information

Article

Visualizing Social Media Research in the Age of COVID-19

Panagiotis D. Michailidis

Department of Balkan, Slavic and Oriental Studies, University of Macedonia, 54636 Thessaloniki, Greece; pmichailidis@uom.edu.gr

Abstract: During the last three years, numerous research papers have been reported which use social media data to explore several issues related to the COVID-19 pandemic. Bibliometric methods in this work are used to analyze 1427 peer-reviewed documents from the last three years extracted from the Web of Science database. The results of this study show that there was high growth in publications in open access journals with an annual rate reaching 19.3% and they also identify the top cited journals and research papers. The thematic analysis of papers shows that research topics related to social media for surveillance and monitoring of public attitudes and perceptions, mental health, misinformation, and fake news are important and well-developed, whereas topics related to distance-learning education with social media are emerging. The results also show that the USA, China, and the UK have published many papers and received a high number of citations because of their strong international collaboration.

Keywords: bibliometric analysis; social media platforms; coronavirus; COVID-19; Biblioshiny; Web of Science; text analysis; social network analysis

Citation: Michailidis, P.D. Visualizing Social Media Research in the Age of COVID-19. *Information* **2022**, *13*, 372. https://doi.org/10.3390/info13080372

Academic Editor: Robin Haunschild

Received: 29 June 2022
Accepted: 2 August 2022
Published: 3 August 2022

1. Introduction

With the outbreak of the coronavirus 2019 (COVID-19) in Wuhan in December 2019 and the declaration of a pandemic by the World Health Organization in March 2020, social media platforms began to be used by users and governments to produce and disseminate important information. The information generated by users can be opinions or other subjective or inaccurate information which may include misinformation, rumors, and conspiracy theories, whereas the information generated by governments can be accurate health information for the general public in order to control the coronavirus. In recent years, researchers from various countries and disciplines have published numerous research papers that use social media data to interpret users' attitudes and perceptions about coronavirus and public health, as well as to detect outbreaks of infectious diseases or evaluate government messages related to the protection of public health.

A recent and systematic review of the role of social media platforms during the COVID-19 pandemic has been reported [1]. Due to the qualitative nature of the systematic review study [1], as well as the large number of publications on social media research and COVID-19 in recent years, it is difficult for scholars to study this large volume of literature manually to gain an overall and deeper insight into the intellectual structure of the research field. Therefore, a unified quantitative and qualitative review study of research articles published in the last three years on social media research on COVID-19 is proposed. Inspired by this, the main goal of this paper is to reveal and understand the big picture of social media research on COVID-19 by visualizing the publication and citation trends as well as the research trends or topics using bibliometric methods. To the best of our knowledge, no such bibliometric study has been undertaken in this research field until now.

Bibliometric analysis is a macroscopic tool to extract and uncover knowledge and patterns from a large amount of research literature very quickly compared to a traditional systematic review. In recent years, bibliometric analysis has attracted the interest of many

researchers for a variety of reasons, such as the emergence of digital technologies or bibliometric software such as VOSviewer, CiteSpace, and Biblioshiny and the development of academic databases such as Web of Science, Scopus, and Google Scholar [2–4]. Bibliometric analysis uses quantitative and qualitative methods. Quantitative methods include the descriptive and performance metrics for the research output of a field (the number of publications or citations, etc.) as well as identifying the most important research constituents (top cited papers, top productive sources, etc.). Qualitative methods include science mapping analysis to explore the relationships between research constituents [2]. The science mapping analysis is conducted through text and social network approaches, such as co-word analysis, co-citation analysis, and collaboration analysis [2,3].

This paper is organized as follows: Section 2 presents a literature review of the impact of COVID-19 on mass media. Section 3 presents the bibliometric methodology used in this paper. Section 4 provides empirical results from the bibliometric analysis. Finally, Section 5 gives a short overview of the main findings and conclusions of the paper.

2. Literature Review

Mass media includes a variety of media platforms (i.e., newspapers, magazines, radio, television, and the Internet) that reach a mass audience. In December 2019, a coronavirus emerged and within weeks it led to the emergence of the biggest global health crisis seen to date. During the current COVID-19 pandemic, mass media was playing a very important role in disseminating information and interacting with people. More specifically, mass media covered news about coronavirus such as the number of confirmed infected cases and death cases, the risks of coronavirus, the measures of governments (lockdowns and quarantines), and the social and economic hardships. A majority of people began to use social media more than traditional media during the COVID pandemic [5]. Social media platforms (Facebook, YouTube, Twitter, Instagram, Myspace, etc.) provide a venue for a large audience to express their opinions, sentiments, and subjective information in an unfiltered and uncensored manner. However, much of this information can be inaccurate, and social media platforms can become fertile lands for infodemics [6]. Infodemic refers to misinformation, fake news, and rumors. However, a lot of COVID-19-related news caused mental health problems such as stress, loneliness, depression, etc. On the other hand, the governments and public health authorities used social media to disseminate public health messages to the audience in order to gain public support and compliance with public health preventive measures [7].

Several recent research works have studied mass media use (traditional and social media) in disseminating information about the COVID-19 pandemic. Some studies addressed the impact of media news on individuals' positive and negative emotions and their psychological stability [8]. Another study demonstrated that 52% of the news headlines in English-language media evoked negative sentiments and only 30% evoked positive sentiments [9]. Such information can cause a series of mental health problems [9]. Some analyses of public reactions or the attitudes of citizens on social media platforms (i.e., Twitter) about COVID-19 were performed in these studies [10,11]. Other research studies addressed the issues of the COVID-19 misinformation [12] and the public's compliance with the exposure to risk communication messages about COVID-19 in the mass media [7]. Furthermore, with the increased amount of misinformation and fake news circulating on social media platforms, some machine learning works were developed to detect COVID-19 disinformation such as [13].

Finally, an overview study [1], systematically examined papers on the topic of the role of social media platforms during the COVID-19 pandemic. This study analyzed 81 papers between November 2019 and November 2020 from a qualitative perspective and identified six themes, such as infodemics, public attitudes, mental health, detection or prediction of COVID-19 cases, government responses to the pandemic, and quality of health information in videos [1]. However, the findings of this study required a lot of time and the interpretation of the results may be subjective. With the increased number

of publications on social media research on COVID-19 in recent years, it is important to update the literature review study. At the same time, the quantitative and qualitative review study presented in this work will provide more insightful and useful patterns for scholars in an objective way. The current study will also provide a way to discover research focus and gaps in previous studies.

3. Materials and Methods

In this section, we describe the research methodology used in this paper. The research methodology was a bibliometric analysis that involves statistical analysis and visualizing of the research output of a field. The bibliometric analysis was done using Biblioshiny, an open-source software, supported by the R environment, which provides tools for computing performance metrics and the visualization of various bibliometric networks [14]. The steps of workflow that were followed for the bibliometric analysis were: study design, data collection, data analysis, and visualization and interpretation of the findings [2,3].

3.1. Study Design

The main goal of this paper is to examine and visualize the research output on social media use for COVID-19 issues within the last three years, specifically from 2020 to 2022. The research questions of this study are the following:
RQ1: What is the evolution of publications and citations on social media research for COVID-19?
RQ2: What are the most relevant and influential sources, countries, and publications on social media research?
RQ3: What are the most common research topics and keyword trends on social media research for COVID-19?
RQ4: What are the main clusters of co-citations related to social media research for COVID-19?
RQ5: What is the collaborative network of countries on social media research for COVID-19?

3.2. Data Collection

For this bibliometric research, bibliometric data were retrieved from the Web of Science (WoS) Core Collection database. Keywords were used to identify publications related to social media research for COVID-19. The following list of keywords for representing the topic of social media was used: "social media", "social network*", "Web 2.0", "online communit*", "Facebook", "YouTube", "Twitter", "LinkedIn", "Instagram", "Pinterest", and "Myspace". Furthermore, the following list of keywords for representing the topic of COVID-19 was used: "covid", "covid19", "covid-19", and "coronavirus". These two lists of keywords were used in conjunction to search the titles of the publications within the WoS database. Those keywords were used only for title searches because they represent the relevant topic. Therefore, the following search criteria were used: ("social media" OR "social network*" OR "Web 2.0" OR "online communit*" OR "Facebook" OR "YouTube" OR "Twitter" OR "LinkedIn" OR "Instagram" OR "Pinterest" OR "Myspace") AND (covid OR covid19 OR "covid-19" OR coronavirus) in the title field. The search was launched on 11 June 2022 and 1753 documents were extracted.

The search was then refined by document type (conference paper, article, review) and language (English). As a result, the filtering stage returned 1427 documents analyzed in this paper. Furthermore, the 2021 journal impact factors were obtained from the Journal Citation Reports from Clarivate.

Table 1 presents a summary of the main information about the dataset. Specifically, our dataset contains 1427 articles published between 2020 and 2022. These articles were published in 680 scientific sources. The research dataset consists of three types of documents: articles (1396), conference papers (4), and review papers (27). The average number of years taken for the article to be cited is 0.92 (or 1) approximately and each article has 9.6 citations. The total number of cited references for all articles in the research dataset is 50,970. Furthermore, our articles contain 3297 author's keywords and 1470 keywords plus.

The author's keywords are keywords defined by authors to determine the content of their publications, whereas keywords plus are keywords generated by the WoS database from titles, keywords, and abstracts of publications. Our dataset covers 4969 unique authors, and the total number of contributing authors is 5791. 113 and 4856 authors have written single-authored and multi-authored articles, respectively. On average, each article is written by about 3 authors (i.e., authors per document is 3.48) and 117 articles are written by a single author. Therefore, the collaboration index is around 3.71.

Table 1. Main information about the bibliographic dataset.

Description	**Results**
Timespan	2020–2022
Sources (Journals, Books, etc.)	680
Documents	1427
Average years from publication	0.923
Average citations per document	9.601
Average citations per year per document	3.844
References	50,970
Document Types	
Article	1396
Proceedings papers	4
Review	27
Document Contents	
Keywords Plus (ID)	1470
Author's Keywords (DE)	3297
Authors	
Authors	4969
Author Appearances	5791
Authors of single-authored documents	113
Authors of multi-authored documents	4856
Authors Collaboration	
Single-authored documents	117
Documents per Author	0.287
Authors per Document	3.48
Co-Authors per Documents	4.06
Collaboration Index	3.71

3.3. Data Analysis and Visualization

After loading our dataset into Biblioshiny [14], two forms of bibliometric analysis were performed to answer the research questions. First, a performance analysis was conducted to show the publication and citation patterns, the productive sources, authors, and countries, as well as the most cited papers. Second, a science mapping analysis was performed to explore topic and keyword trends through co-word analysis, the clusters of co-citations through co-citation analysis, and the country collaboration structure through collaboration analysis. These latter results were visualized as networks.

3.4. Interpretation of the Results

The results of the performance and science mapping analysis are presented and interpreted in Section 4.

4. Results

In this section, we present our results based on the research questions in the study design section.

4.1. Publications and Citations Evolution

In this subsection, results are presented to answer the RQ1: What is the evolution of publications and citations on social media research for COVID-19? Table 2 shows the distribution of publications and citations. The highest number of publications was published in 2021, representing 56% (797) of the total number of publications. In the year 2022, there was a decrease in the number of publications compared to the previous year due to the fact that the papers were retrieved from the WoS database only up until June 2022. The number of publications is expected to increase by the end of 2022. On the other hand, the number of citations decreased as the year increased. This phenomenon makes sense because the citation time period is decreased as the year of publication is increased.

Table 2. Distribution of publications and citations.

Year	N	TC
2020	260	9105
2021	797	4288
2022	370	307

N = number of papers, TC = number of citations.

4.2. Most Relevant Sources, Countries, and Publications

In this subsection, results are presented to answer the RQ2: What are the most relevant and influential sources, countries, and publications on social media research? Table 3 presents the top 10 productive sources based on the number of publications. These sources cover 23% of the total number of publications in our dataset. The top three journals that cover articles on social media research for COVID-19 issues are the *Journal of Medical Internet Research*, the *International Journal of Environmental Research, and Public Health and Sustainability*. We also can see that the first nine journals were open access except for *Computers in Human Behavior,* and they also had a high impact as shown by the journal impact factors. This shows that most authors preferred to publish papers in high-impact and open-access journals during the COVID pandemic to be accessible to the research community quickly. Table 4 shows the top 10 most local cited sources. The local cited sources measure the number of local citations received by each cited reference within the reference lists of publications in the dataset. From these results, we can see that the *Journal of Medical Internet Research* is the most cited source among researchers in the social media research field for COVID. This source has been cited 1492 times. The second highest cited source is *Computers in Human Behavior* (1081 times), followed by *PLOS One* (1072 times). This shows that these top journals are the main references in this research domain.

Table 3. Top 10 productive sources.

Source Title	Documents	JIF *
Journal of Medical Internet Research	92	7.093
International Journal of Environmental Research and Public Health	75	4.614
Sustainability	31	3.889
PLOS One	29	3.752
JMIR Public Health and Surveillance	23	14.557
Frontiers in Psychology	22	4.232
BMC Public Health	15	4.135
IEEE Access	14	3.476
Vaccines	14	4.961
Computers in Human Behavior	13	8.957

* JIF = Journal Impact Factor, 2021 Journal Impact Factor, *Journal Citations Reports* (Clarivate, 2022).

Table 4. Top 10 most local cited sources.

Source Title	Citations
Journal of Medical Internet Research	1492
Computers in Human Behavior	1081
PLOS One	1072
International Journal of Environmental Research and Public Health	673
Lancet	563
The Journal of the American Medical Association	348
Health Communication	340
Public Relations Review	312
The New England Journal of Medicine	305
Science	300

Table 5 presents the top 10 most influential sources based on total citations. The total citations of a source is the number of citations received by published papers within the source in the dataset. We also note that Table 5 presents those sources that have at least one citation. From these results, we can see that the top three journals, the *Journal of Medical Internet Research*, *PLOS One* and the *International Journal of Environmental Research and Public Health*, have published many citable articles and received a high number of total citations with a high h-index. However, there are also journals, such as *Phychological Science*, *Cureus*, *Human Vaccines and Immunotherapeutics* and *Phychological Medicine,* that have a high number of citations with a limited number of published articles on social media research for COVID.

Table 5. Top 10 most influential sources based on total citations.

Source Title	h-Index	Total Papers	Total Citations	JIF *
Journal of Medical Internet Research	24	79	1906	7.093
PLOS One	10	26	1279	3.752
International Journal of Environmental Research and Public Health	13	44	566	4.614
Phychological Science	1	1	392	-
Computers in Human Behavior	5	11	336	8.957
Scientific Reports	4	10	326	4.996
Cureus	3	4	316	-
Sustainability	8	25	309	3.889
Human Vaccines and Immunotherapeutics	4	5	287	4.526
Phychological Medicine	2	2	255	-

* JIF = Journal Impact Factor, 2021 Journal Impact Factor, *Journal Citations Reports* (Clarivate, 2022).

Figure 1 demonstrates the country-specific scientific production of publications on social media use for COVID-19. Table 6 shows the details of the top 20 productive countries; SCP is Single Country Publication, MCP is Multiple Country Publication, and MCP ratio is the proportion of the total number of publications. These results show that the USA and China have published more than 150 papers and 67 of these involved international collaborations. The third highest production of publications is in the UK and 31 of these involved international collaborations. However, some observations can be made regarding the MCP ratio. Countries such as Japan, Malaysia, and Germany have higher degrees of international collaboration than other countries.

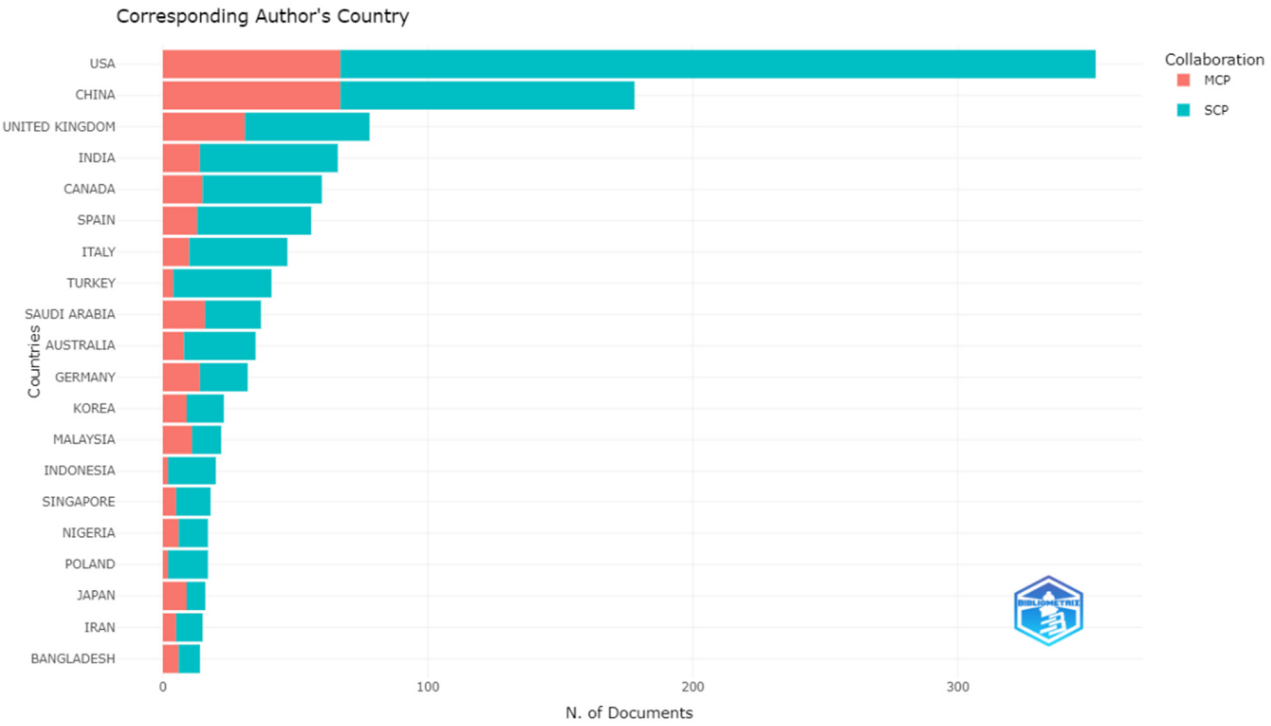

Figure 1. Top 20 most productive countries (this chart was generated by Biblioshiny).

Table 6. Top 20 most productive countries.

Country	Documents	SCP	MCP	MCP Ratio
USA	352	285	67	0.1903
China	178	111	67	0.3764
UK	78	47	31	0.3974
India	66	52	14	0.2121
Canada	60	45	15	0.25
Spain	56	54	13	0.2321
Italy	47	37	10	0.2128
Turkey	41	37	4	0.0976
Saudi Arabia	37	21	16	0.4324
Australia	35	27	8	0.2286
Germany	32	18	14	0.4375
Korea	23	14	9	0.3913
Malaysia	22	11	11	0.50
Indonesia	20	18	2	0.10
Singapore	18	13	5	0.2778
Nigeria	17	11	6	0.3529
Poland	17	15	2	0.1176
Japan	16	7	9	0.5625
Iran	15	10	5	0.3333
Bangladesh	14	8	6	0.4286

Figure 2 shows the most cited countries in the social media research field for COVID, and Figure 3 demonstrates the normalizing of the number of publications per country to their respective citations. Figure 2 shows that China is the leader in this research field, followed by the USA, UK, and Canada. The first four countries have a high impact because they have contributed a lot of research publications in this field. However, there are also countries such as Lebanon, Iraq, Switzerland, Vietnam, Bahrain, Bolivia, and Canada which seem to receive a high number of citations with a limited number of published articles (such as 3, 2, 6, 4, 1, 1, and 60 publications, respectively), as is shown in Figure 3.

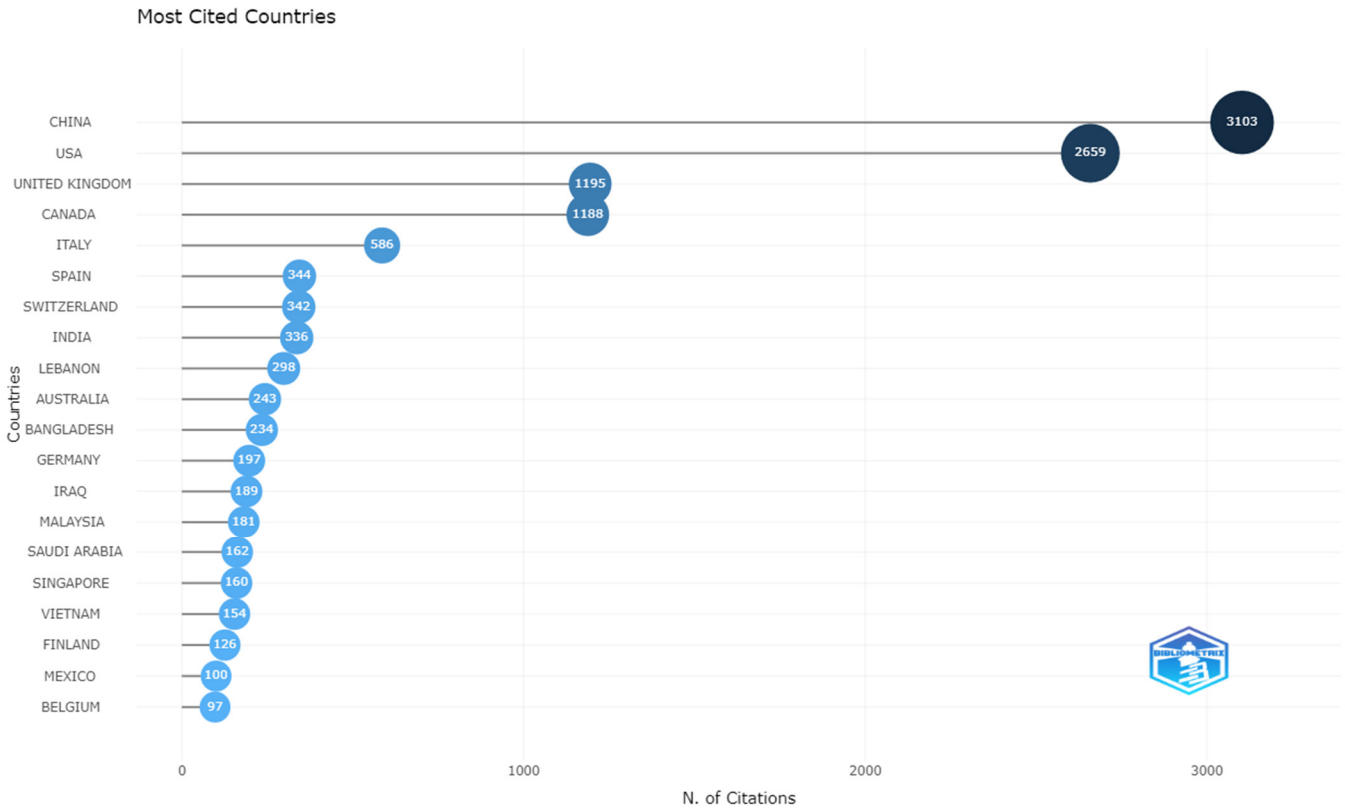

Figure 2. Most cited countries (this chart was generated by Biblioshiny).

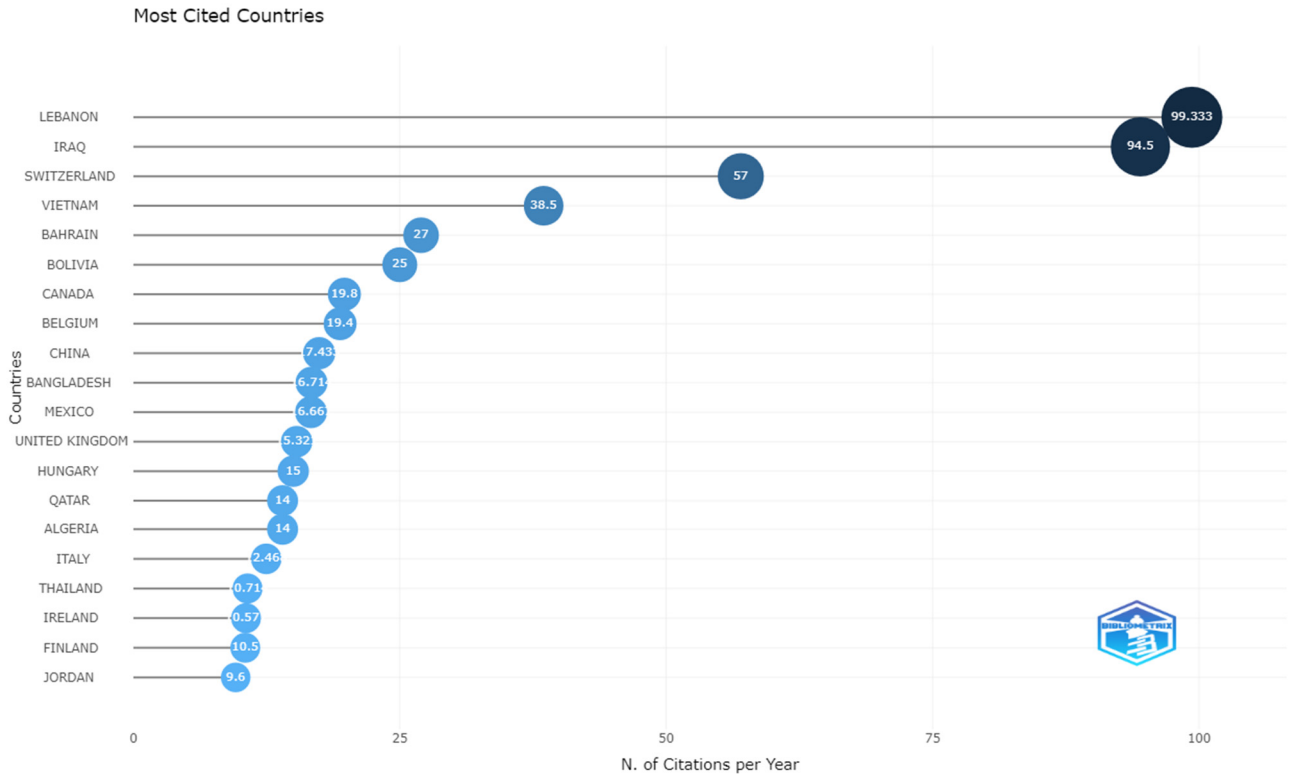

Figure 3. Citations per paper ratio (this chart was generated by Biblioshiny).

We divided the articles into research articles and review articles. The top 10 cited research and review papers based on total citations are presented in Tables 7 and 8, respec-

tively. As we can observe from the results reported in Table 7, the study by Gao, J. et al. (2020) [15] has received the highest number of citations followed by the ones written by Pennycook, G. et al. (2020) [16] and by Elmer, T. et al. (2020) [17]. Based on content analysis of the top research articles, three themes are identified: 3 of the 10 papers focused on accessing mental health, 6 of the 10 focused on identifying infodemics, and 1 of the 10 focused on the evaluation of health information quality. Below is the analysis of these research papers.

The studies by Gao, J. et al. (2020) [15], Elmer, T. et al. (2020) [16], and Ni, M.Y. et al. (2020) [18] refer to the impacts on mental health of social media exposure during the COVID pandemic. More specifically, the work by Gao, J. et al. (2020) conducted an online survey of 4872 Chinese citizens and showed that frequent social media exposure was associated with increased anxiety [15]. The work by Elmer, T. et al. (2020) used two cohorts of Swiss undergraduate students, one that experienced the COVID crisis and one that did not, and they found that exposure to social networks, lack of interaction, and physical isolation were associated with negative mental health problems such as anxiety, stress, and loneliness [16]. Finally, the work by Ni, M.Y. et al. (2020) conducted a similar online survey via social media platforms in China and they also reported mental health problems [18].

The papers by Pennycook, G. et al. (2020) [16], Cinelli, M. et al. (2020) [19], Kouzy, R. et al. (2020) [20], Allington, D. et al. (2021) [21], Ahmed, W. et al. (2020) [22], and Islam, M. et al. (2020) [23] refer to infodemics. The term infodemics is a combination of accurate and inaccurate information about an epidemic, such as the COVID-19 pandemic [1]. The inaccurate information may include misinformation, fake news, or rumors. More specifically, the work by Pennycook, G. et al. (2020) performed two studies to examine why people share and believe COVID-19 fake news [16]. One study showed that people share fake news because they do not think about the accuracy of the content before deciding to disseminate it, whereas the second study showed that people with scientific knowledge and thinking can identify false information about COVID [16]. The paper by Cinelli, M. et al. (2020) examined the spreading of fake news about the COVID pandemic on specific social media platforms and by users who have dealt with the topic [19]. The study by Kouzy, R. et al. (2020) quantified the misinformation about the COVID-19 on Twitter by analyzing 673 English tweets. They also showed that 24.8% of tweets were misinformation and 12.3% of tweets from public health accounts were unverifiable information [20]. The paper by Allington, D. et al. (2021) conducted three surveys of social media use in the UK and found a negative link between COVID-19 conspiracy beliefs and health-protective behaviors, as well as a positive link between COVID-19 conspiracy beliefs and the use of social media platforms [21]. The work by Ahmed, W. et al. (2020) investigated 5G and COVID-19 conspiracy theories on Twitter by analyzing the content of 233 tweets and found that 34.8% linked 5G with COVID-19 and 32.2% denounced the conspiracy theory [22]. Finally, the work by Islam, M. et al. (2020) analyzed 2311 reports about COVID in 25 languages from 87 countries and they found that misinformation was mainly driven by rumors, stigma, and conspiracy theories that were discussed on social media platforms [23].

Finally, the paper by Puri, N. et al. (2020) refers to information quality against vaccine misinformation for COVID-19 [24]. More specifically, this study examined the role of the propagation of vaccine hesitancy and proposed digital health strategies to overcome vaccine misinformation on social media platforms [24].

We can also observe from the results reported in Table 8 that the three review studies that received the highest number of citations are Tsao, S.F. (2021) [1], Shani, H.; Sharma, H. (2020) [25], and Gabarron, E. et al. (2021) [26]. Table 8 presents those review articles that have at least 10 citations. Based on content analysis of the top review articles, we identified that one of the five papers focused on an overview of the role of social media in the COVID era, two of the five focused on the impact of social media during the pandemic, one of the five focused on an overview of the misinformation on social media and one of the five papers focused on a review related to distance learning with social media.

Table 7. Top 10 research articles by total citations.

Authors	Article Title	Source Title	TC
Gao, J. et al. (2020) [15]	Mental health problems and social media exposure during COVID-19 outbreak	PLOS One	636
Pennycook, G. et al. (2020) [16]	Fighting COVID-19 Misinformation on social media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention	Psychological Science	392
Elmer, T. et al. (2020) [17]	Students under lockdown: Comparisons of students' social networks and mental health before and during the COVID-19 crisis in Switzerland	PLOS One	332
Cinelli, M. et al. (2020) [19]	The COVID-19 social media infodemic	Scientific Reports	295
Kouzy, R. et al. (2020) [20]	Coronavirus Goes Viral: Quantifying the COVID-19 Misinformation Epidemic on Twitter	Cureus	287
Puri, N. et al. (2020) [24]	Social media and vaccine hesitancy: new updates for the era of COVID-19 and globalized infectious diseases	Human Vaccines & Immunotherapeutics	252
Allington, D. et al. (2021) [21]	Health-protective behavior, social media usage and conspiracy belief during the COVID-19 public health emergency	Psychological Medicine	251
Ahmed, W. et al. (2020) [22]	COVID-19 and the 5G Conspiracy Theory: Social Network Analysis of Twitter Data	Journal of Medical Internet Research	211
Islam, M. et al. (2020) [23]	COVID-19–Related Infodemic and Its Impact on Public Health: A Global Social Media Analysis	The American Journal of Tropical Medicine and Hygiene	211
Ni, M.Y. et al. (2020) [18]	Mental Health, Risk Factors, and Social Media Use During the COVID-19 Epidemic and Cordon Sanitaire Among the Community and Health Professionals in Wuhan, China: Cross-Sectional Survey	JMIR Mental Health	190

The review paper by Tsao, S.F. (2021) conducted a systematic review of 81 papers from three databases (PubMed, Scopus, and PsycINFO) and identified six topics related to the role of social media in COVID-19 [1]. The six topics were infodemics, public attitudes, mental health, detecting or predicting COVID-19 cases, government responses, and quality of health information in prevention education videos [1]. The papers by Shani, H.; Sharma, H. (2020), and Verner Venegas-Vera, A.V. et al. (2020) outlined the positive and negative impact of social media platforms during the COVID pandemic on healthcare professionals and the general public [25,27]. The work by Gabarron, E. et al. (2021) performed a systematic review of empirical publications on a specific topic of infodemics related to misinformation about COVID-19 on social media platforms [26]. Finally, the paper by Cavus, N. et al. (2021) conducted a literature review of papers related to eLearning education with social media platforms during the COVID-19 era and highlighted the eLearning challenges and strategies for the sustainable educational use of social media by both institutions, teachers, and students [28].

Table 8. Top 5 review articles by total citations.

Authors	Article Title	Source Title	TC
Tsao, S.F. (2021) [1]	What social media told us in the time of COVID-19: a scoping review	Lancet	82
Shani, H.; Sharma, H. (2020) [25]	Role of social media during the COVID-19 pandemic: Beneficial, destructive, or reconstructive?	International Journal of Academic Medicine	32
Gabarron, E. et al. (2021) [26]	COVID-19-related misinformation on social media: a systematic review	Bull World Health Organization	19
Venegas-Vera, A.V. et al. (2020) [27]	Positive and negative impact of social media in the COVID-19 era	Reviews in Cardiovascular Medicine	18
Cavus, N. et al. (2021) [28]	Efficacy of Social Networking Sites for Sustainable Education in the Era of COVID-19: A Systematic Review	Sustainability	12

4.3. Research Topics and Keywords Trends

In this subsection, results are presented to answer the RQ3: What are the most common research topics and keyword trends on social media research for COVID-19? We present a thematic analysis to detect the main research topics in the field using a word cloud and a thematic map. To avoid deviant results, we removed the keywords inserted in the search query (such as terms related to social media and COVID-19). Figure 4 shows the word cloud for the 50 most common author keywords in the publications collection. The size of the keyword in the figure indicates the frequency of the keyword in the dataset. As we can see from the figure, the most common words determine the content of most studies in the collection. More specifically, the frequent keyword "pandemic" is the main topic since the papers in the collection address several issues about the COVID-19 pandemic. The keywords "sentiment analysis", "machine learning", "natural language processing" "topic modeling", "content analysis", and "text mining" show their importance and represent the main methodologies based on their conceptual meaning. These methodologies were used to analyze the public attitudes of social media users about the COVID-19 pandemic and vaccines. "Misinformation", "public health", "infodemiology", "mental health", "anxiety", "infoveillance", "vaccination", "social networks", "fake news", "health communication", "vaccine hesitancy" and "education" also show their importance. These keywords address several specific issues about the pandemic such as the dissemination of inaccurate information on social media platforms, the impacts of social media on users' mental health during the pandemic, and the evaluation of the spreading of government messages on social media for the protection of public health.

To achieve further understanding, Figure 5 shows a thematic map based on author keywords, as proposed by Cobo et al. [29]. For the thematic map, some parameters were fixed, such as the number of words (=500) and minimum cluster frequency (=3), while keywords inserted in the search query were removed. The thematic map consists of four quadrants according to their centrality and density values along two axes. The centrality measures the importance of a theme compared with other themes on the map. The density measures the development of internal links within a cluster represented by a theme. The size of the cluster indicates the number of occurrences of the keywords that it contains, and the position of the cluster is set according to the cluster centrality and density. The label of the cluster chosen by the Biblioshiny software corresponds to the most frequent keywords.

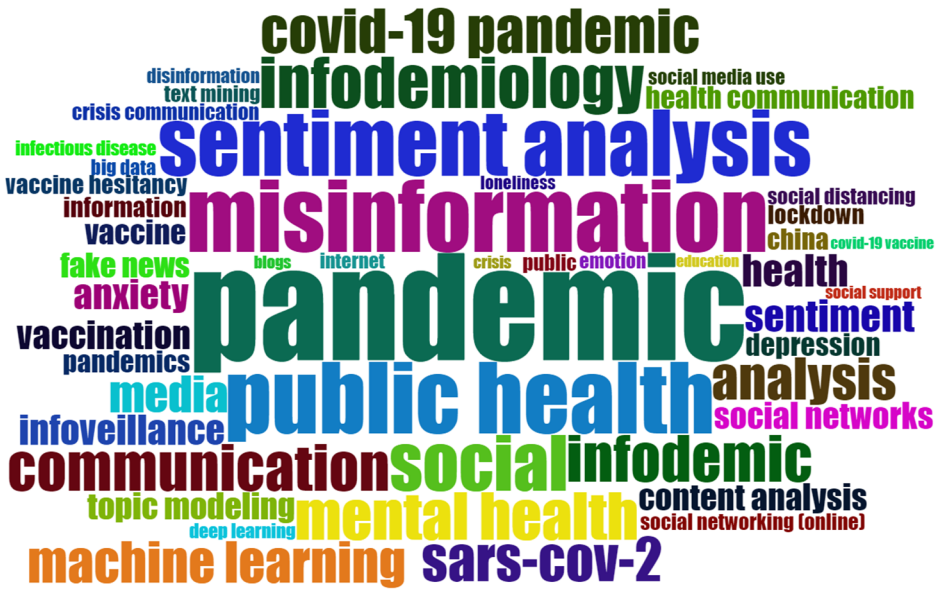

Figure 4. Word cloud based on author keywords (this word cloud was generated by Biblioshiny).

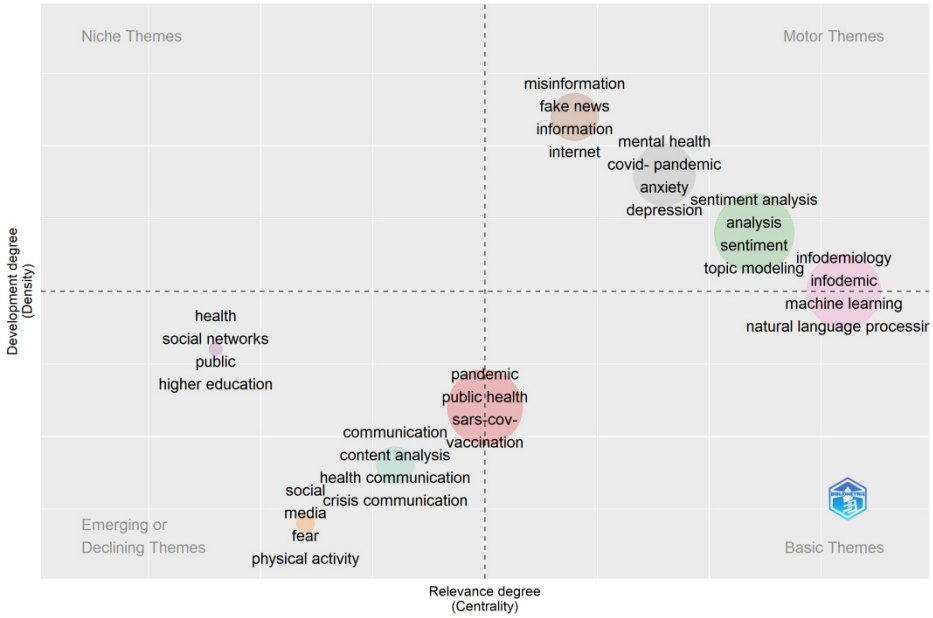

Figure 5. Thematic map of authors' keywords (this map was generated by Biblioshiny).

The upper-right quadrant involves motor themes that are important and well-developed for the field. In this quadrant, there are four clusters. Cluster 1 includes "sentiment analysis", "analysis", "sentiment", and "topic modeling", cluster 2 includes "infodemiology", "infodemic", "machine learning", and "natural language processing", cluster 3 includes "mental health", "covid-pandemic", "anxiety", and "depression", and cluster 4 includes "misinformation", "fake news", "information", and "internet". Therefore, these clusters focus on social media for surveillance and monitoring of public attitudes and perceptions, mental health, and the dissemination of inaccurate information and conspiracy theories on social media as in [20,22,23,30–34]. The attitudes and perceptions of social media users were analyzed using machine learning methods such as topic modeling and sentiment analysis. The lower-right quadrant includes basic themes that are important for the field but not well-developed. In this quadrant, there is a cluster that is marginal with the lower-left quadrant. This cluster involves the keywords "pandemic", "public health", "sars-cov", and "vaccination" and it concerns the sharing of health-related information by social media

users to influence their decision-making about vaccination, i.e., vaccine hesitancy [24]. The lower-left quadrant includes emerging or declining themes with low centrality and density. In this quadrant, there are three clusters. Cluster 1 includes "communication", "content analysis", "health communication", and "crisis communication" and concerns the evaluation of government messages and other health information, and how announcements were consumed on social media platforms [35–37]. Cluster 2 includes "social", "media", "fear", and "physical activity" and it concerns studies addressing the links between the fear of COVID and social network use as well as the impacts of sharing physical activity experiences on social media platforms during COVID-19 lockdown [38,39]. Finally, cluster 3 includes "health", "social networks", "public", and "higher education" and it concerns eLearning education with social media during the pandemic for sustaining usage by students and faculties [28,40].

4.4. Co-Citation Network

In this subsection, results are presented to answer RQ4: What are the main clusters of co-citations related to social media research for COVID-19? Figure 6 shows the co-citation network. This network was performed with a minimum degree of co-citation equal to three and a threshold of 50 network nodes. The node of a network was labeled by the first author and publication year of the paper whereas the edge of the network is the co-citation between two documents. The size of the node indicates the number of local citations received by the documents and the thickness of the edge represents the strength of co-citation ties. The color of the node shows the cluster with which the paper is associated.

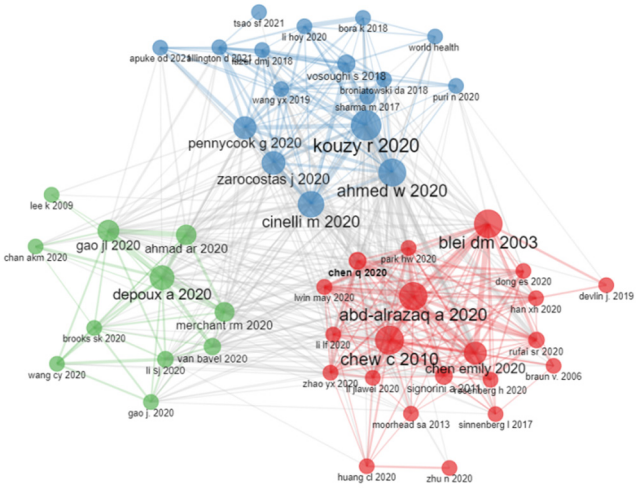

Figure 6. Co-citation network (this network was generated by Biblioshiny).

From the co-citation network in Figure 6, we can observe that there are three clusters of co-citations. The clusters were named based on the majority of the references belonging to them. The first cluster with a red color was named *infoveillance and public attitudes*, focusing on empirical studies monitoring the attitudes and sentiments of social media users for the COVID-19 pandemic and similar viruses (i.e., H1N1). This cluster is the largest and it includes 23 works. The top three cited references of this cluster are represented by Chew et al. (2010) [41], Abd-Alrazaq et al. (2020) [42], and Chen et al. (2020) [43]. The second cluster with a blue color referred to *infodemics*, focusing on works addressing specific issues i.e., misinformation, fake news, and rumors on social media platforms. This cluster consists of 14 articles and the top three cited references are represented by Kouzy et al. (2020) [20], Ahmed et al. (2020) [22], and Cinelli et al. (2020) [19]. The third cluster with a green color was named *impacts on mental health*, focusing on mental health problems caused by exposure to social media. This cluster contains 12 articles and the three top cited references are by Depoux et al. (2020) [44], Gao et al. (2020) [15], and Ahmad et al. (2020) [45]. We can also see that works within each cluster are interconnected.

However, Figure 6 shows that the works of the red cluster linked more with the works of the blue cluster than the works of the green cluster. Therefore, there is a close relationship between the research areas of infodemics and infoveillance.

4.5. Country Collaboration Network

In this subsection, results are presented to answer RQ5: What is the collaborative network of countries on social media research for COVID-19? Figure 7 shows the country collaboration map. This map depicts each country's publication output and the collaboration among countries. In this map, countries with a darker color represent more publications than countries with a light color whereas the thickness of lines represents stronger collaborations among countries. Figure 8 shows the social network of collaboration at the country level in detail. The node of the network represents the country and the edge between two nodes represents the cooperation between countries. The size of the country indicates the degree of its cooperation, and the thickness of the edge indicates the closeness of the collaboration between countries.

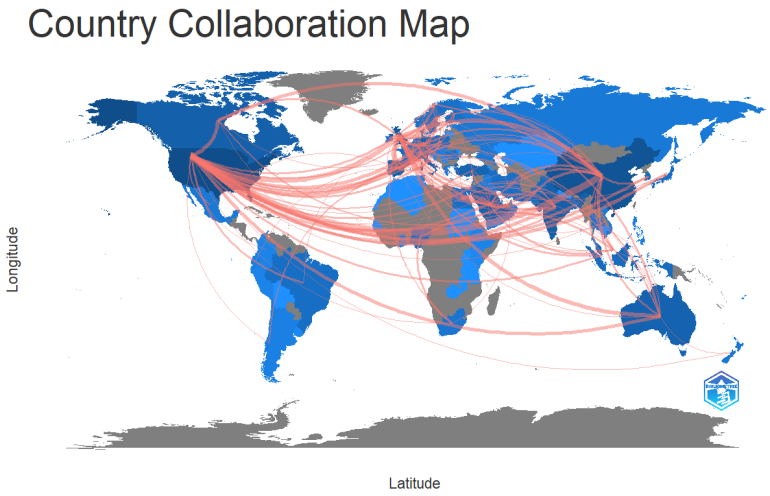

Figure 7. Map of collaboration among countries (this map was generated by Biblioshiny).

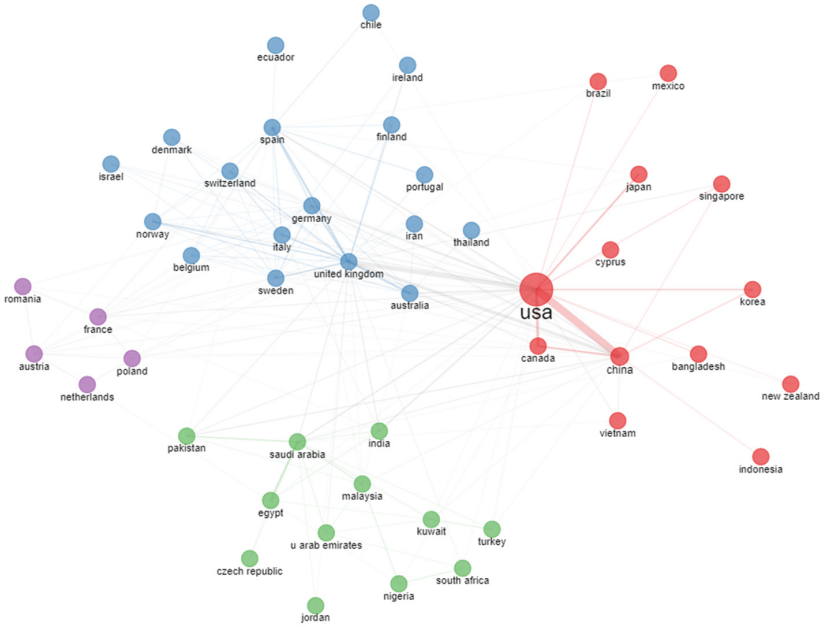

Figure 8. Social network of collaboration at the country level (this social network was generated by Biblioshiny).

From these figures, we can observe that there are many collaborations between countries and the network shows four clusters of collaboration among countries. The largest cluster of collaboration is the cluster with a blue color whereas the smallest cluster is the cluster with a purple color. Also, few countries have many collaborations, and many countries have a few collaborations. The USA has many research collaborations as is shown by the size of the node in Figure 8 and this is due to the high number of publications as is shown in Figure 7 (the region with a darker color). The USA is also a bridge between the four clusters of the country collaboration network. Finally, we can see that there are strong collaborations between countries on the use of social media for COVID (as indicated by the thick line in Figure 8) by countries such as the USA and China, the USA and UK, China and the UK, the USA and Canada, and the UK and Australia.

5. Discussion and Conclusions

In this paper, a bibliometric analysis was adopted to visualize the social media research on the COVID pandemic. Although there are useful qualitative and systematic reviews [1], to the best of our knowledge this is the first work that used quantitative and qualitative methods to explore and analyze a large volume of research publications. The analysis showed that researchers produced a large number of publications about the use of social media data for COVID in a very short period (2020–2022) with an annual growth rate of nearly 19.3%. Most of the papers received a high number of citations in 2020 since their citable time period was greater than in 2021 and 2022. Furthermore, most papers were published in journals rather than in conferences, especially in the first two years (2020 and 2021) and this is due to the measures of governments to prevent the spread of the coronavirus to the public. The first publications in conferences started in 2022.

Using the bibliometric results it was identified that journals such as the *Journal of Medical Internet Research*, the *International Journal of Environmental Research, and Public Health* and *Sustainability* published a large number of papers, and this is due to the high impact of open access journals. The geographical distribution of publications was very fragmented because the COVID pandemic was of global interest to most researchers as shown by the quantitative analysis. The most productive and influential countries were the USA, China, and the UK and this is due to their strong international collaboration as shown by the performance and collaboration network analysis. The performance analysis also showed that China is the leader in social media research for COVID because it has received many citations.

The top 10 most cited research articles and the five review articles which had the most significant impact were identified. The research articles focused on three themes: 3 of the 10 papers concerned accessing mental health, 6 of the 10 concerned identifying infodemics, and 1 of the 10 concerned the evaluation of health information quality consumed on social media. On the other hand, the review articles focused on four themes: one of the five papers concerned an overview of the role of social media in the COVID era, two of the five concerned the impact of social media during the pandemic, one of the five concerned an overview of the misinformation on social media, and one of the five papers concerned a review related to distance learning with social media. The aforementioned research topics were confirmed by the keyword trend analysis and the thematic map as shown in Section 3. From the thematic map analysis, it was found that vital and well-developed themes related to social media for surveillance and monitoring of public attitudes and perceptions, mental health, misinformation, and fake news. On the other hand, the analysis found some emerging themes related to eLearning education with social media and the evaluation of public health-related information on social media platforms.

The intellectual structure of the social media research for COVID issues consists of three main thematic clusters: infoveillance and public attitudes, infodemics, and impacts on mental health. From co-citation analysis, it was revealed that there was a close relationship between the research areas of infodemics and infoveillance. This fact is confirmed by the

upper-right quadrant of the thematic which shows that these two research areas are vital and well-developed.

This study identified similar research topics to the work by [1] and some studies reported in Section 2. However, the findings of this study reveal some research themes or gaps (i.e., distance learning education, etc.) for further research compared to the previous studies. The topics of this current study were detected and synthesized using bibliometric methods which are very fast and less cumbersome than traditional systematic reviews [1]. In addition to the above topics, our study produced additional results that cannot be easily extracted by traditional reviews, such as publication and citation trends, the most cited sources and papers, the most leading and cited countries, the co-citation clusters, and country collaboration patterns. However, the limitation of the current study is that our publications are from the WoS database only. This is because the Biblioshiny software does not allow the merging of multiple files from multiple bibliographic sources such as Scopus and Google Scholar. In this case, the publications from multiple sources may give better visualization and knowledge results in the field. Another limitation of this study is that our research provides results only for the keywords used in Section 3.2. Therefore, other related keywords or terms might not be included in this bibliometric analysis, and this may be the subject of future research.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in the analysis were derived from the Web of Science Core Collection™ from Clarivate [46]. © 2022 Clarivate. All rights reserved. The author of this paper had access to the Web of Science™ database through the affiliate institution's library.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Tsao, S.-F.; Chen, H.; Tisseverasinghe, T.; Yang, Y.; Li, L.; Butt, Z. What social media told us in the time of COVID-19: A scoping review. *Lancet Digit. Health* **2021**, *3*, 175–194. [CrossRef]
2. Donthu, N.; Kumar, S.; Mukherjee, D.; Pandey, N.; Lim, W.M. How to conduct a bibliometric analysis: An overview and guidelines. *J. Bus. Res.* **2021**, *133*, 285–296. [CrossRef]
3. Zupic, I.; Cater, T. Bibliometric methods in management and organization. *Organ. Res. Methods* **2015**, *18*, 429–472. [CrossRef]
4. Moral-Muñoz, J.A.; Herrera-Viedma, E.; Santisteban-Espejo, A.; Cobo, M.J. Software tools for conducting bibliometric analysis in science: An up-to-date review. *Prof. Inf.* **2020**, *29*, e290103. [CrossRef]
5. Dhanashree; Garg, H.; Chauhan, A.; Bhatia, M.; Sethi, G.; Chauhan, G. Role of mass media and it's impact on general public during coronavirus disease 2019 pandemic in North India: An online assessment. *Indian J. Med. Sci.* **2021**, *73*, 21–25. [CrossRef]
6. Nguyen, H.; Nguyen, A. COVID-19 misinformation and the social (media) amplification of risk: A Vietnamese perspective. *Media Commun.* **2020**, *8*, 444–447. [CrossRef]
7. Thanh, P.T.; Tung, L.T. Can risk communication in mass media improve compliance behavior in the COVID-19 pandemic? *Evidence from Vietnam. Int. J. Sociol. Soc. Policy* **2021**, *41*, 1–17. [CrossRef]
8. Giri, S.P.; Maurya, A.K. A neglected reality of mass media during COVID-19: Effect of pandemic news on individual's positive and negative emotion and psychological resilience. *Pers. Individ. Differ.* **2021**, *180*, 110962. [CrossRef]
9. Aslam, F.; Awan, T.M.; Syed, J.H.; Kashif, A.; Parveen, M. Sentiments and emotions evoked by news headlines of coronavirus disease (COVID-19) outbreak. *Humanit. Soc. Sci. Commun.* **2020**, *7*, 23. [CrossRef]
10. Imran, A.S.; Daudpota, S.M.; Kastrati, Z.; Batra, R. Cross-Cultural Polarity and Emotion Detection Using Sentiment Analysis and Deep Learning on COVID-19 Related Tweets. *IEEE Access* **2020**, *8*, 181074–181090. [CrossRef]
11. Ridhwan, K.M.; Hargreaves, C.A. Leveraging Twitter data to understand public sentiment for the COVID-19 outbreak in Singapore. *Int. J. Inf. Manag. Data Insights* **2021**, *1*, 100021. [CrossRef]
12. Brennen, J.S.; Simon, F.M.; Howard, P.N.; Nielsen, R.K. Types, sources, and claims of COVID-19 misinformation. *Reuters Inst. Study J.* **2020**. Available online: https://ora.ox.ac.uk/objects/uuid:178db677-fa8b-491d-beda-4bacdc9d7069 (accessed on 1 June 2022).
13. Song, X.; Petrak, J.; Jiang, Y.; Singh, I.; Maynard, D.; Bontcheva, K. Classification aware neural topic model for COVID-19 disinformation categorisation. *PLoS ONE* **2021**, *16*, e0247086. [CrossRef] [PubMed]
14. Aria, M.; Cuccurullo, C. Bibliometrix: An R-tool for comprehensive science mapping analysis. *J. Informetr.* **2017**, *11*, 959–975. [CrossRef]

15. Gao, J.; Zheng, P.; Jia, Y.; Chen, H.; Mao, Y.; Chen, S.; Wang, Y.; Fu, H.; Dai, J. Mental health problems and social media exposure during COVID-19 outbreak. *PLoS ONE* **2020**, *15*, e0231924. [CrossRef]
16. Pennycook, G.; McPhetres, J.; Zhang, Y.; Lu, J.G.; Rand, D.G. Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention. *Psychol. Sci.* **2020**, *31*, 770–780. [CrossRef] [PubMed]
17. Elmer, T.; Mepham, K.; Stadtfeld, C. Students under lockdown: Comparisons of students' social networks and mental health before and during the COVID-19 crisis in Switzerland. *PLoS ONE* **2020**, *15*, e0236337. [CrossRef] [PubMed]
18. Ni, M.; Yang, L.; Leung, C.; Li, N.; Yao, X.; Wang, Y.; Leung, G.; Cowling, B.; Liao, Q. Mental Health, Risk Factors, and Social Media Use During the COVID-19 Epidemic and Cordon Sanitaire Among the Community and Health Professionals in Wuhan, China: Cross-Sectional Survey. *JMIR Ment. Health* **2020**, *7*, e19009. [CrossRef]
19. Cinelli, M.; Quattrociocchi, W.; Galeazzi, A. Valensise, C.M.; Brugnoli, E.; Schmidt, A.L.; Zola, P.; Zollo, F.; Scala, A. The COVID-19 social media infodemic. *Sci. Rep.* **2020**, *10*, 16598. [CrossRef]
20. Kouzy, R.; Abi Jaoude, J.; Kraitem, A.; El Alam, M.B.; Karam, B.; Adib, E.; Zarka, J.; Traboulsi, C.; Akl, E.W.; Baddour, K. Coronavirus Goes Viral: Quantifying the COVID-19 Misinformation Epidemic on Twitter. *Cureus* **2020**, *12*, e7255. [CrossRef]
21. Allington, D.; Duffy, B.; Wessely, S.; Dhavan, N.; Rubin, J. Health-protective behaviour, social media usage and conspiracy belief during the COVID-19 public health emergency. *Psychol. Med.* **2021**, *51*, 1763–1769. [CrossRef] [PubMed]
22. Ahmed, W.; Vidal-Alaball, J.; Downing, J.; López Seguí, F. COVID-19 and the 5G Conspiracy Theory: Social Network Analysis of Twitter Data. *J. Med. Internet Res.* **2020**, *22*, e19458. [CrossRef] [PubMed]
23. Islam, M.S.; Sarkar, T.; Khan, S.H.; Mostofa Kamal, A.H.; Hasan, S.; Kabir, A.; Yeasmin, D.; Islam, M.A.; Amin Chowdhury, K.I.; Anwar, K.S.; et al. COVID-19-Related Infodemic and Its Impact on Public Health: A Global Social Media Analysis. *Am. J. Trop. Med. Hyg.* **2020**, *103*, 1621–1629. [CrossRef] [PubMed]
24. Puri, N.; Coomes, E.A.; Haghbayan, H.; Gunaratne, K. Social media and vaccine hesitancy: New updates for the era of COVID-19 and globalized infectious diseases. *Hum. Vaccines Immunother.* **2020**, *16*, 2586–2593. [CrossRef]
25. Sahni, H.; Sharma, H. Role of social media during the COVID-19 pandemic: Beneficial, destructive, or reconstructive? *Int. J. Acad. Med.* **2020**, *6*, 70–75.
26. Gabarron, E.; Oyeyemi, S.O.; Wynn, R. COVID-19-related misinformation on social media: A systematic review. *Bull. World Health Organ.* **2021**, *99*, 455–463A. [CrossRef] [PubMed]
27. Venegas-Vera, A.V.; Colbert, G.B.; Lerma, E.V. Positive and negative impact of social media in the COVID-19 era. *Rev. Cardiovasc. Med.* **2020**, *21*, 561–564. [CrossRef]
28. Cavus, N.; Sani, A.S.; Haruna, Y.; Lawan, A.A. Efficacy of Social Networking Sites for Sustainable Education in the Era of COVID-19: A Systematic Review. *Sustainability* **2021**, *13*, 808. [CrossRef]
29. Cobo, M.J.; López-Herrera, A.G.; Herrera-Viedma, E.; Herrera, F. An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the fuzzy sets theory field. *J. Informetr.* **2011**, *5*, 146–166. [CrossRef]
30. Chakraborty, K.; Bhatia, S.; Bhattacharyya, S.; Platos, J.; Bag, R.; Hassanien, A.E. Sentiment Analysis of COVID-19 tweets by Deep Learning Classifiers—A study to show how popularity is affecting accuracy in social media. *Appl. Soft Comput.* **2020**, *97*, 106754. [CrossRef]
31. Boon-Itt, S.; Skunkan, Y. Public Perception of the COVID-19 Pandemic on Twitter: Sentiment Analysis and Topic Modeling Study. *JMIR Public Health Surveill.* **2020**, *6*, e21978. [CrossRef] [PubMed]
32. Park, E.; Kim, W.-H.; Kim, S.-B. Tracking tourism and hospitality employees' real-time perceptions and emotions in an online community during the COVID-19 pandemic. *Curr. Issues Tour.* **2020**. [CrossRef]
33. Xue, J.; Chen, J.; Hu, R.; Chen, C.; Zheng, C.; Su, Y.; Zhu, T. Twitter Discussions and Emotions About the COVID-19 Pandemic: Machine Learning Approach. *J. Med. Internet Res.* **2020**, *22*, e20550. [CrossRef] [PubMed]
34. Apuke, O.D.; Omar, B. Fake news and COVID-19: Modelling the predictors of fake news sharing among social media users. *Telemat. Inform.* **2021**, *56*, 101475. [CrossRef]
35. Li, L.; Zhang, Q.; Wang, X.; Zhang, J.; Wang, T.; Gao, T.-L.; Duan, W.; Tsoi, K.K.-F.; Wang, F.-Y. Characterizing the Propagation of Situational Information in Social Media During COVID-19 Epidemic: A Case Study on Weibo. *IEEE Trans. Comput. Soc. Syst.* **2020**, *7*, 556–562. [CrossRef]
36. Rufai, S.R.; Bunce, C. World leaders' usage of Twitter in response to the COVID-19 pandemic: A content analysis. *J. Public Health* **2020**, *42*, 510–516. [CrossRef] [PubMed]
37. Wang, Y.; Hao, H.; Platt, L.S. Examining risk and crisis communications of government agencies and stakeholders during early-stages of COVID-19 on Twitter. *Comput. Hum. Behav.* **2021**, *114*, 106568. [CrossRef] [PubMed]
38. Lin, C.-Y.; Broström, A.; Griffiths, M.; Pakpour, A. Investigating mediated effects of fear of COVID-19 and COVID-19 misunderstanding in the association between problematic social media use, psychological distress, and insomnia. *Internet Interv.* **2020**, *21*, 100345. [CrossRef] [PubMed]
39. Hayes, M. Social media and inspiring physical activity during COVID-19 and beyond. *Manag. Sport Leis.* **2022**, *27*, 14–21. [CrossRef]
40. Sobaih, A.E.E.; Hasanein, A.M.; Abu Elnasr, A.E. Responses to COVID-19 in Higher Education: Social Media Usage for Sustaining Formal Academic Communication in Developing Countries. *Sustainability* **2020**, *12*, 6520. [CrossRef]
41. Chew, C.; Eysenbach, G. Pandemics in the Age of Twitter: Content Analysis of Tweets during the 2009 H1N1 Outbreak. *PLoS ONE* **2010**, *5*, e14118. [CrossRef]

42. Abd-Alrazaq, A.; Alhuwail, D.; Househ, M.; Hamdi, M.; Shah, Z. Top Concerns of Tweeters During the COVID-19 Pandemic: Infoveillance Study. *J. Med. Internet Res.* **2020**, *22*, e19016. [CrossRef] [PubMed]
43. Chen, E.; Lerman, K.; Ferrara, E. Tracking Social Media Discourse About the COVID-19 Pandemic: Development of a Public Coronavirus Twitter Data Set. *JMIR Public Health Surveill.* **2020**, *6*, e19273. [CrossRef] [PubMed]
44. Depoux, A.; Martin, S.; Karafillakis, E.; Preet, R.; Wilder-Smith, A.; Larson, H. The pandemic of social media panic travels faster than the COVID-19 outbreak. *J. Travel Med.* **2020**, *27*, taaa031. [CrossRef] [PubMed]
45. Ahmad, A.; Murad, H. The Impact of Social Media on Panic During the COVID-19 Pandemic in Iraqi Kurdistan: Online Questionnaire Study. *J. Med. Internet Res.* **2020**, *22*, e19556. [CrossRef]
46. Web of Science Core Collection Database. Available online: https://www.webofscience.com/ (accessed on 11 June 2022).