

# EXPLORING THE SOCIETAL IMPACT OF FAKE NEWS: A Text Analysis Perspective

**Sangeeta Saha**

Senior Project Assistant, NDLI, Central Library, IIT Kharagpur, West Bengal, India

E-mail: sangeeta3637@gmail.com

**Nirupam Biswas**

Librarian, Agradut Sangha Library, Govt. Sponsored Public Library, West Bengal, India

E-mail: nirupamma91@gmail.com

**Moumita Paul**

Library Assistant, Central Library, Munshi Premchand Mahavidyalaya, West Bengal,  
India

E-mail: moumitapaulp3@gmail.com

***Abstract:** The purpose of this study is to explore the research trends regarding the societal impacts of fake news. A total of 224 bibliographic records were collected from Scopus. Text analysis of abstracts using the Voyant tool was applied to identify major research areas, thematic clusters, and the context of the social impact of fake news. The findings reveal that social media plays a crucial role in spreading fake news, with impacts related to psychological and cognitive issues, public trust, misinformation amplification, and technological vulnerabilities. Major strategies found to combat fake news include rumor management, digital literacy, fake news awareness campaigns, and fact-checking. The study also suggests evaluating the effectiveness of these strategies. The findings provide up-to-date information for policymakers, researchers, and library professionals. Extended research and collaboration are essential for developing more effective solutions to fight against fake news.*

***Keywords:** Misinformation research; Text mining; Keywords in Context analysis; t-SNE visualization; Social media use*

## **1. Introduction:**

The rapid transmission of information via internet-based platforms like social networks has significantly changed how we access news. We can access news in the form of text, audio, video, and images. However, fake news or false information propagates widely through digital media. It spreads speedily without any restrictions and without determining the source (Nistor & Zadobrischi, 2022). Gradoń et al. (2021) mentioned the concepts of disinformation, misinformation, and malinformation, while also discussing the notion of fake news. Researchers have also identified how users recognize different types of fake news, including manipulated content, false headings, and satire (Sahu et al., 2023). Fake news has become a serious issue, particularly in the areas of politics, public health, and major global events (Gradoń et al., 2021). The influence of fake news during

the COVID-19 pandemic is among the most obvious examples (İsmailoğlu, 2022; D. Zhang et al., 2021).

Analyzing research trends and the impact of fake news is essential. This study investigates textual data to explore societal consequences, dimensions of spread, and users' perceptions. It states four objectives as follows:

- To identify publication trends and growth in fake news research.
- To examine the dominant research areas within the literature on fake news.
- To analyze thematic clusters and their interrelationships.
- To understand the societal impact of fake news.

## **2. Literature Review:**

Studies exploring misinformation during the COVID-19 pandemic highlighted various facets of this global issue. İsmailoğlu (2022) identified less credible web sources in misinformation-laden tweets and revealed that negative emotions were more pronounced in such tweets compared to regular ones. Similarly, Zhang et al. (2021) focused on misinformation spread on Sina Weibo, such as claims about smoking preventing COVID-19, often framed using pseudo-scientific arguments, making them difficult for the public to discern. Gradoń et al. (2021) provided a broader perspective on infodemics as a threat to health and security while advocating for multidisciplinary approaches like machine learning, complex networks, and WHO-supported infodemic management strategies. Su et al. (2020) reviewed the challenges of misinformation detection and emphasized misinformation's dynamic and evolving nature across domains, languages, and cultures. Jayasekara (2020) added to this perspective by applying text mining to analyze misinformation trends on social media, underscoring the need for data-driven insights and scholarly contributions to address digital information challenges. Verma and Nayak (2024) investigated public misbeliefs about the Sustainable Development Goals (SDGs) on YouTube. Suggestions were made on strategies for moderating misleading content and formulating policy responses. These studies have explored various aspects of fake news; however, research gaps persist. This study aims to examine the societal impact of fake news, providing valuable insights by uncovering emerging trends and themes through abstract-level analyses.

## **3. Methodology:**

This study uses a text analysis approach within an exploratory research framework to analyze bibliographic data on fake news and its societal impact.

### **3.1 Search Strategy and Data Collection:**

The bibliographic data for this study was collected from Scopus, a widely recognized bibliographic database, using a structured search string. Table 1 presents the components

of the search string. The search was conducted on 25/11/2024, obtaining a total of 224 records.

**Table 1:** Components of search strategy

Components	Search String
Selected Keywords	(TITLE ( "fake news" OR misinformation OR disinformation OR "Fabricated news" OR "false information" OR "pseudo-news" OR "manipulative content" OR "misleading information" OR rumors OR "manipulated content*" ) AND TITLE ( impact OR effect OR influence OR consequence ) AND TITLE ( social* OR society* OR psycholog* ) )
Period	2004-2024 (Note: 2024 is an ongoing year, more studies could be found in the future)
Language	English
Document Type	Articles, Chapters, Conference Papers

### 3.2 Data Analysis:

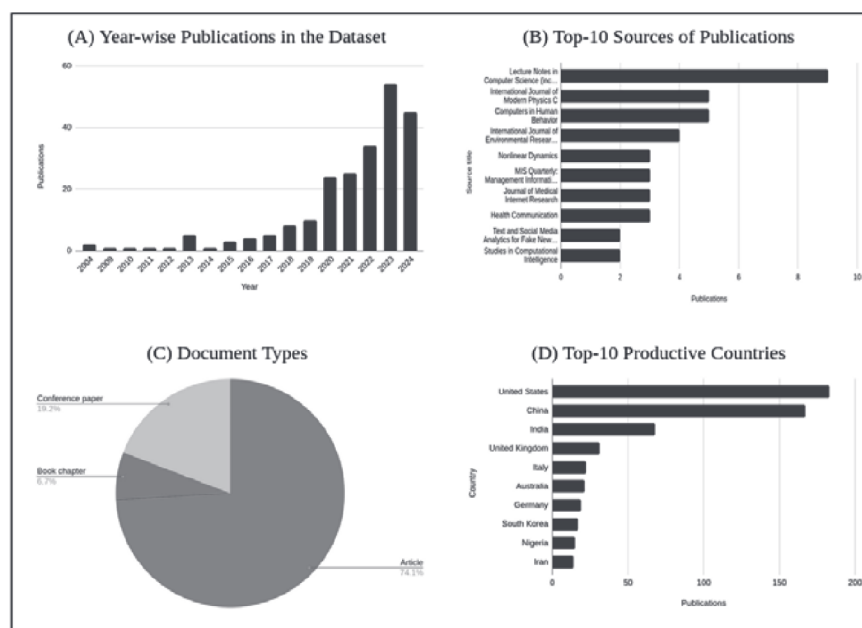
The data analysis shows publication trends, top keywords, dominant themes, and clusters within the dataset. Trend analysis highlighted the growth of research over time, while keyword and thematic analyses revealed major research areas. Cluster analysis uncovered hidden patterns and relationships, providing insights into the progression and key themes of fake news research. Google Sheets was used for primary quantitative data analysis and visualization. The abstracts of 222 records (two records lacked abstracts) were used as the primary corpus for text analysis. The analysis was conducted using the Voyant Tool (Sinclair & Rockwell, 2020). The Voyant Tool automatically performed text preprocessing, such as removing stopwords, converting text to lowercase, and tokenizing. Abstracts for text mining align with prior studies' methodologies (e.g., Mazumder & Barui, 2024), which utilized abstracts as concise representations of research content.

## 4. Findings and Analysis:

### 4.1 Research Trends:

Based on the dataset used in this study, publications show increase over the years, with a rise from 2019 (n= 10) onward (Figure 1 (A)). The surge is observed in 2023 (n=54), followed by 2024 (n=45), while earlier years (2004–2013) had low research outputs. This indicates a growing research focus in recent years. This study found a total of 177 unique sources of publications. Figure 1 (B) showcases only the top ten sources of publications. The most frequent source is Lecture Notes in Computer Science (n=9), followed by International Journal of Modern Physics C and Computers in Human Behavior (n=5

each). Other sources include the International Journal of Environmental Research and Public Health (n=4) and several with 3 publications, such as MIS Quarterly and Nonlinear Dynamics. It reflects the multidisciplinary research approach. Types of documents are shown in Figure 1 (C). The majority of publications are articles (n=166, 74.1%), followed by conference papers (n=43, 19.2%) and book chapters (n=15, 6.7%). It can be said that there is a strong preference for journal articles as the primary mode of scholarly communication. Country-level research production reveals (Figure 1 (D)) that the United States (n=183) has the most number of publications. China (n=167) and India (n=68) have a good number of publications. Other significant contributors include the United Kingdom (n=31), Italy (n=22), Australia (n=21), Germany (n=19), South Korea (n=17), Nigeria (n=15), and Iran (n=14) This particularly suggests a substantial global distribution of research efforts.



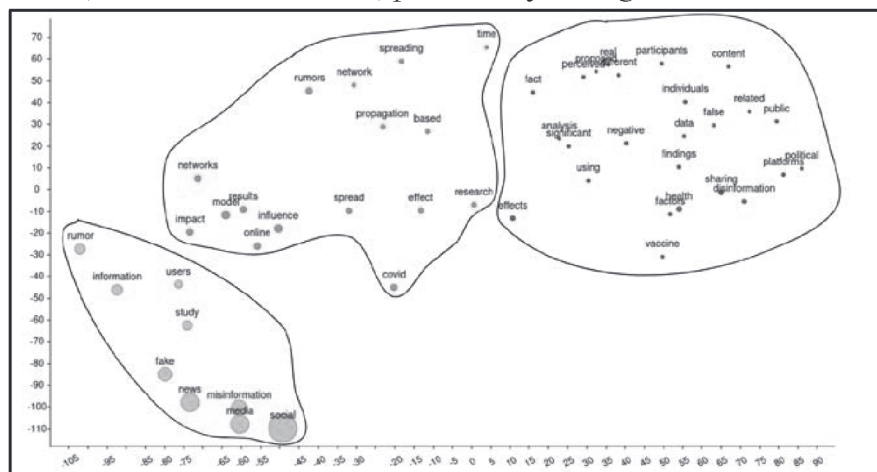
**Figure 1:** Research Trends: (A) Year-wise Publications in the Dataset, (B) Top-10 Sources of Publications, (C) Document Types, and (D) Top-10 Productive Countries

#### 4.2 Major Research Areas:

This study explores the dynamics of fake news and its propagation, user engagement, and societal consequences. Figure 2 is a combination of subplots. First, Figure 2 (A) highlights “social,” “news,” “media,” “misinformation,” “fake,” and “information,” reflecting a research focus on aspects related to digital platforms and the spread of fake news or information. Other terms such as “rumor,” “networks,” “impact,” and “study” signify an exploration of how fake news spreads and the mechanisms behind it. Second, the network visualization (Figure 2 (B)) exposes the interconnectedness of terms like “news,” “media,” “social,” “fake,” and “misinformation.” Secondary terms such as “platforms,” “users,” “propagation,” and “information” point how misinformation is shared and consumed on social platforms. Additionally, the presence of rumor and model suggests an interest in examining rumors as a subset of fake news. Third, the topics



“disinformation,” “sharing,” and “platforms.” It explores consequences on public health, political contexts, and societal behavior, particularly during COVID-19.



**Figure 3:** Scatter plot of t-SNE-based clusters

#### 4.5 Societal Impacts and Strategies for Combating Fake News:

Context analysis using “Keywords in Context” for the keywords “impact, effect, consequence, and influence” helps to identify the social impact of fake news from the abstracts. In this section, major social impacts have been presented (Table 2). Psychological and cognitive issues include mental health problems, skepticism, and reduced self-efficacy. Social trust and relationship consequences are the other forms of impact. Behavioral impact manifests in vaccine hesitancy, rumor sharing, and altered consumer decisions, influenced by perceived credibility and social influence on social media. Economic and policy implications include disrupted decision-making and financial consequences. Misinformation amplification through echo chambers and illusory truth effects further set up societal problems. Lastly, technological impact deals with vulnerabilities in social media systems, fact-checking inefficiencies, and threats from deepfake technology. How can these challenges be addressed? To answer this question, this study highlights some strategies used or proposed in research studies (see Table 2) to counter the consequences. These strategies aim to combat the spread of fake news, enhance public trust, and promote critical thinking.

**Table 2:** Strategies discussed by the researchers

<b>Strategies</b>	<b>Reference</b>
Rumor management strategy	Li et al. (2023)
Community-based anti-rumor seeds	Dong et al. (2024)
Cross-platform governance	Jing et al. (2023)
Warning flags and explanatory text	Barman and Colan (2023)
Educating influencers and fact-checker bots	Yalabadi et al. (2024)
Fact-checking with empathy	Sun and Ma (2023)

Warning labels	B. Zhang et al. (2024)
Digital literacy campaigns	Máñez et al. (2023)

## 5. Conclusion:

This study identifies the fake news research trends, key research topics, and thematic clusters, and highlights major strategies to combat fake news. These strategies have significant implications for policymakers, researchers, and library professionals in developing frameworks to reduce fake news and misinformation. The study is limited by its selective use of keywords and search fields on Scopus, which may have restricted the range of research papers. Future studies could extend the scope by incorporating additional search terms and search fields to assess the long-term effectiveness of these strategies. Finally, this study emphasizes that continued research and collaborative efforts will be crucial in developing more effective measures to address the growing issue of fake news.

## Conflict of Interest:

The authors declare no conflicts of interest regarding this study.

## References:

1. BARMAN (D) and COLAN (O). Does explanation matter? an exploratory study on the effects of covid-19 misinformation warning flags on social media. *International Conference on Behavioural and Social Computing (BESC)*. 10th. 2023, p1-7. <https://doi.org/10.1109/BESC59560.2023.10386371>.
2. DONG (C), XU (G Q) and MENG (L). CRB: a new rumor blocking algorithm in online social networks based on competitive spreading model and influence maximization. *Chinese Physics B*. 33, 8; 2024. <https://doi.org/10.1088/1674-1056/ad531f>
3. GRADONĆ (K T), HOŁYST (J A), MOY (W R), SIENKIEWICZ (J) and SUCHECKI (K). Countering misinformation: a multidisciplinary approach. *Big Data & Society*. 8, 1; 2021. <https://doi.org/10.1177/20539517211013848>
4. ISMAILOĞLU (F). A text mining analysis on misinformation regarding the covid-19 pandemic. *Bilecik Şeyh Edebali Üniversitesi Fen Bilimleri Dergisi*. 9, 1; 2022. <https://doi.org/10.35193/bseufbd.959259>
5. JAYASEKARA (P K). Knowledge discovery of scholarly publications on misinformation on social media: a text mining approach. *Sri Lanka Library Review*. 34, 2; 2020. <https://doi.org/10.4038/sllr.v34i2.30>
6. JING (J), ZHANG (Z), CHOO (K K R), FAN (K), SONG (B) and ZHANG (L). Inference of user desires to spread disinformation based on social situation analytics and group effect. *IEEE Transactions on Dependable and Secure Computing*. 20, 3; 2023, p1833-48. <https://doi.org/10.1109/TDSC.2022.3165324>.
7. LI (Z), LI (Z) and HAJIYEV (A). Investigating how the content semantic features

- influence the social media rumor refutation effectiveness. *E3S Web of Conferences*. 409, 2023. <https://doi.org/10.1051/e3sconf/202340906007>
8. MÁÑEZ (M T M), CANO (A M) and DÍEZ (F). Impact of fake news on social networks during covid-19 pandemic in Spain. *Young Consumers*. 25, 4; 2023, p439-61. <https://doi.org/10.1108/YC-04-2022-1514>
  9. MAZUMDER (S) and BARUI (T). Evolution of technologies in libraries: a text mining study. *Librarian*. 27, 1 & 2; 2024. <http://eprints.rclis.org/45779/>
  10. NISTOR (A) and ZADOBRISCHI (E). The influence of fake news on social media: analysis and verification of web content during the covid-19 pandemic by advanced machine learning methods and natural language processing. *Sustainability*. 14, 17; 2022. <https://doi.org/10.3390/su141710466>
  11. SAHU (L), MAZUMDER (S) and DEB (S). How students identify and evaluate fake news: a case study of a management college. *Qualitative and Quantitative Methods in Libraries*. 12,1, 2023. <https://qqml-journal.net/index.php/qqml/article/view/794>
  12. SINCLAIR (S) and ROCKWELL (G). Voyant-tools. *Информационные Системы*. 2020. <http://ojs.iculture.spb.ru/index.php/systems/article/view/5>
  13. SU (Q), WAN (M), LIU (X) HUANG (C R). Motivations, methods and metrics of misinformation detection: an NLP perspective. *Natural Language Processing Research*. 1, 1–2; 2020, p1-13. <https://doi.org/10.2991/nlpr.d.200522.001>
  14. SUN (M) and MA (X). Combating health misinformation on social media through fact-checking: the effect of threat appraisal, coping appraisal, and empathy. *Telematics and Informatics*. 84, 2023, p102031. <https://doi.org/10.1016/j.tele.2023.102031>
  15. VERMA (A) and NAYAK (J K). Understanding public sentiments and misbeliefs about sustainable development goals: a sentiment and topic modeling analysis. *Journal of Information, Communication and Ethics in Society*. 22, 2; 2024, p256–74. <https://doi.org/10.1108/JICES-05-2023-0073>
  16. YALABADI (A K), YAZDANI-JAHRAMI (M), ABDIDIZAJI (S), GARIBAY (I) and GARIBAY (O O). Controlling the misinformation diffusion in social media by the effect of different classes of agents. 2024. In: YANG (Z) and KREJCI (C), Eds. *Proceedings of the 2023 International Conference of The Computational Social Science Society of the Americas*. Springer Nature, Switzerland. 2023, p143–57. [https://doi.org/10.1007/978-3-031-64193-0\\_9](https://doi.org/10.1007/978-3-031-64193-0_9)
  17. ZHANG (B), CHEN (L) and MOE (A). Examining the effects of social media warning labels on perceived credibility and intent to engage with health misinformation: the moderating role of vaccine hesitancy. *Journal of Health Communication*. 29, 9; 2024, p556-65. <https://doi.org/10.1080/10810730.2024.2385638>
  18. ZHANG (D), FANG (B), YANG (L) and CAI (Y). Disinformation of text mining online about tobacco and the covid-19 discussed on Sina Weibo. *Tobacco Induced Diseases*. 19, 83; 2021. <https://doi.org/10.18332/tid/142776>