Social Media Crowdsourcing a Co-Citation Bibliometric Analysis

Anna D. Kyosova

Graduate School of Technology Management, Ritsumeikan University, Osaka, Japan
Email: a.kyosova@gmail.com (A.D.K.)
Manuscript received October 10, 2023; revised November 9, 2023; accepted December 9, 2023; published November 13, 2024

***Abstract*—Social media crowdsourcing has rapidly expanded into multitudes of academic spheres. This article aims to create an exhaustive bibliometric analysis of the expanding field of social media crowdsourcing with the objective of mapping the scholarly impact, thematic concentrations, and intellectual structure of already existing literature and their fields of research. A comprehensive analysis of the literature on crowdsourcing and social media was conducted using advanced statistical methodologies and tools such as Bibliometrix and Biblioshiny. The publications closely related to these topics were extracted from the Web of Science database, and co-author, co-word, and co-citation analyses were performed. These findings were then visualized graphically. The study highlights the most impactful sources, articles and their authors as well as the frequently used terms and their relatedness in order to examine the past, current and future trends, with the ultimate goal of finding unexplored or under-explored niches. The examination of various factors revealed significant research gaps in both the geographical and platform aspects. The geographic analysis shed light on areas where there is a lack of research, indicating potential for further exploration into the utilization of social media in crowdsourcing in the countries where it is yet to be explored. Additionally, the overreliance on Twitter as a social media platform has in a way constrained research opportunities and directions using local less popular, but potentially more efficient for the gathering of information in the respective field.**

***Keywords*—bibliometric analysis, co-citation, crowdsourcing, social media**

I. Introduction

Crowdsourcing is a term coined by Jeff Howe in 2006 [1], referring to the practice of obtaining services, ideas, or content by seeking contributions from a large group of people via an open call.

The mechanics of a successful crowdsourcing campaign are heavily reliant on the concept of wisdom of the crowd and its collective intelligence, which allows a large group of people to make predictions or to provide insights and solutions, more accurately than individual experts [2].

Although it can be done by individually asking the opinions of a large group of participants through traditional means such as cold-calling, stopping passersby on the streets or hoping people open and answer an email survey from a random email, it can significantly be fastened by implementing modern technology with social media platforms. Social media has revolutionized the way crowdsourcing is conducted, allowing for efficient and widespread participation and solving geographic, time and language barriers that were previously limiting factors.

Due to its nature crowdsourcing can be implemented across a variety of disciplines, including but not limited to, science, technology, research, marketing, and social issues. Therefore, it is crucial to examine the effectiveness and impact of crowdsourcing multidimensionally to gain a comprehensive understanding of its applications in different fields.

In order to access multitudes of peer-reviewed articles in different disciplines, the Web of Science database was used for the search and retrieval of relevant literature. Followed by implementing a R-tool called Bibliometrix, and its platform Biblioshiny, for comprehensive mapping analysis and visualizing the bibliometric data [3].

This paper will first provide an overview of social media crowdsourcing and its significance in different fields. Then, it will implement co-citation biblio-metric analysis to analyze the scholarly impact and interconnections of social media crowdsourcing within the academic literature. Identifying the gap in the literature, this study aims to contribute to the understanding of the current state of knowledge and identify potential avenues for further research in the field of social media crowdsourcing.

II. Method and Data

The co-citation bibliometric analysis, which involves analyzing the citations within academic literature to identify patterns and relationships between research articles, has been used in this article.

Biblioshiny, an app for Bibliometrix, which is a specially designed package for bibliometric analysis in R [3]., the respective versions are R 4.2.3 and RStudio 2023.03.0. To access the Biblioshiny app the following script needs to be run in RStudio:

library(bibliometrix)
bibliometrix::biblioshiny().

To gather and retrieve the needed information, the Core Collection of the Web of Science database was used. The search term includes both the main keywords "social media" and derivatives and alternative spellings of "crowdsourcing", resulting in the following search:

(crowdsourc OR "crowd-sourc*"OR "collective intelligence" OR "collaborat* innovation") AND "social media"*

The initial search retrieved 1711 entries altogether, followed by a 3-step refining process: 1) only articles with the search terms mentioned in the abstract (945) and either in the title (112) or among the authors' keywords (291) were designed 2) selecting the full articles in English and 3) manual removal of duplicates resulted in 340 academic articles.

III. Initial Data Statistics

The overall composition of the 340 extracted publications,

covering 2010 until June 2023) is as seen on Table 1. The publications were authored by 1,141 authors, with only 37 authors working individually. The annual growth rate indicates that on average the number of documents, covering the topic of social media crowdsourcing, has been increasing by 11.25%. The average citations are 24.49 times and there are 12,846 references within this dataset. When it comes to the content of the documents, there are 557 unique keywords associated with the documents as well as 1,121 unique author-generated keywords.

Table 1. Main information

Description	Results
Timespan	2010:2023
Sources (Journals, Books, etc)	282
Documents	340
Annual Growth Rate %	11.25
Document Average Age	5.56
Average citations per doc	24.49
References	12846
Keywords Plus (ID)	557
Author's Keywords (DE)	1121
Authors	1141
Authors of single-authored docs	37
Single-authored docs	37
Co-Authors per Doc	3.77
International co-authorships %	30.29
Article	196
Article; book chapter	3
Article; early access	3
Proceedings paper	128
Review	10

The dataset is comprised of 196 articles, 3 articles that are also book chapters, 3 articles marked as early access, 128 proceedings papers, and 10 reviews.

The number of works covering social media crowdsourcing in our dataset by the year of their publication can be seen in Fig. 1. The number of publications shows a rapid increase from the starting year of 2010 until 2016 illustrating the growing interest in the subject, afterwards the production plateaus with the only notable exception being year 2019 when a record number of works were published. It should be noted that 2023 is incomplete in our dataset, therefore the smaller number should not be seen as a downward trajectory in research.

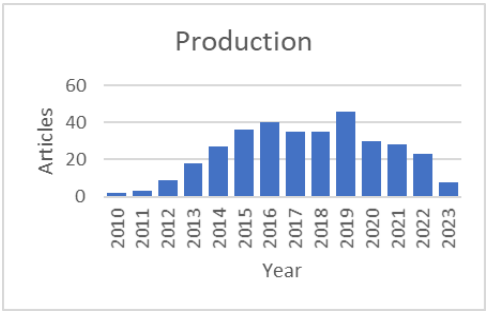

Fig. 1. Number of publications related to crowdsourcing and social media.

IV. Sources

A. Most Relevant Sources

Building on the overview, an analysis was conducted to identify the key sources that have significantly influenced research in this field. Utilizing the appropriate tools for bibliometric analysis such as Biblioshiny and the Bibliometrix R package, a comprehensive examination of relevant sources was performed [3]. This rigorous methodology enabled us to identify the most prominent academic references that played vital roles in shaping the intellectual landscape of this domain.

Table 2. Most relevant sources

Sources	Articles	H index
International journal of disaster risk reduction	9	7
Isprs international journal of geo-information	8	6
Journal of medical internet research	5	3
Sustainability	5	3
Computers environment and urban systems	4	4
Ecosystem services	4	2
Computers in human behavior	3	3
Frontiers in public health	3	1
Www'15 companion: proceedings of the 24th international conference on world wide web	3	3
3rd international workshop on social sensing (socialsens 2018)	2	3

Table 2 highlights the key journals and the area allowing for the areas of research to emerge and it becomes evident the prevalence of disaster risk reduction, encompassing fields such as geography, healthcare, technology, and sustainability. To enhance the comprehensiveness of our analysis, we have integrated the h-index into the table. The h-index is a measure that evaluates both the productivity and citation impact of published works. Its addition offers a nuanced insight into each source's localized influence. Out of the top 10 most productive articles, "Frontiers in Public Health" is an outlier among the journals, as it does not appear on the list of most influential publications determined by h-index. This contrast highlights the importance of distinguishing between mere quantity and meaningful impact. Based on the data, we can clearly see that the emerging dominant topics are in the areas of disaster management, technology and health.

B. Most Local Cited Sources

This section comprehensively analyzes the most frequently cited local sources, providing a detailed understanding of how they influence different areas of academia.

Table 3. Most cited local sources

Sources	Number of Articles
PLOS ONE	132
Lecture Notes in Computer Science	100
Science	96
Proceedings of the National Academy of Sciences USA	83
Computers in Human Behavior	79
Environmental Modelling & Software	75
Government Information Quarterly	75
International Journal of Disaster Risk Reduction	74
GeoJournal	73
ISPRS International Journal of Geo-Information	73
International Journal of Information Management	70

The diversity of themes depicted in Table 3 is astounding. It showcases a wide range of scientific applications of social media crowdsourcing from computer science journals such as "Lecture Notes in Computer Science", boasting 100 articles, to environmental studies such as "Environmental Modelling

& Software," featuring 75 pieces and government and public administration represented by "Government Information Quarterly", having also published 75 works. Specialized research areas are represented in journals like the "International Journal of Disaster Risk Reduction" and "GeoJournal", boasting 74 and 73 articles respectively. However, with a remarkable 132 articles, "PLOS ONE" with its diversity in subject matter claims the top spot on the list and solidifies its impact on academia.

The significant number of citations these publications have received suggests a rising significance for fields such as geographical studies and disaster risk reduction.

C. Sources with the highest Local Impact

An important aspect of evaluating academic sources is their local impact, which sheds light on how influential they are within particular research circles. The local impact of the local sources, as seen in Table 4, is measured using several bibliometric indicators such as h-index, g-index, m-index, total citations (TC), number of papers published (NP) and the year of publication (PY start).

Table 4. The most impactful local sources

Sources	h-index	g-index	m-index	TC	NP	PY_start
International Journal of Disaster Risk Reduction	7	9	0.875	351	9	2016
ISPRS International Journal of Geo-Information	6	8	0.600	128	8	2014
Computers, Environment and Urban Systems	4	4	0.500	148	4	2016
Computers in Human Behavior	3	3	0.273	1208	3	2013
Journal of Medical Internet Research	3	5	0.600	46	5	2019
Sustainability	3	5	0.375	48	5	2016
WWW'15 Companion: Proceedings of the 24th International Conference on World Wide Web	3	3	0.333	47	3	2015
3rd International Workshop on Social Sensing (SocialSens 2018)	2	2	0.333	7	2	2018
Advanced Engineering Informatics	2	2	0.333	11	2	2018
Computer Supported Cooperative Work-The Journal of Collaborative Computing and Work Practices	2	2	0.200	103	2	2014

The impact and productivity of a source can be effectively evaluated through the h-index or g-index. Table 5 reveals that with an h-index of 7 and a g-index of 9, the "International Journal of Disaster Risk Reduction" demonstrates remarkable levels both in terms of impact and productivity. Its m-index score stands at 0.875 alongside its total citations count - which has soared up to about 351 since inception in the year 2016- further corroborating its profound influence on current literature surrounding disaster risk reduction. The journals "Computers, Environment and Urban Systems" as well as the "ISPRS International Journal of Geo-Information" highlight the increasing significance of niche publications. Even though they possess comparatively lower h-indexes, their combined citation counts of 148 and 128 respectively. Despite having a remarkable 1208 overall citations, "Computers in Human Behavior" exhibits a rather humble h-index of 3. This indicates that although the journal may contain some articles with significant recognition, its comprehensive influence evaluated by the h-index remains confined. "Journal of Medical Internet Research" and "3rd International Workshop on Social Sensing", despite their lower h-indices, display potential due to their recent establishment in 2019 and 2018 respectively.

V. Countries

The location of corresponding authors in scholarly publications presents an opportunity to explore the worldwide research panorama. This part endeavors to deliver a thorough evaluation of which countries are most commonly represented as corresponding authors, assessing not only article quantity but also collaborative tendencies. Fig. 2 is a visualization of not only article quantity, but also collaborative tendencies denoted by Single Country Publications (SCP) and Multiple Country Publications (MCP).

With 99 articles, a frequency of 0.291 and an MCP ratio of 0.202, the United States is the undisputed leader in the field. However, despite its dominant volume output in research papers, it exhibits comparably lower levels of global cooperation (0.202) with other nations. In comparison, China exhibits a robust research output with 27 articles, but it also boasts of a substantial international collaboration evidenced by its MCP ratio of 0.407. Both Australia and Germany have made comparable contributions with 24 and 18 articles, respectively. Nevertheless, the MCP ratio of Australia (0.167) is lower than that of Germany (0.278), indicating a greater emphasis on domestic research for Australia while highlighting Germany's inclination towards international partnerships. The other members of the European Union, Italy, Greece and Spain, also follow Germany's trend by exhibiting a greater inclination towards international collaboration. Their MCP ratios range from 0.364 to 0.455.

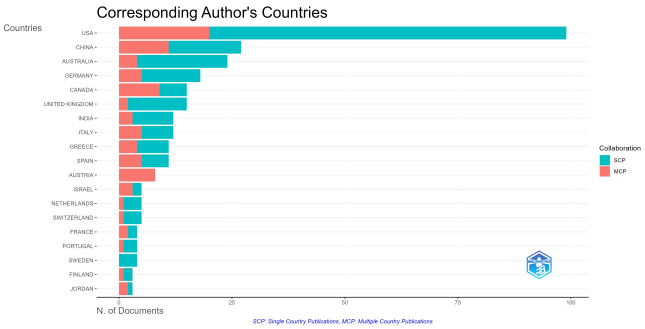

Fig. 2. Top countries by production and collaboration, generated via Biblioshiny.

Despite its lower overall output, Canada distinguishes itself as a center for international collaboration with 15 articles and an MCP ratio of 0.600. The UK exhibits a domestic focus with an MCP ratio of 0.133, whereas India demonstrates more balanced proportions with an MCP ratio of 0.250.

While countries such as China and the USA stand out for their volume, Canada and various European nations shine with regards to international partnerships. For researchers looking to form global connections or policymakers working

towards encouraging cross-border collaborations between academics, grasping these patterns is key.

Fig. 3 not only shows the collaboration links, but it also allows us to view which countries and areas haven't participated in the academic discord on the subject. Those yet unexplored locations may become a great inspiration for future research.

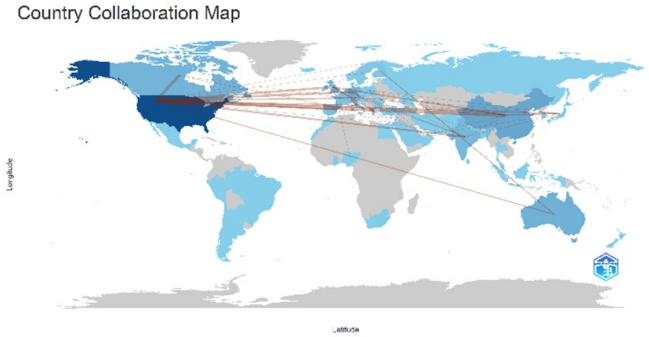

Fig. 3. Research collaboration by country, generated via Biblioshiny.

VI. References

A. Most Locally cited References

Table 5 identifies the most cited authors in the field of social media crowdsourcing as well as the citation number in our dataset. Those are the authors who have greatly contributed to shaping the academic world's understanding in the area of crowdsourcing and social media. Unsurprisingly, Howe and Suroweicki stand out as the ultimate sources of intellectual influence, with their fundamental theories receiving widespread citations. Howe's concept of "crowdsourcing" has provided a basis for further research, earning 39 mentions in academic literature [1]. Similarly, Suroweicki's innovative exploration into "collective intelligence" has been cited on 32 occasions - demonstrating its crucial role in comprehending crowdsourcing mechanics [2]. Estelles et al. is another highly cited reference, the Spanish authors were cited 32 times in our dataset, the article provides valuable insights into the different academic definitions of crowdsourcing research [4]. On the other hand, Braham (21 citations) focuses on the practical applications and real-world of crowdsourcing in public health science [5].

Table 5. Most locally cited references

Reference	Citations	Country
Howe, J. (2006).	39	US
Estelles Arolas E., González-Ladrón-De-Guevara, F. (2012).	32	Spain
Surowiecki, J. (2004).	32	US
Goodchild M.F. (2007).	31	US
Brabham D. C. (2008).	21	US
Howe, J. (2008).	21	US
Kaplan, A. M., & Haenlein, M. (2010).	19	Germany
Sakaki, T., Okazaki, M., & Matsuo, Y. (2010).	16	Japan
Doan, A., Ramakrishnan, R., & Halevy, A. Y. (2011).	16	US
Goodchild, M. F., & Glennon, J. A. (2010).	16	US

With 19 and 16 citations, Kaplan and Doan have contributed to the academic landscape by exploring technology's role. In particular, Kaplan investigates how social media platforms spur crowdsourcing [6] while Doan takes a systemic viewpoint in discussing Internet-based crowdsourced systems' architectures [7]. Goodchild has entered the top 10 most cited references, with 2 works focused on gathering geographic data as well as the gathering and using of geographic data for disaster management [8, 9]. Sakaki et al, the only entry from Japan, also focus on disaster response by crowdsourcing through Twitter [10].

The table also shows the significant domination of the US authors 7 out of 10, which could be attributed to the fact that they were the ones to coin and define the discipline.

B. Co-Citation Network in Social Media Crowdsourcing Research

Visualizing the concentration and interconnectedness between the references in academic literature can be effectively seen in Fig. 4 where the primary objective is to comprehensively show dominant scholarly discourse themes by evaluating their relative significance and how they are interconnected with each other within clusters.

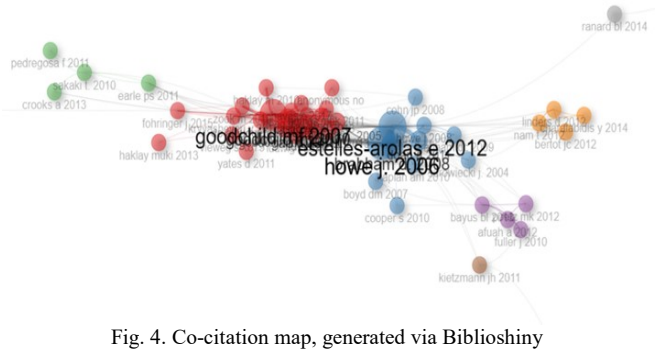

Fig. 4. Co-citation map, generated via Biblioshiny.

The authors with high interconnectedness are illustrated by the following bigger nodes. For example, Howe J.'s [1] report, shows that the work serves as an important hub in the network due to its high Betweenness score of 172.25, which facilitates connections between different thematic clusters. Additionally, with a PageRank value of 0.05 also highlights its significance within the larger context. Estelles-Arolas' *et al.* [4] is also vital in connecting various parts of the network, as evidenced by its Betweenness rating of 127.32. The two publications also have high PageRank nodes of 0.05 and 0.045 respectively. Brabham [5], on the other hand, exhibits high closeness but low moderate betweenness of 62.72, but the Closeness score of 0.0079 indicates that it has numerous close connections with other nodes within its cluster, giving it a pivotal role in the network.

Surowiecki's [2] has a Betweenness of 0 but a Closeness of 0.0029, indicating that while it doesn't serve as a bridge, it is not isolated.

A nuanced landscape of interconnectedness and influence among seminal works in crowdsourcing research is revealed by the co-citation network analysis. The pivotal nodes are represented by the works of Howe [1] and Estelles-Arolas *et al.* [4], signifying their foundational position within this domain.

VII. Co-Author Analysis

A. Most Relevant Authors and Co-Authors

Table 6 lists the most relevant authors within the dataset, based on the number of articles they have published and their contributions in co-authored publications, reflected by the "Articles Fractionalized" category.

Table 6: Most relevant authors

Authors	Articles	Articles Fractionalized
Ghermandi A	6	1.92
Yigitcanlar T	5	1.21
Kankanamge N	4	0.96
Mei L	4	0.63
Pernici B	4	0.46
Sinclair M	4	1.42
Starbird K	4	1.07
Xu Z	4	0.63
Agarwal N	3	1.00
Charalabidis Y	3	0.92

Ghermandi A. is the most prolific author and co-author on the list, with 6 articles published offers a new perspective to crowdsourcing for geolocational data [11–15]. This author has a fractionalized article score of 1.92, indicating a significant contribution to the field. Yigitcanlar T follows closely behind with 5 articles and a fractionalized value of 1.21 [16–18]. The majority of the other top authors have written 4 articles or less on the subject of crowdsourcing and social media. The data illustrates that the authors with the most publications are also the ones leasing the industry in co-authorship, thereby increasing their influence in the sphere.

B. Authors' Production over Time

It is essential to comprehend the academic input and influence of authors across time in a particular research field for assessing its development. Fig. 5 illustrates an all-encompassing evaluation of production metrics pertinent to scholars, such as publication frequency, total citations (TC), and yearly citation count per paper (TCpY) throughout several years.

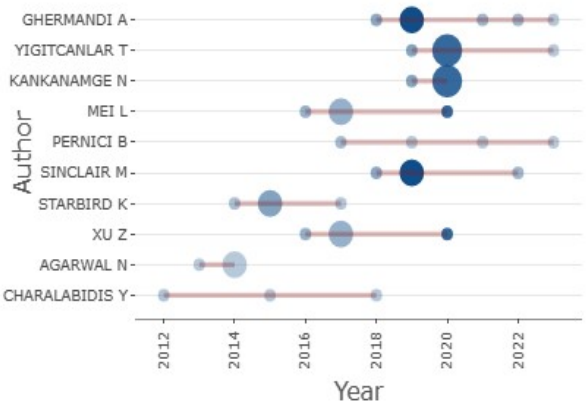

Fig. 5. Authors production over time, generated via Biblioshiny.

Ghermandi heads the top most productive authors in our dataset with the highest performance in 2019 [11–15]. Nevertheless, it appears that the influence dwindles over consecutive years. Both Kankanamge and Yigitcanlar [16–18], have exhibited a noteworthy surge in their citation counts per paper from 2019 to 2020, suggesting an expanding impact within their respective domains. Meanwhile, Mei and Xu have maintained a steady level of impact over time, with a significant spike in 2020 [19–21]. Sinclair exhibits the same trends as Ghermandi A. [12–15].

Starbird has demonstrated a fluctuating impact over time, reaching its pinnacle in 2015 [22]. Agarwal's scholarly activity in 2013 and 2014 was also notable, however, their works did not receive any citations indicating that their impact is yet to be recognized.

C. Lotka's Law

Bibliometrics requires a comprehensive grasp of how scholarly output is distributed among authors. Lotka's Law acts as the fundamental framework for investigating this distribution by suggesting that a few writers account for most publications. Fig. 6 illustrates Lotka's Law which predicts that most authors will make minimal contributions, which is confirmed by the fact that an astounding 90.6% of writers have only submitted one document.

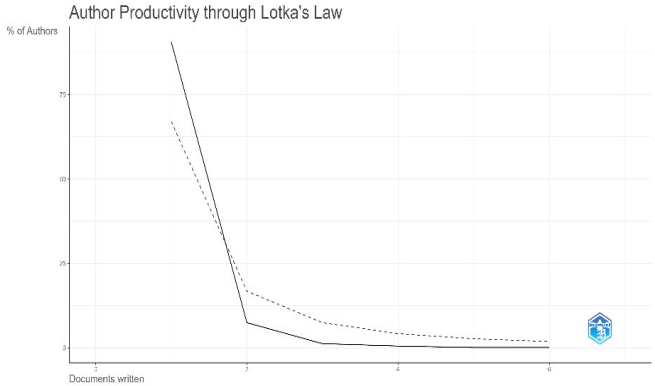

Fig. 6. Authors productivity through Lotka's law, generated via Biblioshiny.

The group of authors who have authored two or three documents accounts for 7.4% and 1.2% of the overall tally, respectively. Despite being the smallest demographic the researchers with multiple works are a vital subset that plays an important role in grasping the dynamics surrounding academic output distribution.

D. Authors' Local Impact

An important measure that sheds light on a scholar's influence is their local impact within academic communities (see Table 7).

Ghermandi exhibits remarkable scholarly impact and productivity, as evidenced by an h-index of 5 and a g-index of 6. The m-index value at 0.833, along with the total citations count of 288 since 2018 further substantiate their substantial influence in the field. Mei and Xu have an h-index and g-index score of 4, with a combined citation count of 244 [19–21]. Their m-indices at 0.500 demonstrate consistent impact from the year 2016 till now. Emerging talents in their respective fields, Kankanamge and Yigitcanlar began making an impact starting from 2019 [16–18]. Their promising potential is demonstrated by the h-indices of 4 for Kankanamge and g-indices ranging between 4 to 5 for

Yigitcanlar. Although Starbird's h-index and g-index are similar to many of their peers at 4, they have a comparatively lower total citation count of 152 [22]. This implies that while the author's work has significance, it may not be as extensively referenced as others in their field. Goonetilleke, Hu, and Liu exhibit consistent contributions to their respective fields [16–21]. Their m-indices span between 0.375 and 0.600 which suggests a moderate yet steady level of influence.

Table 7. Influence of authors within the field

Sources	h-index	g-index	m-index	TC	NP	PY_start
Ghermandi A	5	6	0.833	288	6	2018
Kankanamge N	4	4	0.800	181	4	2019
Mei L	4	4	0.500	244	4	2016
Sinclair M	4	4	0.667	269	4	2018
Starbird K	4	4	0.400	152	4	2014
Xu Z	4	4	0.500	244	4	2016
Yigitcanlar T	4	5	0.800	181	5	2019
Goonetilleke A	3	3	0.600	153	3	2019
Hu CP	3	3	0.375	197	3	2016
Liu YH	3	3	0.429	182	3	2017

E. Co-Citation Network in Social Media Crowdsourcing Research

An effective way to study the intellectual terrain of a scientific field is through co-citation networks. Fig. 7 represents a co-citation network map of the most prominent authors from our dataset by visualizing their influence and interconnectivity within the research in the area of social media crowdsourcing.

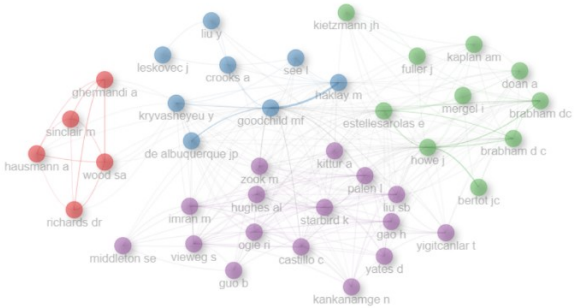

Fig. 7. Co- citation network of the most influential researchers in our dataset, generated via Biblioshiny.

Although the co-citation network is complex and interconnected, the pivotal node appears to be Goodchild MF 2007, revealing its foundational significance within the field, exhibiting both a high betweenness (144.42) and high PageRank (0.082), indicating that this work has a high level of influence and plays a crucial role in the network by serving as a significant link between diverse thematic clusters [8]. Goodchild MF's work in 2010 also plays a significant role in linking sub-themes within the cluster as evidenced by its moderate Betweenness value of 12.84 and PageRank of 0.055 [9].

VIII. Co-Word Analysis

A. Most Frequent Words

Fig. 8 is a visual representation of the frequency and prevalence of various terms. As expected, social media (26 occurrences) is the most popular term. The prevalence of "social media" indicates the significant role of social platforms in diverse research settings. The fact that "Twitter" is mentioned 20 times indicates the platform's importance in research, particularly pertaining to fields such as social networking, public opinion, and data analysis platforms such as Twitter (20 occurrences) in the collection and dissemination of geographic information. This suggests that social media has been used primarily as a source of location data and engagement.

Fig. 8. Word cloud of the most prevalent words, generated via Biblioshiny.

The term "Participation" has appeared 13 times, indicating an increased focus on community involvement, democratic procedures and user engagement in the design of systems. With 12 instances, the term "Volunteered Geographic Information" highlights a growing emphasis on voluntarily provided geospatial data that is frequently useful in fields such as environmental research, urban planning and disaster response.

The occurrence of the terms Networks and Communication, which appear 12 and 11 times respectively, may indicate an interdisciplinary connection that combines social sciences with computer science, engineering or information systems.

Co-creation, Crisis and Cultural Ecosystem Services - have been mentioned five times each. It suggests that these topics are gaining academic attention as specialized areas of study which may represent niche subjects.

B. Word Clusters

By utilizing advanced clustering techniques, it is possible to gain a better understanding of the complex connections among significant concepts in academic literature. The "Clustering by Coupling" approach used in Fig. 9 to show the interrelationships based on measures like frequency, centrality, and impact.

The dataset comprises labels that indicate the terms and their confidence levels, along with the clusters, their frequency, centrality and impact. Cluster 5 holds paramount significance in shaping academic discourse, as evidenced by its highest impact score of 2.102 and a substantial centrality value of 0.5294; it encompasses terms such as "cultural ecosystem services" and "spatial patterns. With terms such as "information," "social media," and "twitter" at its nucleus, Cluster 6 boasts an impact score of 2.451; however, it exhibits a marginally lower centrality of 0.5064 that indicates widespread yet not completely centralized influence. Cluster 7, marked by its, includes terms such as "continued influence" and "fake." Although it has a lower frequency and impact

compared to other clusters, it is notable for its thematic distinctiveness. The fourth cluster, which contains words such as "social media," "participation," and "networks," is highly frequent with a score of 102. Its centrality metric of 0.3621 suggests that it's both well-connected and influential. Cluster 3 is characterized by "mouth," "quality-of-life," and "sales" terms. Its frequency may be moderate, but with a centrality of 0.2939, it appears as a specialized area with particular attention to its specific focus.

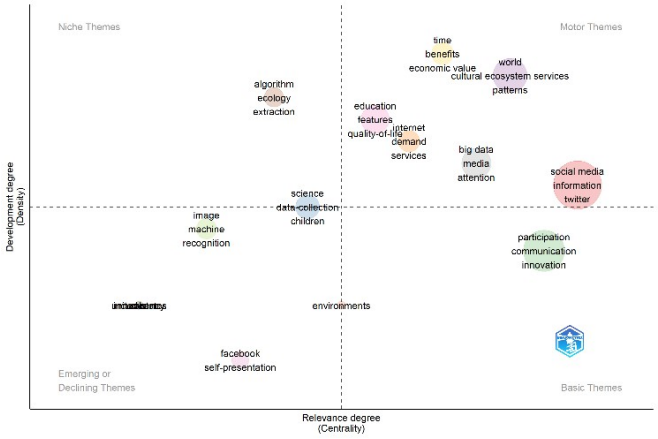

Fig. 9. Word clusters, generated via Biblioshiny.

With the "Clustering by Coupling" technique, one can gain a comprehensive understanding of the diverse thematic clusters forming academic research. This method enables the categorization of these clusters according to their impact level, emergence status, coupling strength and specialization areas; each has its own significant influence on scholarly pursuit.

C. Co-Word Map

Analyzing the most common words offers a subtle comprehension of the subjects at hand and develops patterns in scholarly topics. It acts as an indicator to evaluate which fields are currently receiving academic scrutiny, along with ones that could benefit from more investigation. Fig. 10 helps visualize the correlation between the most popular terms.

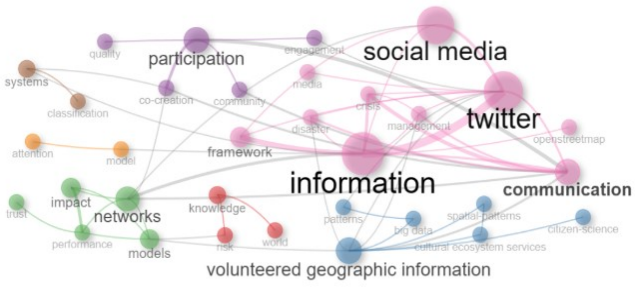

Fig. 10. Word co-relation, generated via Biblioshiny.

It is crucial to note that although Twitter has a prominent place among the most frequently used words, there is a significant lack in the utilization of other social media platforms.

IX. Limitations

The comprehensive bibliometric analysis of social media crowdsourcing presented in the study offers valuable insights into its academic landscape. Although the study has conducted a comprehensive analysis and identified the major contributors, influential sources, and emerging trends in the field, it is important to mention some of the limitations.

The analysis focuses solely on English language publications, which may limit our exploration of contributions from non-English speakers in this domain.

The scope of this research is also constrained to the Web of Science database, which may not include all relevant literature in this field. Additionally, the exploratory research only includes publications until June 2023 and may not cover the most recent advancements.

X. Conclusion

This comprehensive study explored the complex terrain of crowdsourcing on social media, utilizing a meticulous analysis of bibliometric data to uncover its academic influence, thematic concentration, and intellectual framework. Our assessment was supported by cutting-edge statistical techniques and methodologies such as Bibliometrix and Biblioshiny platforms, co-citation networks, Lotka's Law along with various others.

The study's results indicate that the field displays considerable dynamism and cross-disciplinary involvement. In terms of research production, the United States leads as a major contributor; however, it trails other countries in international partnerships - an area where China and Canada thrive. Notably, two publications ("International Journal of Disaster Risk Reduction" and "ISPRS International Journal of Geo-Information") have played pivotal roles in shaping discussions on disaster management and geographic information systems.

The analysis of the network of co-citations emphasizes the crucial contributions made by Howe and Suroweicki, whose ideas on collective intelligence and crowdsourcing provide essential intellectual foundations. Nonetheless, it is worth noting how authors such as Ghermandi and Yigitcanlar are increasingly influential in academia; despite a short period dedicated to scholarly work, both have impressively high h-indices and total citation counts that demonstrate consistent impact over time.

Lotka's Law validates that the distribution of academic output is imbalanced, whereby a limited group of authors are responsible for a substantial portion of scholarly work. This finding goes beyond mere statistics and sheds light on how the field progresses over time; influential thought leaders can strongly steer research direction.

Although the field has experienced strong growth and greater complexity, there are still gaps in both subject matter and location which provide potential for future investigation. Such inadequacies raise concerns about how complete and inclusive current scholarly discussions truly are.

To sum up, the scholarly landscape of social media crowdsourcing is intricate but predictable as a rapidly growing field. Its practical use cases are diverse and relevant across different fields including healthcare and disaster risk mitigation. It holds universal importance for research in this era of complexity that calls for continuous evaluation to meet contemporary challenges while it develops further collaborative networks with experts within its domain.

The study adds depth to the current knowledge of social media crowdsourcing research by combining complex layers of information, providing a sophisticated understanding. It also establishes an essential groundwork for future scholarly studies.

CONFLICT OF INTEREST

The author declares no conflict of interest.

ACKNOWLEDGMENT

The author expresses gratitude to Professor Tetsuaki Oda for his invaluable advice, guidance, and understanding.

REFERENCES

[1] J. Howe, "The rise of crowdsourcing," *Wired*, vol. 14, 2006.
[2] J. Surowiecki, *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies, and Nations*, Doubleday & Co., 2004.
[3] M. Aria and C. Cuccurullo, "Bibliometrix: An R-tool for comprehensive science mapping analysis," *Journal of Informetrics*, vol. 11, no. 4, pp. 959–975, 2017.
[4] E. Estellés-Arolas and F. González-Ladrón-De-Guevara, "Towards an integrated crowdsourcing definition," *Journal of Information Science*, vol. 32, no. 2, pp. 189–200, 2012.
[5] D. C. Brabham, "Crowdsourcing as a model for problem solving," *Convergence*, vol. 14, no. 1, pp. 75–90, 2008. doi:10.1177/1354856507084420
[6] A. M. Kaplan and M. Haenlein, "Users of the world, unite! The challenges and opportunities of Social Media," *Business Horizons*, vol. 53, no. 1, pp. 59–68, 2010.
[7] A. Doan, R. Ramakrishnan, and A. Y. Halevy, "Crowdsourcing systems on the World-Wide Web," *Communications of the ACM*, vol. 54, no. 4, p. 86, 2011.
[8] M. F. Goodchild, "Citizens as sensors: The world of volunteered geography," *GeoJournal*, 2007, vol. 69, pp. 211–221. doi:10.1007/s10708-007-9111-y
[9] M. F. Goodchild and J. A. Glennon, "Crowdsourcing geographic information for disaster response: a research frontier," *International Journal of Digital Earth*, vol. 3, no. 3, pp. 231–241, 2010.
[10] T. Sakaki, M. Okazaki, and Y. Matsuo, "Earthquake shakes twitter users: Real-time event detection by social sensors," in *Proc. the Nineteenth International WWW Conference*, pp. 851–860, 2010.
[11] A. Ghermandi, "Geolocated social media data counts as a proxy for recreational visits in natural areas: A meta-analysis," *Journal of Environmental Management*, vol. 317, 2022.
[12] M. Sinclair, A. Ghermandi, S. A. Moses, and S. Joseph, "Recreation and environmental quality of tropical wetlands: A social media based spatial analysis," *Tourism Management*, vol. 71, pp. 179–186, 2019.
[13] A. Ghermandi and M. Sinclair, "Passive crowdsourcing of social media in environmental research: A systematic map," *Global Environmental Change-Human and Policy Dimensions*, vol. 55, pp. 36–47, 2019.
[14] A. Ghermandi, Y. Depietri, and M. Sinclair, "In the AI of the beholder: A comparative analysis of computer vision-assisted characterizations of human-nature interactions in urban green spaces," *Landscape and Urban Planning*, vol. 217, 2022.
[15] A. Ghermandi, M. Sinclair, E. Fichtman, and M. Gish, "Novel insights on intensity and typology of direct human-nature interactions in protected areas through passive crowdsourcing," *Global Environmental Change-Human and Policy Dimensions*, vol. 65, 2020.
[16] N. Kankanamge, T. Yigitcanlar, and A. Goonetilleke, "How engaging are disaster management related social media channels? The case of Australian state emergency organisations," *International Journal of Disaster Risk Reduction*, vol. 48, 2020.
[17] N. Kankanamge, T. Yigitcanlar, and A. Goonetilleke, "Public perceptions on artificial intelligence driven disaster management: Evidence from Sydney, Melbourne and Brisbane," *Telematics and Informatics*, vol. 65, 2021.
[18] N. Kankanamge, T. Yigitcanlar, A. Goonetilleke, and M. Kamruzzaman, "Can volunteer crowdsourcing reduce disaster risk? A systematic review of the literature," *International Journal of Disaster Risk Reduction*, vol. 35, 2019.
[19] Z. Xu, Y. H. Liu, J. Y. Xuan, H. Y. Chen, and L. Mei, "Crowdsourcing based social media data analysis of urban emergency events," *Multimedia Tools and Applications*, vol. 76, no. 9, 2017.
[20] Z. Xu, Y. H. Liu, N. Y. Yen, L. Mei, X. F. Luo, X. Wei, and C. P. Hu, "Crowdsourcing based description of urban emergency events using social media big data," *IEEE Transactions on Cloud Computing*, vol. 8, no. 2, 2020.
[21] Z. Xu, H. Zhang, C. P. Hu, L. Mei, J. Y. Xuan, K. K. R. Choo, V. Sugumaran, and Y. W. Zhu, "Building knowledge base of urban emergency events based on crowdsourcing of social media," *Concurrency*, 2016.
[22] D. Dailey and K. Starbird, "Journalists as crowdsourcerers: responding to crisis by reporting with a crowd," *Computer Supported Cooperative Work-the Journal of Collaborative Computing and Work Practices*, vol. 23, no. 4–6, 2014.