on this fundamental understanding, our research aims to apply author co-citation analysis to determine the most influential authors in the fake news domain. Figure 3 illustrates the co-citation network of these prolific authors. Notably, the minimum citation

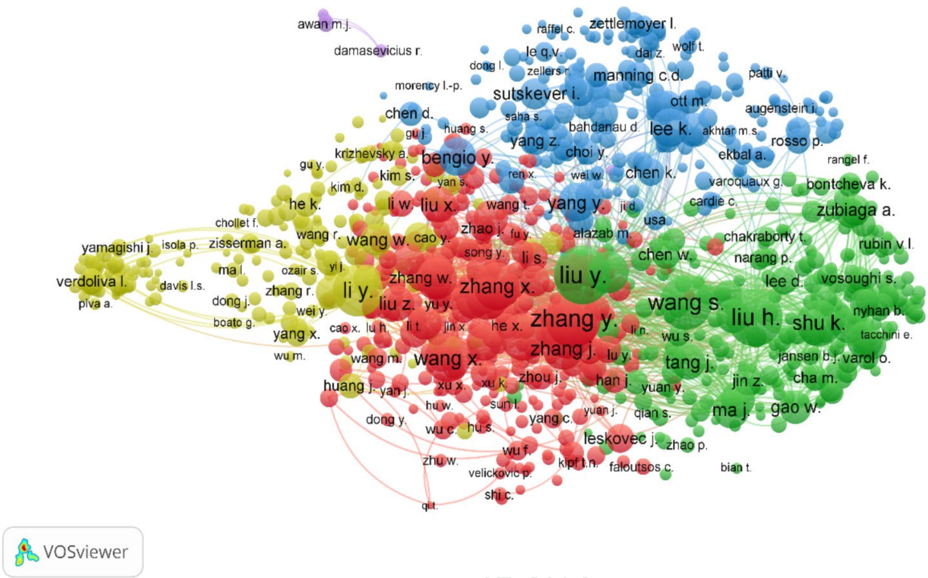

Fig. 3 Visualization's network of fake news research filed; co-citation analysis by authors(Source Vos-Viewer 1.6.20)

threshold for an author to be included was set at 20, and of the 230,931 authors analyzed, 5,211 met or exceeded this criterion.

Table 6 provides a comprehensive overview of the 10 most influential authors on fake news, establishing their respective rankings based on citation values. Liu Yao (University of Southern California, USA) emerges as a leading figure in this scientific landscape, boasting a remarkable citation value of 999, a total link strength of 2,219, and 2,156 total citations. Zhng Yongdong (China University of Science and Technology, China) follows close behind, with an equally impressive citation value of 999, alongside 255,103 total links, a total link strength of 2,156, and 2,125 total citations. Huan Liu (Arizona State University, USA), Suhang Wang (Pennsylvania State University, USA), and Yajun Wang (Microsoft Corporation, USA) occupy the following positions, highlighting their significant contributions to the discourse on fake news. Yuezun Li (Ocean University of China, China), Jian Ping Li (University of Electronic Science and Technology of China, China), Xiang Wang (University of Science and Technology of China, China), and Xichen Zhang (Saint Mary's University, Canada) round out this elite group of authors. Finally, Kai Shu (Illinois Institute of Technology, USA) deserves a special mention, although his number of citations is lower.

Based on total link strength. Liu Yao emerges as a prominent figure in this scholarly landscape, boasting a remarkable citation value of 999 and a total link strength of 2,219 and 2,156 total citations. Zhang Yongdong follows closely, commanding an equally impressive citation value of 999, alongside substantial figures of 255,103 total links, 2,156 total link strength, and 2,125 total citations. Huan Liu, Suhang Wang, and Yajun Wang secure subsequent positions, highlighting their significant contributions to the discourse surrounding fake news.

Table 6 Top 10 most influential authors based on co-citation analysis in the fake news domain

Value label	Links	Total link strength	weights. Citations	Affiliation
Liu, Yao	999	257,753	2219	University of Southern California, 3650 McClintock Ave, Los Angeles, 900,089, CA, United States
Zhang Yongdong	999	255,103	2156	University of Science and Technology of China, Hefei, China
Huan Liu	999	236,802	2125	Arizona State University, Tempe, United States
Yajun Wang	999	215,233	1981	Microsoft Corporation Verified email at microsoft.com
Suhang Wang	999	227,560	1929	Pennsylvania State Universi, University Park, United States
Yuezun Li	999	211,973	1732	Assistant Professor, Ocean University of China Verified email at ouc.edu.cn
Jian Ping Li.	999	179,362	1499	University of Electronic Science and Technology of China, School of Computer Science and Engineering, Chengdu, 611,731, China
Xichen Zhang.	999	171,573	1508	Assistant Professor, Sobey School of Business, Saint Mary's University
Xiang Wang	999	175,040	1500	University of Science and Technology of China
Kai Shu	2	997	165,775	Illinois Institute of Technology Chicago IL, USA

4.3 Bibliographic coupling

The second bibliographic research concentrated on bibliographic linking, occurring when two items in their bibliographies mention a third standard piece. This linkage demonstrates the possibility that the two articles are about the same issue; the "strength of linkage" of the two papers grows as the number of shared sources they cite grows (Kessler 1963). Bibliographic linking has been criticized when writers used it to demonstrate future research based on existing patterns. Despite this critique, Ferreira (2018) claims that bibliographic association is still valuable for situating recent contributions to the subject.

Indeed, Garfield (2009), given the likelihood that two papers in the group may belong to a different topic, said that the co-citation method is the most accurate predictor of topic proximity. Ferreira (2018) contends that the approaches are complementary since the bibliographic linking method is "backward-looking," and the co-citation approach is "forward-looking."

4.3.1 Bibliographic coupling by documents

Bibliographic coupling is a technique used to measure the similarity between scientific publications based on the number of shared references they cite. Two publications are considered bibliographically coupled if they share many references. This coupling indicates that the publications are likely related to content or research topic.

The bibliographic coupling of the top 10 authors is presented in Fig. 4. Considering the large number of cited documents, an author's minimum number of citations was set to 20, and among the 4000 documents, only 587 met the threshold.

The provided references offer a diverse landscape of contributions to fake news research. Notably, (Vosoughi et al. 2018) stand out with numerous links and citations, reflecting its considerable impact on the scholarly community. Shu et al. (2017) and Allcott and Gentzkow (2017) have also made significant contributions, evidenced by their solid linkages and citation counts, suggesting their influence on the ongoing discourse around fake news. Castillo Camacho and Wang (2021) earlier publication maintains relevance, showcasing foundational concepts in the field. The more recent papers by (Hanshal et al. 2023)d Shu et al. (2020a, b, c) have gained attention, signifying their current impact on fake news research. Technological aspects are addressed by Devlin et al. (2018), while (Barnabò et al. 2023) recent work is part of ongoing discussions, as indicated by its link strength. Mikolov et al. (2013) earlier contribution is foundational, and Pérez-Rosas et al. (2017) work, despite fewer citations, likely offers specific insights into the broader landscape of fake news research. These references form a rich tapestry of knowledge, influencing and shaping the research trajectory in the dynamic domain of fake news (Table 7).

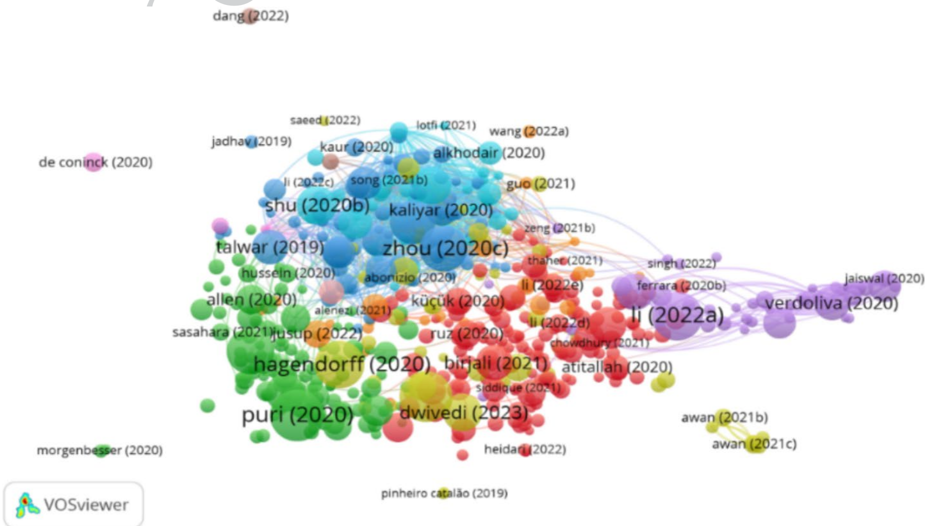

Fig. 4 Visualization's network of fake news research filed; bibliographic coupling analysis by documents (VOSviewer 1.6.20)

4.4 Cluster analysis and research themes

4.4.1 Cluster 1 (green nodes): social media data and datasets for misinformation analysis

According to a comprehensive bibliometric investigation, Cluster 1 is the convergence point for 82 works on various themes. This cluster includes research on fake news, social media, and deep learning for language comprehension. The cluster contains data mining, stylometric analysis, graph-based approaches, and deep learning models. Castillo and Mendoza (2011) significant finding is a 539-link strength. This project aims to create a multidimensional data library for false news investigations. Due to social media disinformation, Castillo emphasizes the need for comprehensive datasets incorporating news content, social context, and location data. The article describes the Fake News Net repository containing two distinct datasets. This resource aids in identifying, evolving, and ablating fake news. The paper also provides a perceptive exploratory review of the datasets, highlighting the benefits and prospective paths for social media misinformation research. Cluster 1's work, "Advances in Fake News Detection, Social Media Analysis, and Deep Learning for Language Understanding." highlights their extent and depth.

4.4.2 Cluster 2 (blue nodes): misinformation dynamics and spread on social media platforms

Cluster 2, with 73 papers, addresses disinformation and fake news, notably on social media. Misinformation is prevalent on social media; therefore, this cluster examines its causes and solutions. The primary paper (Vosoughi et al. 2018) of information on Twitter. This is especially true with fake political news. The study undermines the idea that machines propagate disinformation, demonstrating that people do it more. This research emphasizes the importance of understanding misinformation spreading factors and their possible effects on society. This detailed view encourages more investigation into disinformation dynamics and mitigation methods.

The cluster's various documents and strong linkages help us comprehend disinformation in the changing social media world, especially during critical events like the COVID-19 epidemic and the 2016 US presidential election.

Table 7 Top 10 most influential publications in Fake News Research; bibliographic coupling analysis using (VOSviewe 1.6.20r)

Reference	Links	Total link strength	Citations
vosou	295	1128	214
(Shu et al. 2017)	241	773	100
(Castillo et al., 2011)	214	669	78
(Hanshal et al. 2023)	217	618	67
(Allcott and Gentzkow 2017)	201	583	110
(Kai Shu, Deepak Mahudeswaran, et al., 2020)	236	539	62
(Devlin et al. 2018)	217	523	111
(Barnabò et al. 2023)	202	407	49
(Mikolov et al. 2013)	198	390	85
(Pérez-Rosas et al. 2017)	169	368	40

4.4.3 Cluster 3 (purple nodes): word embeddings and neural language models

Cluster 3, with 63 documents, centers on a significant study by Mikolov et al. (2013), the top paper with 390 links. This landmark study provides two novel model architectures for continuous vector representations of words from large datasets. The research aims to improve word representations by comparing neural network-based methods.

The authors note that existing natural language processing (NLP) systems generally interpret words as atomic units without considering their similarity. The research claims straightforward methods may fail, especially when dealing with restrictions such as limited domain-specific data or scaling up models. The authors argue that machine learning has made training sophisticated models on massive datasets possible, resulting in better performance than simpler models. This top article contributes to Cluster 3's word representation discourse with its solid link strength. It promotes more sophisticated NLP systems and large-scale, complicated models.

4.4.4 Cluster 4 (red nodes): machine learning and deep learning models for fake news detection

Cluster 4, with 61 publications, focuses on creating innovative approaches and algorithms to detect bogus news. This cluster uses advanced machine learning and deep learning methods like neural networks (LSTM, CNN, and BERT). Data mining is also popular, helping extract relevant characteristics and trends from social media and news information. These publications all use benchmark datasets tailored to false news detection research. Cluster 4 models and approaches may be rigorously evaluated using these datasets as a benchmark.

Cluster 4 study emphasizes the importance of contextualizing approaches in the changing social media ecosystem, given the prevalence of fake news. These studies evaluate false news detection algorithms using accuracy, precision, recall, and F1-score. These metrics give a consistent framework for comparison, revealing the robustness and dependability of the models in Cluster 4's different texts. This cluster's research promotes false news detection by demonstrating a dedication to several approaches, contextual factors, and stringent assessment criteria. We still rank (Vosoughi et al. 2018) first in cluster 4.

4.4.5 Cluster 5 (Orange nodes): Social Media Mining and data mining for fake news detection

Cluster 5, with 60 papers, investigates innovative false news-detecting methods. The work "Fake News Detection on Social Media: A Data Mining Perspective." (Hanshal et al. 2023) lead this cluster with a Total Link Strength (TLS) 773. This study discusses the complex issues of detecting false news on social media and using auxiliary information such as user social activities. It highlights the limitations of using content alone to detect false news, highlighting data quality and analytical difficulties from users' social activities.

The top study reviews algorithms, fake news characterization, assessment criteria, and datasets. Its primary purpose is to provide deep insights and future research paths for identifying false news on social media. This pivotal study in Cluster 5 lays the stage for a detailed investigation of methods to identify and battle false news on social media platforms.

4.4.6 Cluster 6 (yellow nodes): information credibility and rumor detection on social media

The leading paper (Castillo and Mendoza 2011), titled "Information Credibility on Twitter," anchors Cluster 6, with 42 publications. This fundamental article examines Twitter news believability. The authors leverage user behavior, post text, and external source citations to identify tweets as trustworthy automatically. Their approach focuses on "trending" microblog posts to understand information credibility dynamics.

The researchers test their approaches with thorough human assessments, reaching 70–80% precision and recall. Beyond methodological issues, the research discusses social media credibility, arguing that social media signals might help consumers evaluate information reliability.

The authors contextualize their results by examining credibility studies in traditional media, blogs, and Twitter news sources. This thorough review begins to explain credibility evaluation across information distribution channels. In addition, the study addresses disinformation and false rumors in the Twitter ecosystem, shedding light on the difficulty of maintaining information accuracy in the dynamic social media scene.

4.4.7 Cluster 7 (Pink nodes): deepfake detection and media forensics

Cluster 7, with 17 publications, explores deepfake detection and the issues of detecting modified or synthetic information. These articles include face modification, forgery detection, video forensics, and deep learning to combat deepfakes. These publications discuss various deepfake detection approaches, benchmark datasets, and deep learning models, notably Generative Adversarial Networks (GANs) and Recurrent Neural Networks (RNNs).

"Very Deep Convolutional Networks for Large-Scale Image Recognition" (Simonyan and Zisserman 2014) has the most excellent Total Link Strength (TLS) in this cluster, suggesting significant influence. While this publication has the greatest TLS, it may not directly relate to Cluster 7's deep fake detection subject. The research focuses on how network depth affects convolutional neural networks (ConvNets) for large-scale image recognition and advances computer vision.

The presented bibliographic coupling analyses comprehensively map the research landscape in fake news and misinformation detection on social media platforms. Identifying clusters and influential publications provides valuable insights into the key research themes, methodologies, and interconnections. This analysis can guide future research efforts by fostering interdisciplinary collaborations and synergies among diverse research communities. As misinformation continues to evolve, the insights gained from this bibliometric study can contribute to developing more effective and comprehensive solutions, ultimately leading to a more informed and resilient digital ecosystem.

4.4.8 Author bibliographic linkage

Author bibliographic linkage, or bibliographic coupling, examines how authors are interconnected through shared references in their works. It quantifies the frequency at which authors cite the same documents, revealing scholarly connections and potential collaborations within a specific field.

In our analysis, authors were selected based on a minimum of one document and 20 citations. From a pool of 13,002 authors, this criterion yielded 13 authors, prioritizing quality over quantity to highlight influential contributors. The top 10 authors were ranked based on their total link strength, which measures the number of shared references between each pair of authors.

Based on their total link strength, the bibliographic coupling analysis reveals the most impactful authors (Table 8). Imran Ashraf (Cluster 2, Yeungnam University, South Korea), with expertise in localization, signal analysis, sentiment classification, and object detection, emerges as the top author with a link strength of 1317. Alexander Gelbukh (Cluster 1, National Polytechnic Institute, Mexico), specializing in stylometric, cyberbullying detection, rumor analysis, and authorship attribution, closely follows with 882. Grigori Sidorov (Cluster 1, National Polytechnic Institute, Mexico), also focused on cyberbullying, sentiment analysis, and NLP tasks, secured the third position with 783.

The bibliographic coupling network visualization in Fig. 5 highlights three distinct clusters of authors based on their shared references. Cluster 1 (Blue) consists mainly of authors working on stylometrics, cyberbullying detection, rumor analysis, and natural language processing tasks related to fake news and disinformation. Cluster 2 (Green) includes authors focused on sentiment classification, object detection, and signal analysis, often applied to social media data or multimedia content. Cluster 3 (Blue) appears minor, with authors contributing to diverse topics like crop modeling, evapotranspiration analysis, and COVID-19 radiological findings.

These rankings and cluster patterns offer valuable insights into author significance within their research communities, reflecting their impact on the field through bibliographical connections and shared references. Authors with higher link strengths are more interconnected and influential within their specific clusters and subdomains of fake news research.

4.5 Bibliographic coupling analysis by source

Bibliographic linkage by source analysis offers insight into the intellectual landscape of fake news research by examining shared references among academic sources. This method, leveraging metrics like total link strength, enables exploring connections between scholarly works. Total link strength, measuring shared references between sources, is the primary criterion, offering advantages such as representativeness, influence assessment, and trend identification. While other criteria like document count and citations can complement the analysis, total link strength remains critical for identifying key players and trends in the field (Van Eck and Waltman 2010). While other criteria like document count and citations can complement the analysis, total link strength remains critical for identifying key players and trends in the field.

The analysis focused on sources in which articles were published between 2017 and 2023, considering a source article threshold and requiring at least 10 citations per article. Consequently, the largest group of related articles comprises 70 sources. After analyzing bibliographic links by source, the most relevant sources were identified based on the highest links and total link strength, as outlined in Table 9.

Table 8 Top 10 most influential authors in fake news research; bibliographic coupling analysis by authors (2017–2023) source (VOSviwer 1.6.20)

Author	Cluster	Links	Total link strength	Affiliation
Ashraf, Imran	2	13	1317	Yeungnam University: The institution will open in a new tab, Gyeongsan, South Korea.
Gelbukh, Alexander	1	13	882	Instituto Politécnico Nacional , Mexico, Mexico.
Sidorov, Grigori	1	13	783	Instituto Politécnico Nacional , Mexico, Mexico.
Rustam, Furqan	2	13	756	University of Management and Technology Lahore The institution Lahore, Pakistan
Umer, Muhammad	2	12	723	The Islamia University of Bahawalpur , Bahawalpur, Pakistan
Zubiaga, Arkaitz	1	12	471	Queen Mary University of London
Damaševičius, Robertas	3	13	444	Vytautas Magnus University, Kaunas, Lithuania
Awan, Mazhar Javed	3	7	403	University of Management and Technology Lahore, Lahore, Pakistan
Rosso, Paolo	1	12	345	Universitat Politècnica de València, Valencia, Spain
Vishwakarma, Dinesh Kumar	1	12	293	Delhi Technological University, New Delhi, India
Ferrara, Emilio	1	11	194	USC Viterbi School of Engineering, Los Angeles, United States
Saeed, Faisal	1	13	151	Birmingham City University, Birmingham, United Kingdom
Gupta, Brij B.	1	11	27	Asia University, Taichung, Taiwan

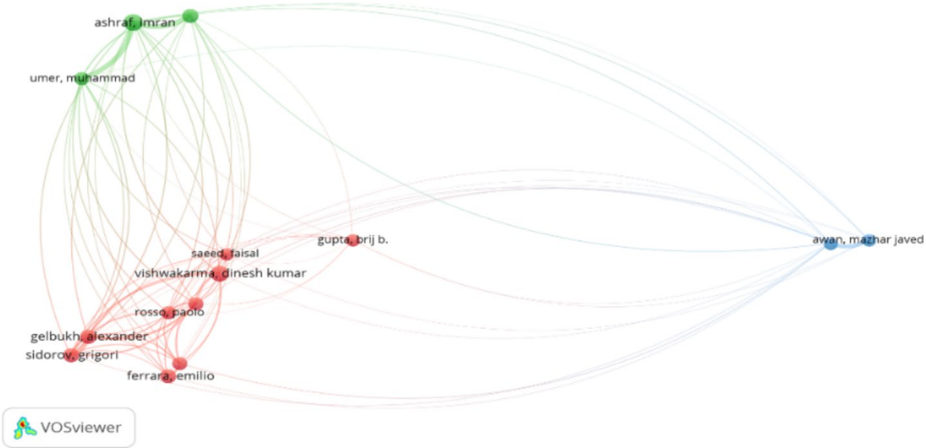

Fig. 5 Visualization's network of fake news research filed; bibliographic coupling by author (2017–2023) Source (VOSviewer 1.6.20)

4.6 Cluster analysis and research themes

A cluster analysis was conducted on academic sources related to the research theme Fig. 6. The study revealed four distinct clusters, each representing a research theme or area of interest within the broader theme of fake news, dissemination, and detection.

4.6.1 Cluster 1 (green nodes): machine learning and deep learning for misinformation detection

This cluster demonstrates a robust publication output, prominently led by Applied Sciences (Switzerland), with 115 documents and an impressive Total Link Strength of 17,483. Information (Switzerland) and Electronics (Switzerland) contribute actively to this cluster, with Total Link Strengths of 14,098 and 0, respectively. The research within this cluster primarily focuses on developing and evaluating machine learning and deep learning models for detecting misinformation, fake news, and rumors on social media platforms. Techniques such as natural language processing, sentiment analysis, and multimodal approaches are explored. Furthermore, investigations into the spread and impact of misinformation during events such as the COVID-19 pandemic are conducted within this cluster.

4.6.2 Cluster 2 (blue nodes): Multimedia and Computational Intelligence approaches

Sources such as Multimedia Tools and Applications, Expert Systems with Applications, and Knowledge-Based Systems contribute significantly to this cluster, with Total Link Strengths of 15,712, 25,982, and 13,143, respectively. The research in this cluster revolves around applying computational intelligence techniques, including neural networks, deep learning, and ensemble methods, to address challenges in misinformation detection. Additionally, multimedia data such as images and videos are considered, and techniques for detecting deepfakes and manipulated media content are explored.

Table 9 Top influential sources by bibliographic coupling (based on total link strength) source (VOSviwer 1.6.20)

Source	Cluster	Total link strength	weights. Documents
Ieee Access	2	55,569	189
Information Processing And Management	3	28,070	73
Expert Systems With Applications	2	25,982	79
Social Network Analysis And Mining	3	20,801	57
Applied Sciences (Switzerland)	1	17,483	115
Multimedia Tools And Applications	2	15,712	60
Plos One	3	15,453	47
Information (Switzerland)	1	14,098	39
Knowledge-Based Systems	2	13,143	54
ACM Computing Surveys	4	12,902	24

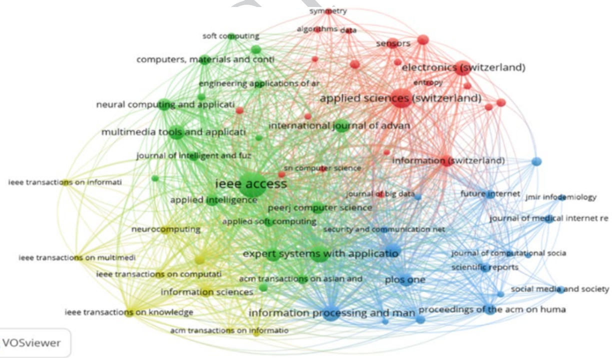

Fig. 6 Visualization's network of fake news research filed; bibliographic coupling by source (VOS Viwer 1.6.20)

4.6.3 Cluster 3 (red nodes): Social Network Analysis and Information Diffusion

This cluster exhibits a relatively high publication rate, with PLOS ONE, Social Network Analysis and Mining, and Information Processing and Management being the notable contributors, with Total Link Strengths of 15,453, 20,801, and 28,070, respectively. The focus here is on analyzing the diffusion patterns and dynamics of misinformation on social networks. Techniques such as network analysis, graph-based models, and user behavior analysis are employed to understand the propagation of rumors and fake news. Moreover, investigations into the influence of bots, coordinated campaigns, and echo chambers on the spread of misinformation are carried out within this cluster.

4.6.4 Cluster 4 (yellow nodes): surveys, reviews, and future directions

As expected for a cluster focused on surveys and reviews, the publication rate seems lower than the other clusters. However, ACM Computing Surveys and IEEE Transactions on Knowledge and Data Engineering are prominent sources within this cluster, with Total Link Strengths of 12,902 and 0, respectively. This cluster primarily consists of survey papers, literature reviews, and studies that critically analyze the state-of-the-art in misinformation detection and identify future research directions. These sources provide a comprehensive overview of the field, highlighting challenges, open issues, and potential solutions.

These clusters offer opportunities for interdisciplinary collaborations that could drive innovative solutions for detecting and mitigating fake news online. For example, integrating AI techniques from the first cluster with social media analysis methods from the third cluster could lead to novel approaches for identifying and counteracting misinformation online. Similarly, combining multimedia processing expertise from the second cluster with social network analysis techniques from the third cluster could facilitate the development of multimodal fake news detection systems.

Funding agencies and policymakers can leverage these insights to align research agendas and priorities with emerging trends, societal concerns, and regional strengths in specific areas of fake news research. Encouraging knowledge dissemination through interdisciplinary conferences, workshops, or special issues in relevant journals could foster collaboration, knowledge sharing, and the exchange of methodologies and techniques among researchers from different domains.

Furthermore, incorporating the identified research themes and influential sources into curriculum development for relevant academic programs could train the next generation of researchers and professionals with a comprehensive understanding of the interdisciplinary nature of fake news research.

Targeted literature reviews within each cluster could identify gaps, challenges, and potential research directions, informing future efforts and addressing the most pressing issues in fake news detection, analysis, and mitigation. Additionally, analyzing the geographical distribution of influential sources and research institutions within each cluster could reveal potential collaboration networks, facilitate international research partnerships, and identify regional strengths or gaps in specific areas of fake news research.

Finally, performing a temporal analysis of bibliographic cluster patterns could shed light on the evolution of research topics, emerging trends, and the impact of societal events or technological advancements on the fake news research landscape.

By leveraging the insights gained from this bibliographic cluster analysis, researchers, funding agencies, and policymakers can foster interdisciplinary collaborations, align research agendas, disseminate knowledge, and drive innovative solutions to combat the growing challenge of fake news and online misinformation.

5 Discussion

This research aims to map the evolution of the Fake News domain comprehensively. We examined the Fake News domain using co-citation analysis from 1986 to 2023 and discussed it using bibliographic linkage from 2016 to 2023. The study's primary limitations are article category and impact factor limitations. The study did not cover all scientific journals and other publications. We believe a future bibliometric examination of scientific conference papers, book chapters, books, and journals not yet indexed will be required without an impact factor. Another difficulty is determining thresholds for bibliometric analysis, such as the number of citations. The main issue is that this may result in omitting potentially fascinating material.

Based on the data, overall visualization, and analysis, we conclude that, over the previous decade, social media (e.g., Facebook, Twitter, Reddit) have become the primary information-sharing and large-scale communication platforms. Researchers believe that these platforms are the primary drivers of the BigData revolution. These sites have also become unmanageable sources of information for their users. The authors have demonstrated that significant political events, like the US elections, Brexit, or pandemics like COVID-19, provide a fertile ground for creating and disseminating misinformation or disinformation.

We also discovered that DeepFake, a deep generative technique for producing and modifying the look of faces, has resulted in a broad spectrum of benign and malicious applications during the last five years. These programs use film visual effects to spread falsehoods about celebrities and political leaders (Galli et al. 2022). The dark side of this new approach presents another problem for researchers: recognizing DeepFake to distinguish fraudulent faces from real ones. With the fast expansion of DeepFake-related studies in the community, both sides (i.e., DeepFake creation and detection) have drawn academic interest, driving them on a new path, which explains the evasion of DeepFake detection research (Shahid et al. 2022).

On the other hand, researchers are trying to find solutions to free and protect society from fake news. However, the challenges are enormous in front of this flood, which is growing by taking advantage of technological progress. According to this bibliographic study, researchers face considerable challenges in reducing fake news damage.

Even if LD-based approaches are more accurate than other methods, they may be more acceptable. The choice of features and classifiers significantly influences the model's performance. Previous research has focused little emphasis on element and classifier selection. Researchers should focus on determining which classifiers are most suited to specific characteristics. Although few research studies have examined this, extended textual features need sequence models (RNNs). We believe that feature and classifier selection research might improve performance.

The notion of feature engineering is unusual in deep learning-based research. The characteristics of news content and headlines are commonly employed in detecting false news. Additional elements such as user behavior (Ding et al. 2020), profile, and social network activity must be investigated. Political or religious biases in profile attributes and lexical, syntactic, and statistical aspects can increase detection rates. One of the critical challenges in the field of fake news categorization is a lack of data on false news. Using CNN, LSTM, and ensemble models, several studies have achieved great accuracy in identifying precision. However, just a few studies employed SeqGAN, and even fewer used the Deep Belief Network (DBN).

Transformers have supplanted RNN models like LSTM as the preferred model for NLP workloads. BERT has been used to detect false news. However, the generative pre-trained transformer (GPT) has yet to be applied. As a result, present algorithms make critical judgments while failing to provide accurate information about the reasons behind the specific decisions, projections, recommendations, or actions (Fauvel et al. 2023). XAI (Explainable Artificial Intelligence) is a branch of study that aims to make AI system findings more intelligible to humans (Adadi and Berrada 2018). The XAI technique is gaining traction and garnering the attention of academics in the field of false news.

Organizations from many areas struggle to discover viable methods for detecting fake information online. False online information is difficult to distinguish since it frequently attempts to fool consumers. New paradigms of information distribution in the mainstream media have emerged due to the digital era. The new digital mass media facilitates user-driven communication; multimedia material is the latest information identity. As a result, the media landscape has changed away from mass media and toward individualized social media. This progress has benefits, but it also can harm society owing to the spread of misinformation and disinformation in the form of "fake news."

In this age of big data and the internet, the proliferation of social media platforms hastened the spread of fake news. Detecting false news on social media is a complex challenge worldwide due to the speed with which rumors circulate, their fragmented content, and their broad reach. In the next part, we use Fake news on social networks as an example, describing the many forms of Fake news, their dissemination techniques, and other detection and prediction solutions given by researchers on this subject.

5.1 Social media's impact on information consumes

In the age of social media, everyone may own their media. The obstacles to sharing text, photographs, or videos, colloquially known as "user-generated content," are vanishing due to freely available publishing systems, low-cost content generation, and smartphones. User-generated content creation is not restricted to social media sites. Traditional media firms have also embraced this progression. Social media erodes people's pluralistic news intake since they only see information from their chosen sources. Content that contradicts their views, on the other hand, may never surface in their stream. This concept has seen various variations, including echo chambers (Kaliyar et al. 2021), balkanization, and bubble filters (Rowland 2011). Despite its importance, factual evidence for this theory is sparse.

The logic behind the echo chamber is divided into two parts. Social media users typically interact with persons who are in training or working. These groups frequently have similar ideals and interests. As a result, most social network transactions occur between people with similar interests. The other component is the way social media organizes material. Every engagement on social media sites reveals something about our personality by commenting, liking, or sharing. Social media use this information to forecast precisely what we want and suggest material that matches our expectations.

Indeed, online behavior is prone to feedback loops in which users value and agree on specific content formats. In turn, they receive more comparable material and less content, contradicting their stance. Fake news is highly linked to echo chambers. Indeed, local societies united by predetermined ideology are unlikely to respond favorably to misleading news, mainly if they include information targeted at demeaning their political opponents. As a result, eliminating pluralistic critique of the entire political scene is a critical technique for combating incorrect news flow.

According to data specialists at Facebook, the diversity of users' friends exceeds preconceived ideas, including political opinions. Other research suggests that the echo chamber notion itself is amplified. Further research indicates that the echo chamber notion is becoming more popular. Indeed, social media users and search engines consume more diverse news and information than those who do not utilize such services (Kaliyar et al. 2020).

Social media have influenced news consumption since its inception; according to a study by Hasebrink and Hölig (2020), In every country, slightly more than a quarter (28%) of those polled prefer to start their information journey with a website or an app. The 18-24-year-olds (dubbed "Generation Z") have an even weaker connection with websites and apps and are more than twice as likely to access news through social media.

5.2 Fake news in the COVID-19 era

The emergence of COVID-19 has resulted in a return to traditional media consumption patterns reminiscent of the mass media period (Yang and Pan 2021). The State Presidents gave multiple talks and live broadcasts. Additionally, experts have returned to the stage. In addition to politicians, senior advisers engage in formal press conferences. Chris Whitty, Chief Medical Officer, and Sir Patrick Vallance, Chief Scientific Adviser, have become household names in the United Kingdom. Anthony Fauci, head of the National Institute of Allergy and Infectious Diseases, has become well-known in the United States (Shearer and Mitchell 2021).

However, considering the predictive informational environment, the COVID-19 pandemic is fertile for fake news. From the theory that 5G internet was the real cause of the epidemic to Trump thinking bleach injection could be a solution. Those rumors have resulted in lawsuits against mobile antennas; as soon as the World Health Organization authorized the COVID-19 vaccines, an increase in anti-vaccination movements and whispers of conspiracy theories began (Amith and Tao 2018; Hayawi et al. 2022; Weinzierl and Harabagiu 2021).

According to the CCHR-commissioned study report, around one in every six Britons refuses vaccination against SARS-CoV-2, while a comparable number still hesitates. According to the report, persons who get their knowledge about the pandemic from social media are less likely to receive the vaccine. According to the report, 31 million people follow anti-vaccine groups on Facebook. It also implies that the anti-vaccine campaign may generate $1 billion in income for social media corporations yearly. $989 million might be earned on Facebook and Instagram, owing to advertising that targets the 38 to 7 million followers of anti-vaccine pages (Burki 2020).

The COVID-19 crisis has put evidence-based technocratic policymaking into doubt, paving the way for populist political (Gradoń et al. 2021) views inspired by conspiracy theories and mistrust (Khan et al. 2021). This crisis highlighted the complexity of truth in science and policymaking and its simplistic expression in political discourse. Politicians have often claimed that they follow scientific recommendations when imposing restrictions on the public, which is a rational strategy to disassociate themselves from unpopular policy decisions. However, science's uncertainty and rapid evolution have created controversy among experts. This fact has rarely been reported, even by the mainstream media.

Finally, COVID-19 also impacted the 2020 US presidential election, which Donald Trump lost. Donald Trump plunged throughout his campaign and administration, spreading voter fraud, misinformation, and spoofing the absentee vote (Preston et al. 2021; Zhang & Ghorbani, 2020). Most adult Americans report having seen false information about COVID-19, and many have seen claims about alleged cures. However, relatively few Americans are confident in their ability to verify the accuracy of the information they find about the coronavirus. Around three in ten US adults (28%) express high confidence in the accuracy of the information on the coronavirus epidemic. Smaller proportions (22%) report feeling unsafe, while for the majority, nearly half (49%), the description is somewhat accurate (Hölig & Hasebrink 2020).

However, social media differs from traditional media regardless of its echo chamber effect. While social media fights misleading content on its platforms, including elections and COVID-19, Americans continue to use it for information—about half of US adults (53%) report seeking information through social media. Today, social media attracts a large audience seeking information (30%) from Facebook and (31%) from Twitter or joining journalists (10%) from Facebook and (24%) from Twitter. While (54%) on Facebook and (41%) on Twitter say they have discovered trending news content through the sharing of their friends within their networks (B. Y. A. Mitchell et al., 2021). Fake news is a severe problem in several places throughout the world. The reasons for their rise are even more problematic in light of the COVID-19 problem and the political orientations in this context, which have generated uncertainty. Fake news and the policies that go with it continue to pose concerns.

5.3 Deep fake, the newborn of the fake news family

Synthetic media technology is progressing rapidly, making creating non-verifiable media with an increasingly realistic appearance and sound easier. The “deep fakes” (because they relied on deep learning) frequently depict someone saying or doing something she has not said or done. The proliferation of “deep fakes” poses a new challenge to the trustworthiness of visual experiences and has already resulted in negative consequences such as consensual pornography, political disinformation, and financial fraud (Galli et al. 2022). Deepfakes can harm viewers by deceiving or intimidating them and their subjects, jeopardizing their reputation and society, and undermining societal values such as trust in institutions.

Indeed, for decades, it has been possible to convincingly alter digital photos and videos with visual effects, except that recent breakthroughs in deep learning increase the realism of artificial materials and how easily we can create them (Rana et al., 2022; Shahid et al. 2022). This AI-generated material, dubbed “deep fakes,” falls into one of three categories:

Face-swapping involves changing a person’s visage in a video with another person. This method has allowed the insertion of celebrities into video footage that never appeared. (De Carvalho et al., 2013) And produce non-consensual pornography where another person replaces a person’s likeness in an original film. (Ismail et al., 2021; Malik et al., 2022)

- Lip sync Consists of editing a video source to match an arbitrary audio recording to the mouth region. For example, actor and director Jordan Peele presents a particularly striking example of this type of media, editing a video of President Obama to say statements such as “President Trump is a total and complete dip-**** ;(Malik et al., 2022).

The puppeteer, an artist, sits in front of a camera and acts out what he wants his puppet to say and do by animating a target person (head, eye, and facial expressions).

While these methods have pleasant and non-malicious applications, some concerns focus on their potential militarization (Vaccari & Chadwick, 2020). For example, there has been a disturbing increase in the catastrophic effects of misinformation in recent years, ranging from violence against our citizens to election meddling (Vaccari & Chadwick, 2020; Wu et al., 2019). Fake, sophisticated, convincing videos will likely make misinformation tactics even more deadly.

Although we are just four years into the deep fake age, it is necessary to speculate freely about the future of this prevalent kind of digital manipulation. Early in 1994, Tom Cruise recently became the victim of several high-profile, deep fake films that proliferated on TokTok and YouTube. These videos, created by Chris Ume in collaboration with Tom Cruise impersonator Miles Fisher, are clever and compelling. Still, they also necessitate a significant amount of work, time, and skill and demonstrate how deep fakes' discourse has yet to catch up with the reality of their creation (Wu et al., 2019).

Indeed, the terminology is unclear: is deep fake one word or two? Deepfakes have followed a creative path that began with sexist face substitution in pornography, then cascaded into a public arena of joyful experimentation, where their applications have flourished. Such applications are possible through remarkable advances in machine learning code and passionate software sharing on open-source platforms. The free flow of this technology rests on strong democratic ideals. Still, unlike the rise of the World Wide Web, no utopian preachers have come forward to champion the common good of open-source image publishing.

Suppose the dominating reactions to deep fakes have been motivated by fear, with research focusing on detection and regulation. May we expect a brief period of free access and use, which will be closed - if feasible - by political authorities? Or will this reactionary tendency be tempered by economic interests seeking to capitalize on the numerous commercial opportunities created by deep fakes?

5.4 Fake news, an illustration of disinformation

"Fake news" generally refers to false and verifiable information that may mislead readers. This section outlines using social media data to describe, detect, and attribute fake news.

5.5 Characterization

Persuading individuals and persuading them to reject other people, groups, or future ideas or actions; arousing emotions such as fear, anger, or excitement about a person, group, concept, nation, or simply future activities to gain support or opposition; educating, for example, about the dangers of vaccination; preventing belief in degrading or illegal behavior; emphasizing the seriousness of certain words or actions; confusing historical occurrences with activities, as in the famous example of the US landing at the lunar surface or simply in a desert on Earth.

Explain the significance of using the internet to engage with the public and demonstrate the importance of identifying misleading information on social platforms, such as Elizabeth Warren and Mark Zuckerberg (Giansiracusa, 2021; Montesi, 2020a). End-to-end models enrich causal relationships between claims and evidence to detect intent, such as persuasive influence. (Hidey & McKeown, 2016). Once we know the meaning of a misleading news article, we can better understand its success, i.e., the likelihood that the misleading article will achieve its goals. Virality measures based on social theories can help characterize this.

In social psychology, the factors of social influence (the degree of dissemination of the news article) and personal impact (the user's pre-existing knowledge) are reliable proxies for the spread of misinformation (Zhou et al. 2019). They increased social and emotional influences on a customer's behavior and perception, encouraging them to trust a news piece and promote its spread. Social media analysis allows researchers to assess the societal influence on individuals' attitudes and/or opinions toward misleading and false information. Perception and behavior encourage them to trust a news piece and promote its spread. Social network analysis enables researchers to evaluate how social influence influences individuals' attitudes and/or opinions about misleading and false information (Zhou et al. 2019). A social dynamic arises that supports misinformation propagation by considering a whole information ecosystem rather than individual consumption patterns. In line with social hemophilia theory, social media use tends to follow people with the same affinities as their friends. As a result, it receives news that promotes its existing stories, causing an echo chamber phenomenon. Treating propagation networks in a hierarchical structure allows for a finer-grained analysis, including a macro-level, such as display and transfer, and a micro-level, such as response (Shu, Mahudeswaran, Wang, & LiShu et al. 2020a, b, c). thus illustrating the structural and temporal characteristics of hierarchical information diffusion networks statistically differing between disinformation and correct information. (Shu et al. 2019). This approach could complement a strictly intent-based characterization approach, emphasizing the importance of misinformation having immediate negative consequences, such as comic relief, when disseminated with positive intent (Shu, Mahudeswaran, Wang, & LiShu et al. 2020a, b, c).

5.6 Detection

Detection aims to identify fake news quickly and efficiently based on explanatory variables. As fake news spreads false claims, the easiest way to remember it is to examine the sincerity of the central claims in a news story to assess its authenticity. For traditional media, detecting fake news depends mainly on the ability to explore the content of the information. Content can take many forms, such as text, images, or videos. Research focuses on learning features from individual or mixed modalities to build machine learning models for fake news detection. In addition to elements directly related to the content of news stories, social context-related features will be gathered from the customer's social engagement when reading news on social media (Sahoo & Gupta, 2021; Ulizia et al., 2021).

Social interactions are evidence of the spread of information over time and provide additional information critical to authenticity. There are three main components to consider in social media: the users, the content created, and the network. Initially, fake news is probably made and distributed via automated accounts and cyborgs. Consequently, recording user-profiles and behavior through user-based features can provide valuable information for spotting fake news. Therefore, registering profiles and behavior from user-based elements can give helpful information for spotting fake news. (Sahoo & Gupta, 2021). Second, online consumers convey their feelings or ideas regarding false news via social media content, including skepticism and sentimentality.

Consequently, extracting message-based attributes to identify probable false news from the general public's responses to the messages becomes meaningful. Finally, social media consumers from different networks discuss their preferences, themes, and relationships. Moreover, false news distribution processes establish an echo chamber cycle, underscoring the importance of network-based feature extraction in detecting fake news (Ruchansky et al., 2017; Sahoo & Gupta, 2021; Shu et al. 2020).

Fake news frequently includes text, images, and videos, among other media types. Consequently, utilizing multimodal data has the potential to enhance detection. First, work has focused on extracting linguistic features, including lexical features, lexicon, emotion, and readability for classification or defining neural-linguistic characteristics through neural network structures, including convolutional neural network (CNN). And recurrent neural networks (RNN). Second, critical points are primarily derived from statistical visual characteristics, visual content capabilities, and neural visual features (Cheng et al., 2021).

Furthermore, newer research focuses on building visual scene graphs from photographs to find common-sense information. (Razniewski et al., 2021), Which substantially improves graphs of scenes issued from visual content. Z.

5.7 Attribution

The objective of attribution is to validate purported sources or producers and to accompany attribution evidence. Due to the absence of a central authority or process for storing and verifying social media data sources, the search for attribution is an increasing issue. Defining the origin of network dispersion includes finding a collection of critical nodes to improve information transmission.

Attribution entails confirming the purported sources or suppliers and the related proof for attribution. Attribution in social media is a new issue due to the lack of a centralized authority or system for keeping and verifying the source of data from social media. From the standpoint of network dispersion, determining the origin is locating a key group of nodes to optimize information transmission (Shu, Bernard, et al., 2019).

Provenance routes show information flows among sources and other hubs, including intermediary nodes. Social traits can pinpoint information's origin (Feng et al., 2013). Following hypotheses of degree propensity and proximity propaganda (Azadegan & Teich, 2010), Nearby and central nodes are more likely to be forwarded to these nodes. Thus, we consider the optimal senders via graph optimization based on a given set of possible source nodes. In addition, new methods that integrate information across core networks, such as node properties and temporal data, are also being developed to improve source discovery (de Rosa & Papa, 2021; Paschen, 2020).

Developing deep learning and intense generative models significantly increased the likelihood that machine-generated texts would produce clear, legible, and engaging false news, resulting in new modalities of attribution. The researchers present a variety of language generation models that rely on adversarial learning, SeqGAN (M.-Y. Chen et al., 2021; Yan et al., 2021), ScratchGAN (Wu et al., 2019), MaskGAN (Fedus et al., 2018), RankGAN (Lin et al., 2017), RL-GAN (Wu et al., 2019), and LeakGAN (Guo et al., 2018), and unsupervised models based on Transformer (Vaswani et al., 2017) for language generation, such as Grover (Zellers et al. 2019) and GPT-2 (Ang, n.d.).

Considering synthetic texts created by machines and the ways offered to differentiate the models used to generate these texts is a big challenge. Classifying various text creation method data and investigating the decision limits is feasible. The collected information can be derived from representative language generation models such as Grover, SeqGAN, VAE, TextGAN, GPT-2, MaliGAN, etc. Moreover, meta-learning enables us to anticipate new sources of text generation based on specific learning instances. In addition, certain generative models, such as the Ctrl model, have limitations; SentiGAN (Shirish Keskar et al.,

2019) and PPLM (Dathathri et al., 2019) Allow the production and encoding of specific text kinds, such as appealing or dynamic text. Reducing misleading correlations in the prediction model, such as decoupling style characteristics from synthetic text via reverse learning and constructing prediction models that can recover transferable features across various text generation algorithms, is crucial. (Paschen, 2020).

Using a social network, individuals may connect and interact with anybody, anytime, from anywhere. It is becoming increasingly crucial for large-scale information dissemination to use SOCIAL MEDIA. Due to their convenience and accessibility, social media attracts an increasing number of users, who consider these networks as a source to search and receive news and information. In 2018, about 68% of adults in the U.S. received their news via social media (E. S. and A. Mitchell, 2021). However, social media is also a source of misinformation and disinformation, including fake news. During the 2016 U.S. election, 8,711,000 comments, reactions, and shares of the most discussed fake news on Facebook compared to 7,367,000 most concerned factual information(E. S. and A. Mitchell, 2021).

End-user contributions to creating information, such as news articles, posts, and comments, and the sharing and recommending of new items through social media entail implicit user evaluations of the content and aid in identifying false news. However, compared to traditional data, Social Media data is crowded, incomplete, messy, deconstructed, and rich in social affiliation. Social networks possess limited tagged data but have unique characteristics that lend themselves to creating weak supervision, leading to an emerging form of weak supervision, namely weak social supervision (WSS). This new data type demands new computational analytical approaches combining social theories with statistical data mining tools (Apuke & Omar, 2021; Shu et al. 2019).

Due to the increased use and ease of social media, more people seek and receive news information online. Even as social media combats misinformation spread across its platforms related to elections, the COVID-19 pandemic disease, and many other topics, many Americans continue to rely on these sites for their information. Approximately (53%) of U.S. adults are told that social media was their "frequent" or "occasional" source of news from August 31 to September 7, 2020 (SHEARER 2021). Also, according to (Tandoc, 2021), 47.2% of Singaporean social media users reported getting their news regularly or very frequently through Facebook, compared to 37.8% from local television channels, 37.2% from local online newspapers, and 36% via local newspapers.

Identifying fake news on social networks has become critical to preventing its consumption and creating a healthy and trustworthy information ecosystem. Even though social networks have little labeled data, they possess unique attributes conducive to creating weak supervision, resulting in an emerging form of inadequate control, i.e., weak social supervision.

5.8 Disinformation by WSS

Humans are incapable of distinguishing erroneous information from other types of information. Several cognitive theories explain this phenomenon, including naïve realism and confirmation bias. Misinformation exploits the customer's susceptibility as an emerging consumer to approach them. Given these cognitive biases, certain forms of misinformation, including " fake news, "are often passed off as genuine. Interdisciplinary research studies focus on human exposure to "fake news" to develop more effective detection systems (Preston et al. 2021).

Various cognitive theories, such as realism and naïve confirmation bias, explain this phenomenon. Fake news primarily targets new customers by exploiting their inherent weaknesses. This sensitivity of people to "fake news" is the topic of interdisciplinary study, the results of which serve as the basis for developing more effective detection systems (D'Ulizia et al. 2021).

Different strategies for identifying transmission based on several forms of WSS, such as sources, have been developed to understand better how misinformation and false news affect social media. These techniques include accountability, trust, global perspective, and intent(Abbasi and Liu 2013; Ang et al. 2021); the social group targeted by its bias, population, stature, and worldview (Shu et al. 2018a, b); the characteristics of the linguistic, visual, contextual content, emotional tone, density, duration, and cohesion (Shu et al. 2018a; Zhou et al. 2019) as well as the nature of the relationship with their network such as cohesion and points of separation (Shu, Bernard, et al., 2019). It is possible to quantify the impact of these theories by measuring user metadata (Abbasi and Liu 2013), which could answer the question "Why are individuals susceptible to fake news? "or "Do particular populations of people show more vulnerability to certain forms of fake news? " (Shu et al. 2018).

According to Social Theories such as the Acceptance Preference Theory, social identification and recognition are essential to a user's identity, driving people to choose "socially trusted" news consumption and broadcasting alternatives. According to social homophily theory, social media users who follow and befriend like-minded others get material that reinforces their current narratives, resulting in an echo chamber effect. One quantitative poll is essential for determining whether and how well theories predict people's responses to fake news (Kirchner and Reuter 2020).

Kai Shu, Deepak Mahudeswaran, et al. (2020) reveal that the architectural and chronological insights inside hierarchical dissemination networks might influence the consumption of false news, suggesting that more sources of weak social supervision are advantageous in combatting fake news (Kai Shu, Ahmed Hassan Awadallah, et al., 2020). In hierarchical dissemination, propagation networks are handled for deep study, with macro levels as post, repost, and microlevels reacting to spread networks. There are considerable differences in the structural, temporal, and linguistic aspects of hierarchical transmission networks between Fake and Real news.

5.9 WSS for tracking misinformation

Tracking fake news and misinformation raises specific challenges, making identifying them more challenging. Fake news and misinformation content are diverse in subject matter, style, and media platforms, while fake news disguises the truth through various linguistic types while mocking real news. Therefore, labeling fake news data cannot be improved, and particular integration approaches are inadequate for correctly identifying fake news with annotated data. Consequently, enhancing the collection of annotated false news data is impractical, and specific integration approaches are insufficient for accurately identifying fake news with minimal labeled data. As challenges, disinformation and fake news continue to emerge. Indeed, the propagation of false information is becoming an issue related to significant occurrences (pandemics, elections, natural disasters, war). Learning with little social supervision enables the identification of disinformation and false news in complex settings using detection algorithms that are practical, intelligible, and predictive. The findings of these algorithms offer a method for detecting false information and insight into future outcomes (Kai Shu, Ahmed Hassan Awadallah, et al., 2020).

5.10 Effectiveness of disinformation detection

The objective is to use inadequate social supervision as evidence to identify deception efficiently. Specifically, such interaction networks allow the visualization of the many entities and their interactions to identify misinformation while disseminating information. Data is a significant obstacle because fake news and misinformation content are diverse in topics, styles, and media platforms. Fake news seeks to twist the truth with various linguistic types while mocking factual information. Thus, obtaining observations on fake news remains difficult, while specific integration methodologies cannot detect fake news with limited labeled data. Then, interaction networks depict relationships between entities such as editors, information, and users. By simulating their interaction, interaction networks aim to include multiple concepts linked with entities inside the same latency space. In this manner, the Tri-relation framework (TriFN) for fake news represents information characteristics to identify false information (Kai Shu, Ahmed Hassan Awadallah, et al., 2020; Shu, Wang et al. 2019). Sociology and cognitive theory underpin weak social surveillance norms. These ideas have emphasized findings that suggest the ineffectiveness of social monitoring since social scientists remind us that users with similar tastes and interests are more likely to engage with one another. Users with low credibility are thus more inclined to propagate disinformation, while those with high credit ratings are less likely to do so. Consider the following example for the publishing relationship; we explore the subsequent weak social supervision: publishers with high medical or political bias scores will be more susceptible to disseminating disinformation.

Furthermore, users with low credibility scores become significantly more vulnerable to diffuse disinformation for a broadcast relationship, while those with high credibility scores become less susceptible to propagating fake news. Utilizing non-negative matrix factorization is one method for learning new representations under poor social supervision (NMF). TriFN can achieve an accuracy of 0.89 for misinformation detection in experiments on real-world datasets (Shu, Wang et al. 2019).

5.11 Disinformation recognition thorough explanation

Extracting the most explicable terms from material and user comments is a component of explanation-based misinformation detection. This approach improves the detection process's accuracy and simplifies the findings, particularly for end-users unfamiliar with machine learning. According to Shu, Wang et al. (2019), most news content phrases are accurate, yet they merely bolster fraudulent statement sentences. Therefore, not all news phrases may be equally relevant for identifying and explaining whether or not a news report is untrue. User comments may also provide information regarding crucial reasons why a news report is inaccurate, even though they may be uninformative and disruptive. They used Weak Social Supervision, which involves users' comments linked to original news content, which is valuable in detecting fake news and providing insight into predictive results. Therefore, weak social surveillance can guarantee that user comments connected to the original news content contribute to discovering false news and explaining the prediction findings.

Cui et al. (2019) present dEFEND's explicable method for detecting fake news. This model has four components: a news content coder, a comment user coder, a sentence-comment co-attention element, and a false news predictor. According to their article, a news document includes linguistic indicators at various levels, particularly word and sentence levels, that give variable degrees of explanation on whether the news is fabricated. Consequently, he suggests learning representations of news content through a hierarchical structure. Therefore, they propose a model that starts by understanding the news content through a hierarchical structure. First, they learned the sentence vectors by carefully using the word encoder, and then they learned the sentence representations through the sentence encoder component.

Based on the use of bidirectional LSTMs to learn the encoding of sentences, words, and comments, preceded by a neural network named sentence-comment co-attention. Their experiment on real-world datasets showed that the framework outperforms 5.33% in F1 scores from the previous methods of fake news detection presented by Zhou and Zafarani (2019). Finally, integrating exploratory features such as news decision network visualization and related news and trends by the DEFEND provides an intuitive news presentation and improves the system's interactivity.

5.12 Precocious fake news detection

Early detection of misinformation, such as fake news, often involves recent and time-sensitive events. Also, identifying misinformation in advance requires limited information from user interactions, as high user engagement implies that the misinformation has already influenced more users. Social media are multidimensional, suggesting diverse and varied interactions between information and its disseminators.

Social network content is multidimensional, revealing heterogeneous and multiple connections between news and its broadcasters on these networks. Users' comments and publications offer a wealth of crowdsourced information, such as views, opinions, and emotions, allowing for identifying fake news. Moreover, consumers' degrees of trust differ. According to Shu et al. (2017), Less reputable users are more likely to propagate false information. These results from social media can provide more insight into the correct prediction of false news. Thus, many sources of weak social monitoring produced from social media may be used, and information retrieved to detect false news early.

In fact, during the model-building phase, social context information is employed to create weak rules to get low-labeled instances and restricted self-labeling to promote training. During the prediction stage, understanding information content independent of social involvement is sufficient to swiftly identify false news given any news item in the test set. Therefore, we may use a deep neural network structure in which the network's bottom layers learn the common feature representations of the news items. The network's top layers independently model the mappings of feature representations to each supervisory source. Numerous Weak Social Supervision Sources (MWSS) is an architecture that permits combining multiple WSS sources and clean labels, as shown in Fig. 7. The following elements are sentiments, partiality, and credibility to extract the correct tags.

The MWSS framework aims to learn from the social network based on weak multiple supervision to detect misinformation at an early stage.

(a) In Training: The MWSS conjointly learns proper labels from various weak sources;
(b) MWSS applies the learned representation features and fc function to the prediction labels of (unseen) occurrences identified in testing data during inference.

Research has shown that current events presenting contradictory opinions or feelings are more likely fake news (Jin et al. 2016). sentiments would be more susceptible to constitute fake news. Also, users' opinions on fake news appear more polarized and less neutral (Cui, Wang et al. 2019). After the news has been distributed based on each user's sentiment ratings, the variation and despair of the sentiment scores are computed using their variance. Consequently, the labeling function is poor: when the standard deviation of user sentiment ratings exceeds a threshold of τ1, a news article is weakly classified as Fake News. Social research theorizes news publishers' biased association with news veracity (Zellers et al. 2019). To a certain degree, news post users may represent publishers, and their intolerance for false news against actual news differs significantly (Shu et al. 2018). Therefore, people who are more predisposed toward information are more willing to distribute fake news, whereas users who are less emotionally invested are more likely to share authentic news.

By mining users' interests against their tweet history, Kulshrestha et al. (2017) measure the biased kernels of users. The bias result is between −1 and 1, where −1 represents a leftward bias, and +1 represents a rightward bias. Thus, we get the weak labeling function, according to which content is weakly labeled as Fake News if users' average absolute bias rating exceeds the threshold τ2.

According to research, less reputable individuals, such as malevolent accounts or ordinary people susceptible to fake news, are more apt to disseminate false information (Shu et al. 2017; Shu, Wang et al. 2019). A credibility score signifies the quality of the user's reli-

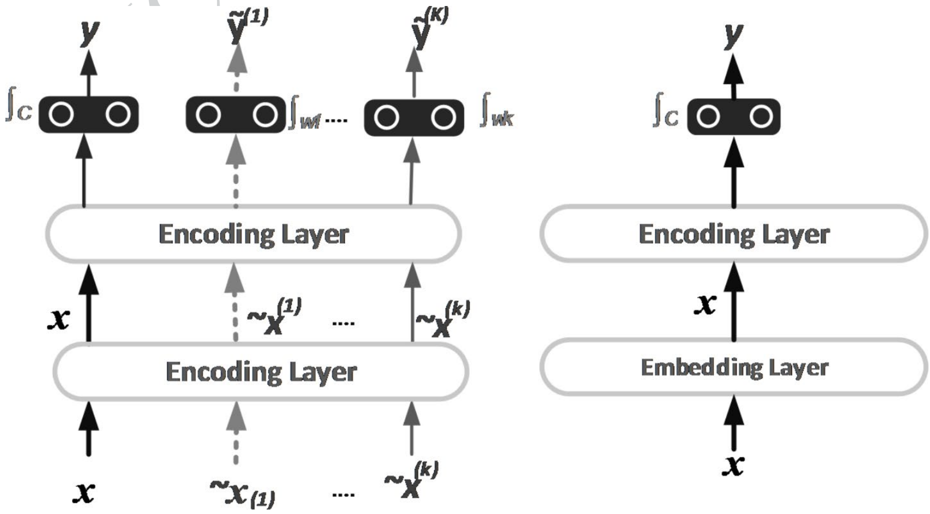

Fig. 7 The MWSS framework to learn from social network data based on weak supervision to catch early misinformation

ability. In their paper Kai Shu, Suhang Wang, et al. (2020), researchers chose the practical approach of measuring the user's credit scores.

Inspired by the idea (Abbasi and Liu 2013), Low-credibility users tend to coordinate and create large groups, while higher-credibility users target small groups. It is important to note that the credibility of social network users comes from their content publication history. Thus, the label-only function: If a piece of news has a mean credibility score under a threshold of $\tau3$, it is labeled as Fake News.

5.13 Analysis of fake news content and context

The quick spread of fake news is so harmful that academics have determined to automate it using machine learning techniques such as deep neural networks (DNN). However, dependability is overshadowed by the black box problem - a lack of transparency in DNN decision-making addressed the "black box problem" of deep learning to identify fake news. They generated a dataset of 24,000 items, 12,000 of which were false and 12,000 of which were genuine. The fake stories were sourced from Kaggle, but the authentic ones were from The Guardian and The New York Times. According to the study's conclusions, DNNs may be used to identify language trends in fake news. Additionally, the approach may detect fake news in creative subject areas.

CSI (capture, score, and integrate) is a further method for tackling the "black box problem" of deep learning in fake news identification. It is a three-step strategy that encompasses the three defining qualities of false news (Zhou et al. 2017). These features include language, source, and user-provided clarifications to fill in missing data. In the first stage, a recurrent neural network (RNN) is used to gather the transitory pattern of user action. The second phase determines the source of suspicions about user behavior. The third hybrid step combines stages 1 and 2 and forecasts fake articles.

Experiments using real-world datasets have shown that predicting false news items is highly accurate. On the other hand, the lack of a manually labeled fake news dataset is a bottleneck for utilizing a computationally demanding model. In their paper, Yang et al. (2012) picked a publicly accessible contemporary dataset named LIAR to solve the restricted availability of labeled datasets and combat false news using statistical methodologies. Linguistic models enabled the use of this dataset to study fake news. Findings rely on a comparison of many methods, including logistic regression (L.R.), Support vector machines, convolutional neural network (CNN) models, and long-term memory networks (LSTM) (SVM). Lastly, combining facts and language considerably improves the identification of fake news. In addition, this data gathering can be used to identify rumors, classify positions, and conduct the study's research modeling, argumentation extraction, and political natural language processing (NLP). Table 10 summarizes the several directions recommended for counting and assessing the content and context of fake news. In the inaugural phase of the 2017 False News Challenge (FNC), which intended to apply artificial intelligence (AI) approaches to tackle the problem of fake news, contestants used location detection. Positional detection is the relative location of two significant pieces of text on an issue, statement, or topic. The opinion of other organizations regarding a specific problem, statement, or topic consists of two relevant text elements. The scoring method used had two levels: a 25% weighting assigned according to whether the text was considered related or unrelated to its title, and

Table 10 Summary of studies conducted in bots, Clickbait, and rumors

Study	Approach	Dataset	Finding
Benevenuto et al. (2010)	This study examines Twitter spam detection. A massive list of tagged people was manually divided into spammers and non-spammers from tweets relating to three significant 2009 news stories (more than 54 million users, 1.9 billion links, and about 1.8 billion tweets).	54 million user accounts, 1.9 billion connections, and 1.8 billion tweets	Characterizing the researchers' tagged collection, users showed various spammer-non-spammer characteristics. Characterization study develops spam detection. They misclassified several legitimate users but detected many spammers utilizing categorization.
Becker et al. (2011)	The study examines Twitter message stream analysis to detect real-world and non-real-world tweets. The authors use a broad family of aggregate statistics from thematically related communications.	25 000 Users, 500 000 Tweets, 49 million Users, followers/ friends	The research identified most spammers but misidentified a few buyers. Results may accurately identify diverse spammers. Spammer classification fared competitively despite having more features and resisting spammers. A web crawler-powered network study showed 89% spam.
Zhao et al. (2015)	The authors uncovered rumors by uniquely clustering tweets with investigative patterns (signal tweets), extracting the claim from each cluster, and using the claim to find the rest of the litigation-related non-signal tweets. We rank clusters by statistically comparing signal tweets to the set—great early performance.	10,240,066 tweets find by keywords	Clustering and classification perform effectively, and earlier detection of rumors.
Biyani et al. (2016)	The study analyzes a web page's content, informality, and header-body similarity from its body, title, and URLs. It identified clickbait articles.	1,349 click-baited and 2,724 non-clickable sites from various sources on Yahoo's homepage.	Definition of 8 types of clickbait -Exaggeration -Teasing -Inflammatory -Formatting -Graphic -Bait-and-switch -Ambiguous -Wrong
Shu et al. (2017)	The authors presented the TI-CNN technique to evaluate pictures and text for false news analysis by combining explicit and latent characteristics into a single feature space.	Publicly available datasets on BuzzFeedNews15. LIAR1 BS Detector CREDBANK	The authors introduced the TI-CNN method for assessing images and text for fake news analysis by merging explicit and latent features into a unified feature space.
Reis et al. (2019)	A machine learning method for detecting false news by analyzing article properties, including source. Assess the prediction performance of the strategies to construct an auto-detection system.	**-2282** BuzzFeed news stories about the U.S. election.	They detect fake news from media-related material and user profile content on social media platforms. KNN and SVM are the two classification algorithms used.
Lemos et al. (2021)	This article explores fake news and clickbait related to the 2019 northeastern Brazil oil spill. My examination of the 2019 Brazilian oil crisis videos linked YouTube deception, the environmental effect of digital media, and algorithmic performance.	- 6335 News articles from news and social media sites.	The M.N. performed worse than LSVC and TF-IDF. To predict, the M.N. likes integer features. Weighted term vectors (TF-IDF vectors) perform better in linear classification. BoW model provided both classifiers with the same accuracy. Bi-gram model accuracy is higher with the M.N. classifier.

Table 10 (continued)

Study	Approach	Dataset	Finding
Sahoo and Gupta (2020)	The researchers suggested ML and DL classifiers for Facebook account features and speech behavior.	> 15,000 Facebook users have shared fake and inaccurate news.	The recommended method detects false news on Facebook using user content and online media sites.
Hakak et al. (2021)	Ensemble classification is recommended for false news identification. Google Colab studies ISOT and Liar, which are major false news databases. Data is classified in an ensemble machine learning model using a decision tree classifier, Random Forest, and additional tree classifier (E.T.).	ISOT and Liar were examined by Google Inc.'s open-source cloud service, Google Colab.	Bagging pooled findings, and the model outperformed state-of-the-art. Accurate feature selection and hyperparameter tweaking improved performance most. The liar and ISOT datasets, the model was 44.15% and 100% correct.
Stevens et al. (2022)	This work uses an augmentation model to find errors in COVID-19 [RD, R.I., R.S., S.R.]. The study also introduces graph neural networks (CGN, GAT, GraphSAGE) (SAmple and aggreGatE).	10,700 Social media posts and their etiquette fall under the classification.	The scientists observed that their false news detection GCN, GAT, and GraphSAGE algorithms improved steadily.
Soetekouw and Angelopoulos (2022)	An online experiment (N=417) examines whether educating social media users to recognize fake news can improve their ability to do so, taking into account skepticism, age, and education.	636 social media users participated in this experiment	The findings demonstrate a strong association between the training procedure and the capacity of social media users to detect false news, implying that the protocol may have a good role in educating social media users to recognize fake news.
(Barnabò et al. 2023)	This study aims to reduce the burden of manual fact-checking while maintaining performance. Active learning strategies with Graph Neural Networks for misinformation detection	Benchmark datasets	Evaluated benchmark datasets Proposed Deep Error Sampling (DES) with uncertainty sampling DES performs equally or better than common active learning strategies Reduces labeling effort by up to 25% without compromising classification performance

a 75% weighting given according to whether the associated pairs identified as agreeing, disagreeing, discussing, or irrelevant.

The winning team in this competition provided a deep convolution neural network (DCNN) and a graded boosting decision tree (GBDT) composite model with a weighted average of 50/50 (Boroumand and Fridrich 2018). Separately, neither the DCNN nor the GBDT had perfect accuracy. The combination of the two approaches, with a score of 82.01, successfully identified the location of each stock. Similarly, the strategy provided by the Athene team (Boroumand and Fridrich 2018) received an 81.97 and placed second in the competition. An MLP classifier and a 100-unit cached layer are used. A RELU (Rectified Linear Unit) activating process allows the layer to be cached, while a soft-max enables the final linear layer to be obtained. The achieved accuracy is 81.72%. In WSDN 2019 (Web Search and Data Mining), another competition detected fake news by classifying article titles. Given the headline "A" of a fake news story and the title "B" of an incoming article, participants were asked to rank "B" in one of three categories: agree, disagree, or irrelevant

(Boroumand and Fridrich 2018). The winning team of this challenge achieved 88.298% weighted accuracy on private leaderboards and 88.098% on public leaderboards (Pham 2019). This ensemble method incorporated neural networks and gradient-based boost trees.

Additionally, the Bidirectional Encoder Representation from the Transformer (BERT) system encoded, converted, and merged news headline pairs into a new symbolic environment. Second place was granted to Liu et al. (2019) for presenting a novel ensemble framework based on the natural language interference (NLI) test. Their approach is a three-stage architecture with a 25-element BERT model, an ensemble mixing method at the first layer, six ML models, and a single L.R. for the definitive classification (Yang et al. 2020). also considered this problem an NLI task and considered both the NLI model and BERT. The BERT and NLI models were also applied by An and Liu (2020) to this issue as an NLI task. Dense CNN, Dense RNN, Gate CNN, and ESIM are the most reliable of the NLI models. Decomposable Attention was trained, and the accuracy obtained was 88.063%.

6 Method against deep fakes

Images and video forensics literature from AI are new (An and Liu 2020). Therefore, few forensic methods find profound forgeries. Rare forensic methods identify profound forgeries. One claim is that first-generation fraud victims did not blink as promised (An and Liu 2020). The generated faces didn't usually match closed-eyelid people. After this forensic strategy was revealed, subsequent generations of synthesis techniques included the blink, making it less effective. An and Liu (2020) detect counterfeit facial identities using 3D head posture estimation differences from the complete face vs. the core (possibly swapped) face area. Face swaps are found, but not lip-synching or puppet simulations. Synthetic pixel artifacts are used in other forensic approaches (Brock et al. 2018; Iuliani et al. 2018; Shetty et al. 2018). Like pixel-based approaches, these can accurately identify many fakes but are sensitive to whitening countermeasures (e.g., additive noise, recompression, resizing).

In "Protecting World Leaders Against Deep Fakes," experts think individuals communicate using unique (but likely non-unique) facial expressions and emotions, as shown in Fig. 8. From one video clip, they estimate action unit strength by recognizing face and head gestures (Brock et al. 2018). Next, they develop a one-class SVM novelty detection model Wang et al. (2019) to distinguish hilarious and profound impersonators. The paper offers OpenFace2, an open-source facial behavior analysis toolkit (Farid 2016). Video face/head motion extraction. The suite includes 2-D and 3-D face landmarks, head position, gaze, and facial movement. Each video frame has a distinct head, regard, and face animation.

Forensics may identify fake videos with consistent facial expressions. Head movements and facial expressions differentiate real persons from fakes. Video compression, length, and speech context determine reliability.

Existing Deepfake algorithms produce low-resolution pictures that must be enhanced to incorporate the faces in the original video, resulting in apparent artifacts in subsequent movies (Li et al. 2019). Using CNN, they want to distort affine faces to differentiate real and fake photographs. Each CNN variant—VGG16, ResNet152, ResNet101, and ResNet50—is trained using training data. Face regions are discovered and extracted using dlib. Matching faces at different levels and blurring using a Gaussian kernel (5 5). Deepfake's output pipe-

(a) (b) (c)

Fig. 8 For four POIs, below are frame samples for (**a**) a real, (**b**) a comedy impersonator, and (**c**) a face change

line deforms the smoothed face to check the original dimensions and reproduce the items. Their model processes training data to save time and money.

Nguyen et al. (2019) examined numerous Deepfake image and video creation and identification technologies. FakeApp, created by OSN Reddit, uses DL-based autoencoder and decoder matching to make deepfake movies. Two independent decoders, A and B, learn and reconstruct latent properties from two facial photos, A and B, using the same autoencoder. By decoding the joint encoder's latent properties, decoder B rebuilds picture B from image A. DFaker, DeepFaceLab, Faceswap, and TensorFlow-based Deepfake-tf employ fake APP-like matching methods.

Before the 2020 U.S. elections, Twitter and Facebook announced plans to erase deepfakes, but disruptions may weaken their detection methods. In particular, adversarial perturbations on source and target faces create deep fakes (Gandhi and Jain 2020). Using CNN and RNN, Montserrat et al. (2020) studied automated face zone weighting and raising algorithms to identify face change in videos. DFDC and ImageNet underpin transfer learning. This three-step technique uses MTCNN to detect faces and extract CNN. Final prediction estimates came from AFW and GRU layers. Nguyen and Derakhshani (2020) used four deep-learning models to discriminate face exchange using biometric eyebrow comparison (LightCNN, Resnet-50, DenseNet-121, and SqueezeNet). After scaling, rotating, and cropping brow photos, they trained and tested their models on the VISOB and CelebDF datasets to match the input conditions for several deep-learning models.

Social verification Tursman et al. (2020) use PCA and hierarchical clustering to assess film facial geometry consistency. LipGAN varied the mouth based on voice variations to create a multi-view dataset with 25 speakers and Deepfakes videos for each original movie. They identified Deepfakes with 75% accuracy using a single mouth, DWT, and PCA+grouping. The mouth signal prevents their method from working when the mouth is tiny relative to the stream. EER of 20.7% and AUC of 0.879% on Celeb-DF were higher than current values. Rashid et al. (2019) used CNN photos to Deepfakes, Face2Face, FaceForensics++, and FaceSwap using DenseNet, Alignment, DenseNet+BiDir, ResNet50, and ResNet50. They discovered Deepfake video flows using reference point alignment and the spatial transformation network and developed the efficient DenseNet+Alignment+BiDir method for movie face alteration detection. Blockchain-based Ethereum intelligent contract architecture verified media (Hasan & Salah, 2019). Their approach maintained digital author and publisher history. Interplanetary file systems collect decentralized media and data (IPFS). Video reliability is checked using technology.

7 Limitations and future directions

7.1 Current state and challenges of fake news research

Nowadays, disinformation campaigns have become more varied and diversified, based on tools and techniques developed at the strongest. In moments of crisis, it can turn to more sinister methods, such as applying false and emotional stories or distorted interpretations of events to affect the behaviors of targeted audiences. We live in complex situations that are becoming increasingly complex every day. Besides the old influence techniques, we must face new strategies based on digital technologies, which are evolving rapidly and becoming increasingly refined, sophisticated, and effective. We observe that many actors with hostile intentions operate in the information environment. Misinformation and fake news constitute new research areas with critical open questions the researcher tries to answer.

Fake news detection via digital technology has yielded promising results in recent decades. Nevertheless, the explicability of fake news detection constitutes a critical element, i.e., the rationale behind spotting a news item as Fake. Recent approaches attempt to generate explanatory factors from user comments (Cui et al. 2019) and natural language claims in an open domain web (Popat et al., 2018). Modeling alternative types of user involvement, including user profiles, can also enhance explanations. Understanding people's gullibility and disseminating fake news is another critical task. However, one way to address this problem is to adopt a causal discovery perspective using directed acyclic graph inferences and then estimate variants of users' processing and dissemination behaviors.

Fake news is a severe social media problem amplified by deep learning models' ability to produce fake neural news (Zellers et al. 2019). The recent advances in neural fake news generation allow malicious users to generate fake news based on limited information. The Generative Adversarial Network (GAN) is a model that can cause long readable texts based on noise (Guo et al., 2018), while the GPT-2 can produce short stories and fictional novels from a primary context (Radford et al., 2019).

7.2 Methodological challenges and validity concerns

In conducting this comprehensive bibliometric analysis, several potential threats to validity must be acknowledged and addressed:

7.2.1 External validity

- The focus on academic literature indexed in Scopus and WoS may limit generalizability.
- Language restrictions (English-only) may exclude relevant research in other languages.
- The time period selected may not capture all relevant historical developments.

7.2.2 Internal validity

- Database selection biases may affect comprehensiveness.
- Search term selection may have missed relevant publications using different terminology.
- The deduplication process may have inadvertently removed unique entries.

7.2.3 Construct validity

- Bibliometric indicators may not fully capture all aspects of research impact.
- Citation patterns may be influenced by factors unrelated to research quality.
- Co-citation and bibliographic coupling assumptions may not always reflect true intellectual relationships.

7.2.4 Conclusion validity

- Results interpretation may be affected by the rapid evolution of fake news research.
- Cluster analysis results may be sensitive to parameter choices.
- The dynamic nature of social media platforms may affect the longevity of findings.

To address these threats, we implemented several mitigation strategies:

- Rigorous data-cleaning procedures.
- Use of multiple databases to ensure comprehensive coverage.
- Application of systematic validation protocols.
- Transparent documentation of all methodological decisions.
- Conduct sensitivity analyses where appropriate.

7.3 Technical challenges and detection methods

Existing fake news creation methods may not provide style and fact-enriched writing with dynamic/attractive stylings and relevant topics. Understanding neurological false news and its detection difficulties is necessary to spot it. Other studies have focused on identifying human-generated from machine-generated fake/real news or generative neural models' counterfeit news production ability. Others have created neural generation detectors that refine classifiers using the generator's primary control point.

Detecting misinformation and fake news early in the conflict is vital. Various academics have examined feature extraction and machine learning models based on news content and social context for binary classification. In contrast, others have focused on user involvement to predict false news. Qian et al. (2018) propose establishing synthetic users' meetings to detect false news, while Wang et al. (2020) provide an invariant neural network model for learning transferable properties to anticipate fake news proactively. Kai Shu, Suhang Wang, et al. (2020) also examine WSS exploitation possibilities for early "fake news" detection and suggest refining these techniques by training in a few steps with less training data.

7.4 Future research directions

Future research in fake news detection and mitigation must address several key challenges across both technical and methodological dimensions. A critical priority is the enhancement of bibliometric analysis frameworks through improved technical integration. This includes developing standardized citation metrics across databases, implementing automated validation protocols for data consistency, and refining deduplication algorithms that maintain data integrity while preserving essential research insights.

The evolution of methodological approaches represents another crucial frontier, particularly in refining AI models and exploring the potential of emerging technologies like GPT. The integration of Explainable Artificial Intelligence (XAI) techniques will be essential for enhancing the transparency and interpretability of decision-making processes in fake news detection. These advances must be complemented by the development of specialized metrics for interdisciplinary research and enhanced protocols for regional and linguistic coverage analysis.

Social media's continuing influence on information consumption patterns necessitates innovative strategies to address echo chambers while promoting diverse content exposure. This includes investigating methods to break down information silos while respecting user privacy and developing more sophisticated approaches to user behavior analysis. Future studies must also consider ethical implications of these technological integrations, particularly in handling user data and profile information.

The field faces significant challenges in addressing the scarcity of labeled fake news data. Research priorities should focus on exploring novel features beyond textual elements, incorporating user profiles and political bias analysis, and developing more robust detection algorithms. Enhanced collaboration between traditional and social media platforms will be crucial, especially in adapting to younger generations' news consumption preferences.

Ongoing technological advancements, particularly in Deep Fake technology, require proactive research responses. This includes developing advanced detection methods, improving cross-platform verification systems, and establishing more robust information integrity frameworks. Such efforts must balance innovation with practical implementation, ensuring that research advances can be effectively deployed in real-world scenarios.

8 Conclusion

This paper provides a map of the field of fake news development. Using co-citation analysis, we investigated the field of disinformation and specifically fake news from 1986 to 2023 and analyzed the period from 2016 to 2023 through bibliographic linkage. Notable research limitations include selecting research and review publications, using the WOS database alone, and evaluating the influence on citations covered in Scopus and WOS databases.

Therefore, the scope of the research was restricted by article category and impact factor, and it did not include all scientific journals and other publications. In our view, future bibliometric research should include scientific conference papers, book chapters, books, and scientific journals without an impact factor. Another limitation of bibliometric analysis These thresholds can lead to the removal of potentially intriguing articles, which is the main problem. Results of the visualization and in-depth study reveal that over the years, researchers have found that during a crisis or significant event, the prerequisite is to determine the situation and have an accurate and clear view of the information space, which is vital in today's complex information environment. Today's environment is challenging and increasingly complicated. Indeed, we face old influence techniques and innovative methods supported by digital technologies, which continue to gain sophistication and effectiveness. A comprehensive and concerted response to disinformation proves an effective remedy against misinformation.

Fake news and disinformation are becoming a growing concern in the information and media field, and risks are raising the attention of media experts, researchers, and customers. Fake News has grown in diversity and reaches various areas on different media platforms. The high cost of tagging often makes it unrealistic for a real-world fake news detection system to obtain much-tagged data for each domain (e.g., entertainment and politics are two different domains). Hence, fake news detection is generally carried out and supervised in a single-domain setting. Unsupervised methodologies deal with restricted or unlabeled fields. The performance, however, is constrained mainly due to excessive fitting over small labeled or unsupervised patterns. In addition, models learned in one domain may be biased and perform poorly in a different target domain. One solution to this problem is to use domain adaptation techniques to explore auxiliary information to transfer knowledge from the source domain to the target domain. In addition, advanced machine learning strategies, such as adversarial learning, provide insight into the representation of invariant subject features to identify emerging false information.

Author contributions All the authors contributed to the design and development of the study. Abdelkebir SAHID, Maleh Yassine, and Karim Ouazane prepared the material, data collection, and analysis. Abdelkebir Sahid wrote the manuscript's first draft, and all authors commented on the previous versions. All authors have read and approved the final manuscript.

Funding No funds, grants, or other support have been received for this study.

Data availability No Data associated with the manuscript.

Declarations

Competing interests The authors have no competing interests to declare concerning the content of this article.

Ethical approval Not applicable.

References

Abbasi, M.-A., Liu, H.: Measuring user Credibility in Social Media. Springer (2013)
Adadi, A., Berrada, M.: Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). IEEE Access. **6**, 52138–52160 (2018)
Allcott, H., Gentzkow, M.: Social media and fake news in the 2016 election [Review]. J. Economic Perspect. **31**(2), 211–236 (2017). https://doi.org/10.1257/jep.31.2.211
Amith, M., Tao, C.: Representing vaccine misinformation using ontologies. J. BIOMEDICAL Semant. **9**, 1–13 (2018)
An, F., Liu, Z.: Facial expression recognition algorithm based on parameter adaptive initialization of CNN and LSTM. Visual Comput. **36**, 483–498 (2020)
Ang, B., Anwar, N.D., Jayakumar, S.: Disinformation & fake news: Meanings, present, future. Disinformation fake news, 3–20. (2021)
Archambault, É., Campbell, D., Gingras, Y., Larivière, V.: Comparing bibliometric statistics obtained from the web of Science and Scopus. J. Am. Soc. Inform. Sci. Technol. **60**(7), 1320–1326 (2009)
Barnabò, G., Siciliano, F., Castillo, C., Leonardi, S., Nakov, P., San Martino, D., G., Silvestri, F.: Deep active learning for misinformation detection using geometric deep learning [Article]. Online Social Networks Media. **33**, 100244 (2023). https://doi.org/10.1016/j.osnem.2023.100244
Becker, H., Naaman, M., Gravano, L.: *Beyond trending topics: Real-world event identification on twitter*. (2011)
Benevenuto, F., Magno, G., Rodrigues, T., Almeida, V.: *Detecting spammers on twitter*. (2010)
Biyani, P., Tsioutsiouliklis, K., Blackmer, J.: 8 amazing secrets for getting more clicks: Detecting clickbaits in news streams using article informality. (2016)
Boroumand, M., Fridrich, J.: Deep learning for detecting processing history of images. Electron. Imaging. **2018**, 211–213 (2018)
Boyack, K.W., Klavans, R.: Co-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately? J. Am. Soc. Inform. Sci. Technol. **61**(12), 2389–2404 (2010)
Bradford, S.C.: Sources of information on specific subjects. Engineering. **137**, 85–86 (1934)

Brock, A., Donahue, J., Simonyan, K.: Large scale GAN training for high fidelity natural image synthesis. *arXiv preprint arXiv:1809.11096*. (2018)

Burki, T.: The online anti-vaccine movement in the age of COVID-19. Lancet Digit. Health. **2**(10), e504–e505 (2020)

Castillo, M., Mendoza, M.: Poblete,. *Information Credibility on Twitter, Information Credibility, WWW.* (2011)

Castillo Camacho, I., Wang, K.: A comprehensive review of deep-learning-based methods for image forensics [Review]. J. Imaging. **7**(4) (2021). Article 69 https://doi.org/10.3390/jimaging7040069

Černe, M., Kaše, R., Škerlavaj, M.: Non-technological innovation research: Evaluating the intellectual structure and prospects of an emerging field. Scand. J. Manag. **32**, 69–85 (2016)

Chen, C.: Science mapping: a systematic review of the literature. *Journal of data and information science, 2*. (2017)

Cui, L., Shu, K., Wang, S., Lee, D., Liu, H.: *defend: A system for explainable fake news detection*. (2019a)

Cui, L., Wang, S., Lee, D.: *Same: sentiment-aware multi-modal embedding for detecting fake news*. (2019b)

Culnan, M.J.: The intellectual development of management information systems, 1972–1982: A co-citation analysis. Manage. Sci. **32**(2), 156–172 (1986)

D'Ulizia, A., Caschera, M.C., Ferri, F., Grifoni, P.: Fake news detection: A survey of evaluation datasets [Article]. PeerJ Comput. Sci. **7**, 1–34 (2021). https://doi.org/10.7717/PEERJ-CS.518

Derouiche, K., Frunza, M.: Study of Tweets' Sentiment Impact on Stock Prices during Class Actions: An Application to Sports Companies. *Available at SSRN 3653125*. (2020)

Devlin, J., Chang, M.-W., Lee, K., Toutanova, K.: Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*. (2018)

Ding, Y., Chowdhury, G.G., Foo, S.: Bibliometric cartography of information retrieval research by using co-word analysis. Inf. Process. Manag. **37**(6), 817–842 (2001)

Ding, J., Hu, Y., Chang, H.: *BERT-based mental model, a better fake neAws detector*. (2020)

Farid, H.: Photo Forensics. The MIT Press (2016)

Fauvel, K., Chen, F., Rossi, D.: A lightweight, efficient and explainable-by-design convolutional neural network for internet traffic classification. Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, (2023)

Ferreira, F.A.F.: Mapping the field of arts-based management: Bibliographic coupling and co-citation analyses. J. Bus. Res. **85**, 348–357 (2018)

Galli, A., Masciari, E., Moscato, V., Sperlí, G.: A comprehensive Benchmark for fake news detection. J. Intell. Inform. Syst. **59**(1), 237–261 (2022)

Gandhi, A., Jain, S.: Adversarial Perturbations fool Deepfake Detectors. IEEE (2020)

Garfield, E.: From the science of science to Scientometrics visualizing the history of science with HistCite software. J. Informetrics. **3**, 173–179 (2009)

Gmür, M.: Co-citation analysis and the search for invisible colleges: A methodological evaluation. Scientometrics. **57**(1), 27–57 (2003)

Gorraiz, J., Melero-Fuentes, D., Gumpenberger, C., Valderrama-Zurián, J.-C.: Availability of digital object identifiers (DOIs) in web of Science and Scopus. J. Informetrics. **10**(1), 98–109 (2016)

Gradoń, K.T., Hołyst, J.A., Moy, W.R., Sienkiewicz, J., Suchecki, K.: Countering misinformation: A multidisciplinary approach. Big Data Soc. **8** (2021). https://doi.org/10.1177/20539517211013848

Hakak, S., Alazab, M., Khan, S., Gadekallu, T.R., Maddikunta, P.K.R., Khan, W.Z.: An ensemble machine learning approach through effective feature extraction to classify fake news [Article]. Future Generation Comput. Syst. **117**, 47–58 (2021). https://doi.org/10.1016/j.future.2020.11.022

Hanshal, O.A., Ucan, O.N., Sanjalawe, Y.K.: Hybrid deep learning model for automatic fake news detection [Article]. Appl. Nanosci. (Switzerland). **13**(4), 2957–2967 (2023). https://doi.org/10.1007/s13204-021-02330-4

Hasebrink, U., Hölig, S.: Audience-based indicators for news media performance: A conceptual framework and findings from Germany. Media Communication. **8**(3), 293–303 (2020)

Hayawi, K., Shahriar, S., Serhani, M.A., Taleb, I., Mathew, S.S.: ANTi-Vax: A novel Twitter dataset for COVID-19 vaccine misinformation detection. Public. Health. **203**, 23–30 (2022)

Hota, P.K., Subramanian, B., Narayanamurthy, G.: Mapping the intellectual structure of social entrepreneurship research: A citation/co-citation analysis. J. Bus. Ethics. **166**(1), 89–114 (2020)

Insights, C.: Attention Spans. In: Microsoft Canada, Spring (2015) AQ2

Iuliani, M., Shullani, D., Fontani, M., Meucci, S., Piva, A.: A video forensic framework for the unsupervised analysis of MP4-like file container. IEEE Trans. Inf. Forensics Secur. **14**, 635–645 (2018)

Jin, Z., Cao, J., Zhang, Y., Luo, J.: *News verification by exploiting conflicting social viewpoints in microblogs*. (2016)

Kaliyar, R.K., Goswami, A., Narang, P., Sinha, S.: FNDNet–a deep convolutional neural network for fake news detection. Cogn. Syst. Res. **61**, 32–44 (2020)

Kaliyar, R.K., Goswami, A., Narang, P.: EchoFakeD: Improving fake news detection in social media with an efficient deep neural network. Neural Comput. Appl. **33**, 8597–8613 (2021)
Kessler, M.M.: Bibliographic coupling between scientific papers. Am. Doc. **14**, 10–25 (1963)
Khan, T., Michalas, A., Akhunzada, A.: Fake news outbreak 2021: Can we stop the viral spread? [Review]. J. Netw. Comput. Appl. **190** (2021). Article 103112 https://doi.org/10.1016/j.jnca.2021.103112
Kirchner, J., Reuter, C.: Countering fake news: A comparison of possible solutions regarding user acceptance and effectiveness. Proc. ACM Hum Comput Interact. **4**(CSCW2), 1–27 (2020)
Kulshrestha, J., Eslami, M., Messias, J., Zafar, M.B., Ghosh, S., Gummadi, K.P., Karahalios, K.: Quantifying search bias: Investigating sources of bias for political searches in social media. Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing, (2017)
Lemos, A.L.M., Bitencourt, E.C., dos Santos, J.G.B.: Fake news as fake politics: The digital materialities of YouTube misinformation videos about Brazilian oil spill catastrophe. Media Cult. Soc. **43**, 886–905 (2021)
Leydesdorff, L.: Theories of citation? Scientometrics. **43**, 5–25 (1998)
Li, H., Li, M., Liu, Q.: Hybrid beamforming with dynamic subarrays and low-resolution PSs for mmWave MU-MISO systems. IEEE Trans. Commun. **68**, 602–614 (2019)
Liu, Y., Liu, Z., Jia, R.: DeepPF: A deep learning based architecture for Metro passenger flow prediction. Transp. Res. Part. C: Emerg. Technol. **101**, 18–34 (2019)
Mas-Tur, A., Roig-Tierno, N., Sarin, S., Haon, C., Sego, T., Belkhouja, M., Porter, A., Merigó, J.M.: Co-citation, bibliographic coupling and leading authors, institutions and countries in the 50 years of Technological forecasting and Social Change. Technol. Forecast. Soc. Chang. **165**, 120487 (2021). https://doi.org/10.1016/j.techfore.2020.120487
McCain, K.W.: Mapping authors in intellectual space: A technical overview. J. Am. Soc. Inform. Sci. (1986–1998). **41**(6), 433 (1990)
Mikolov, T., Chen, K., Corrado, G., Dean, J.: Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*. (2013)
Mining disinformation and fake news: Concepts, methods, and recent advancements, Disinformation, Misinformation, and Fake News in Social Media 1–19 (Springer 2020)
Mongeon, P., Paul-Hus, A.: The journal coverage of web of Science and Scopus: A comparative analysis. Scientometrics. **106**, 213–228 (2016)
Montesi, M.: Understanding fake news during the Covid-19 health crisis from the perspective of information behaviour: The case of Spain. J. Librariansh. Inform. Sci. **53**(3), 454–465 (2021)
Montserrat, D.M., Hao, H., Yarlagadda, S.K., Baireddy, S., Shao, R., Horváth, J., Bartusiak, E., Yang, J., Guera, D., Zhu, F.: *Deepfakes detection with automatic face weighting*. (2020)
Nguyen, H.M., Derakhshani, R.: Eyebrow Recognition for Identifying Deepfake Videos. IEEE (2020)
Nguyen, H.H., Yamagishi, J., Echizen, I.: Capsule-forensics: Using Capsule Networks to Detect Forged Images and Videos. IEEE (2019)
Pérez-Rosas, V., Kleinberg, B., Lefevre, A., Mihalcea, R.: Automatic detection of fake news. *arXiv preprint arXiv:1708.07104*. (2017)
Pham, L.: *Transferring, transforming, ensembling: the novel formula of identifying fake news*. (2019)
Phan Tan, L.: Mapping the social entrepreneurship research: Bibliographic coupling, co-citation and co-word analyses. Cogent Bus. Manage. **8**, 1896885 (2021)
Preston, S., Anderson, A., Robertson, D.J., Shephard, M.P., Huhe, N.: Detecting fake news on Facebook: The role of emotional intelligence. PLOS ONE. **16**(3) (2021). Article e0246757 https://doi.org/10.1371/journal.pone.0246757
Reis, J.C.S., Correia, A., Murai, F., Veloso, A., Benevenuto, F.: Supervised learning for fake news detection. IEEE. Intell. Syst. **34**, 76–81 (2019). https://doi.org/10.1109/MIS.2019.2899143
Rowland, F.: The filter bubble: What the internet is hiding from you. Portal: Libr. Acad. **11**(4), 1009–1011 (2011)
Sahoo, S.R., Gupta, B.B.: Classification of spammer and nonspammer content in online social network using genetic algorithm-based feature selection. Enterp. Inform. Syst. **14**, 710–736 (2020). https://doi.org/10.1080/17517575.2020.1712742
Sánchez, A.D., Del Río, M.C., García, J.: Bibliometric analysis of publications on wine tourism in the databases Scopus and WoS. Eur. Res. Manage. Bus. Econ. **23**(1), 8–15 (2017)
Sarkis-Onofre, R., Catalá-López, F., Aromataris, E., Lockwood, C.: How to properly use the PRISMA Statement. Syst. Reviews. **10**, 1–3 (2021)
Shahid, W., Li, Y., Staples, D., Amin, G., Hakak, S., Ghorbani, A.: Are you a cyborg, bot or human?—a survey on detecting fake news spreaders. IEEE Access. **10**, 27069–27083 (2022)
Shearer, E., Mitchell, A.: News use across social media platforms in 2020. (2021)

Shearer, A.M.M.J.J.B.O.E.: *How Americans Navigated the News in 2020: A Tumultuous Year in Review*. P. C. Research. (2021). https://www.pewresearch.org/journalism/2021/02/22/how-americans-navigated-the-news-in-2020-a-tumultuous-year-in-review/
Shetty, R.R., Fritz, M., Schiele, B.: Adversarial scene editing: Automatic object removal from weak supervision. *Advances in Neural Information Processing Systems*, *31*. (2018)
Shu, K., Sliva, A., Wang, S., Tang, J., Liu, H.: Fake news detection on social media: A data mining perspective. ACM SIGKDD Explorations Newsl. **19**, 22–36 (2017)
Shu, K., Mahudeswaran, D., Wang, S., Lee, D., Liu, H.: Fakenewsnet: A data repository with news content, social context and spatialtemporal information for studying fake news on social media. *arXiv preprint arXiv:1809.01286*. (2018a)
Shu, K., Wang, S., Liu, H.: Understanding user Profiles on Social Media for fake news Detection. IEEE (2018b)
Shu, K., Wang, S., Liu, H.: *Beyond news contents: The role of social context for fake news detection*. (2019)
Shu, K., Awadallah, A.H., Dumais, S., Liu, H.: Detecting fake news with weak Social Supervision. IEEE INTELLIGENT SYSTEMS (2020a)
Shu, K., Mahudeswaran, D., Wang, S., Liu, H.: *Hierarchical propagation networks for fake news detection: Investigation and exploitation*. (2020b)
Shu, K., Mahudeswaran, D., Wang, S.H., 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. **8**, 171–188 (2020c). https://doi.org/10.1089/big.2020.0062
Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*. (2014)
Sługocki, W., Sowa, B.: Disinformation as a threat to national security on the example of the COVID-19 pandemic. Secur. Def. Q., **35**(3). (2021)
Soetekouw, L., Angelopoulos, S.: Digital Resilience through Training protocols: Learning to identify fake News on Social Media. Inform. Syst. Front. (2022). https://doi.org/10.1007/s10796-021-10240-7
Stevens, H., Rasul, M.E., Oh, Y.J.: Emotions and Incivility in Vaccine Mandate Discourse: Natural Language Processing Insights. *JMIR Infodemiology*, *2*. (2022). https://doi.org/10.2196/37635
Studying fake news via network: analysis: detection and mitigation, Emerging research challenges and opportunities in computational social network analysis and mining 43–65 (Springer 2019)
Surwase, G., Sagar, A., Kademani, B.S., Bhanumurthy, K.: Co-citation analysis: An overview. (2011)
Tools of Disinformation: How Fake News Gets to Deceive, Disinformation and Fake News 35–46: (2021)
Tursman, E., George, M., Kamara, S., Tompkin, J.: *Towards untrusted social video verification to combat deepfakes via face geometry consistency*. (2020)
Van Eck, N., Waltman, L.: Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics. **84**, 523–538 (2010)
Vosoughi, S., Roy, D., Aral, S.: The spread of true and false news online. Science. **359**(6380), 1146–1151 (2018)
Wang, T.-C., Liu, M.-Y., Tao, A., Liu, G., Kautz, J., Catanzaro, B.: Few-shot video-to-video synthesis. *arXiv preprint arXiv:1910.12713*. (2019)
Weinzierl, M.A., Harabagiu, S.M.: Automatic detection of COVID-19 vaccine misinformation with graph link prediction. J. Biomed. Inform. **124**, 103955 (2021)
White, H.: Bibliometrics. In *Annual review on information science and tecWhite, H. (1989). Bibliometrics. Annual Review on Information Science and Technology, 119–186.hnology* (pp. 119–186): Elsevier Science Publishers. (1989)
Yang, J.L., Pan, Y.C.: COVID-19 rumor detection on Social Networks based on content information and user response. Front. Phys. **9**, 763081 (2021). https://doi.org/10.3389/fphy.2021.763081
Yang, C.-Z., Ma, J., Wang, S., Liew, A.W.-C.: Preventing deepfake attacks on speaker authentication by dynamic lip movement analysis. IEEE Trans. Inf. Forensics Secur. **16**, 1841–1854 (2020)
Zellers, R., Holtzman, A., Rashkin, H., Bisk, Y., Farhadi, A., Roesner, F., Choi, Y.: Defending against neural fake news. *arXiv preprint arXiv:1905.12616*. (2019)
Zhao, Z., Resnick, P., Mei, Q.: *Enquiring minds: Early detection of rumors in social media from enquiry posts*. (2015)
Zhou, X., Zafarani, R.: *Fake news detection: An interdisciplinary research*. (2019)
Zhou, P., Han, X., Morariu, V.I., Davis, L.S.: Two-stream Neural Networks for Tampered face Detection. IEEE (2017)
Zhou, X., Zafarani, R., Shu, K., Liu, H.: *Fake news: Fundamental theories, detection strategies and challenges*. (2019)
Zupic, I., Čater, T.: Bibliometric methods in management and organization. Organizational Res. Methods. **18**, 429–472 (2015)

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal:
Article:

AQ1. Reference Griffith & Hart, 1971; Tandoc Jr et al. 2018; Tandoc 2021; Lazer et al. 2018; Pranckute, 2021; Zhu et al., 2021; Wang & Waltman, 2016; Baas et al., 2020; Martin-Martin et al., 2021; Visser et al., 2021; Singh et al., 2021; Huang et al., 2020; Glanzel & Moed, 2013; Zhang & Glanzel, 2017; Waltman, 2016; Wang et al., 2020; Zhou et al., 2020; Li et al., 2021; Chen et al., 2022; Aria & Cuccurullo, 2017; Franceschini et al. 2016; Castillo et al., 2011; Kai Shu, Ahmed Hassan Awadallah, et al., 2020; Kai Shu, Deepak Mahudeswaran, et al., 2020; Zhang & Ghorbani, 2020; Mitchell et al., 2021; De Carvalho et al., 2013; Ismail et al., 2021; Malik et al., 2022; Rana et al., 2022; Vaccari & Chadwick, 2020; Wu et al., 2019; Giansiracusa, 2021; Montesi, 2020a; Hidey & McKeown, 2016; Sahoo & Gupta, 2021; Ulizia et al., 2021; Ruchansky et al., 2017; Cheng et al., 2021; Razniewski et al., 2021; Feng et al., 2013; Azadegan & Teich, 2010; Paschen, 2020; Fedus et al., 2018; Guo et al., 2018; Keskar et al., 2019; Popat et al., 2018; Radford et al., 2019 cited in text but not provided in the reference list. Please provide references in the list or delete these citations.

AQ2. Citation for Reference (Insights 2015) missing in the manuscript. Please check.