Quality & Quantity
https://doi.org/10.1007/s11135-023-01626-7

Twenty years of Wikipedia in scholarly publications: a bibliometric network analysis of the thematic and citation landscape

Mohamed M. Mostafa[1]

Accepted: 3 February 2023

Abstract
Wikipedia has grown to be the biggest online encyclopedia in terms of comprehensiveness, reach and coverage. However, although different websites and social network platforms have received considerable academic attention, Wikipedia has largely gone unnoticed. In this study, we fill this research gap by investigating how Wikipedia is used in scholarly publications since its launch in 2001. More specifically, we review and analyze the intellectual structure of Wikipedia's scholarly publications based on 3790 Web of Science core collection documents written by 10,636 authors from 100 countries over two decades (2001–2021). Results show that the most influential outlets publishing Wikipedia research include journals such as *Plos one, Nucleic Acids Research, the Journal of the Association for Information Science and Technology, the Journal of the American Society for Information Science and Technology, IEEE Access,* and *Information Processing and Management*. Results also show that the author collaboration network is very sparsely connected, indicating the absence of close collaboration among the authors in the field. Furthermore, results reveal that the Wikipedia research institutions' collaboration network reflects a North–South divide as very limited cooperation occurs between developed and developing countries' institutions. Finally, the multiple correspondence analysis applied to obtain the Wikipedia research conceptual map reveals the breadth, diversity, and intellectual thrust of the Wikipedia's scholarly publications. Our analysis has far-reaching implications for aspiring researchers interested in Wikipedia research as we retrospectively trace the evolution in research output over the last two decades, establish linkages between the authors and articles, and reveal trending topics/hotspots within the broad theme of Wikipedia research.

Keywords Wikipedia · Bibliometric networks · Intellectual structure · Keyword co-occurrence · Historiography

✉ Mohamed M. Mostafa
mostafa@usa.com

[1] Gulf University for Science and Technology, West Mishref, Kuwait

Published online: 14 February 2023

1 Introduction

Wikipedia was launched in 2001 and since then it has grown to include more than fifty-five million articles in around 300 languages, making it one of the most-widely visited websites around the globe ahead of Reddit, Netflix and Instagram (Gobel and Munzert 2018; The Economist 2021). Each month, Wikipedia welcomes 400 million unique visitors, nearly half of whom visit the English edition (Navarrete and Borowiecki 2016). Wikipedia is considered the largest multilingual web encyclopedia, where a large number of common users regularly update and add new articles. Azad and Deepak (2019, p. 150) argued that "the exponential growth and reliability of Wikipedia make it an ideal knowledge resource for information retrieval." Due to its tremendous growth over time, the encyclopedia is structured as a complex network of articles. In this interconnected network, each article focuses on a single concept and includes hyperlinks connecting to other articles either inside or outside the encyclopedia. Hyperlinks are designed to direct the interested reader to more in-depth descriptions of the major notions mentioned in the articles. Wikipedia hyperlinks have been constantly validated for semantic similarity (Kim et al. 2019). In fact, Wu and Cunningham (2015, p. 474) argued that "in contrast to traditional encyclopedias, where authority derives from expert contributors, Wikipedia depends on collaboration and consensus to produce quality articles." This collaboration is based on the famous "wisdom of the crowds" principle, under which groups of ordinary mass can be "remarkably intelligent" (Surowiecki and Silverman 2007). Thus, Wikipedia has been regarded as "the Library of Alexandria of our time, or as close as we have gotten to it since the time of the grand Egyptian library" (Ezell Forthcoming, p. 4). Wikipedia's large-scale open-collaboration online innovative production system makes understanding its bibliometric network important in its own right.

It has been argued that as scientific production evolves in a research field, it becomes important for researchers to apply quantitative methodologies to understand the "intellectual structure" of the field (Rivera and Pizam 2015). In a similar vein, Portugal-Ferreira et al. (2014, p. 1902) noted that as fields of research mature and become more complex, researchers should aim to "periodically seek some sense of the knowledge produced and accumulated, to identify novel contributions, detect trends and research traditions, understand which topics are addressed, the theories and methods employed, delve into the intellectual structure of the discipline and its knowledge base, and prospect areas of future inquiry." Mapping an entire area of scientific knowledge is known as "scientography" (Francisco et al. 2019). This fairly new area of research aims at revealing "hidden patterns" in the development and structure of scientific knowledge in a graphical form (Qi et al. 2018). Although such landscape mapping shares some intrinsic semantic characteristics with cartography (Chen 2002; Chen and Paul 2001), it depends heavily on information visualization technology, high-performance computers and sophisticated digital databases to track disciplinary development status, research hotspots and research frontiers (Zou et al. 2018).

However, although web information has been investigated extensively (Cai et al. 2010; Leong et al. 2012; Wallace 2018; Kosterich and Weber 2019), no previous research has focused solely on analyzing Wikipedia's bibliometric network. Notable exceptions include a couple of outdated studies on Wikepedia (Jullien 2012; Okoli 2009). This is surprising given the fact that bibliometric analyses have recently been applied to investigate the knowledge structure in several research domains (Block et al. 2020; Bouzembrak et al. 2019; de la Hoz-Correa et al. 2018; Francisco et al. 2019; Gaede and Rowlands 2018; Gupta et al. 2021; Qin

et al. 2020; Ruiz-Alba et al. 2021; Vila-Lopez and Kuster-Boluda 2021; Zhu and Hua 2017). Thus, by investigating Wikipedia's scholarly publications' bibliometric networks, we argue that we make multiple contributions to the existing literature. First, we believe that this is the first comprehensive study to investigate Wikipedia's scholarly publications since its launch on January 15, 2001. By examining the structure and dynamics of Wikipedia's research, we contribute to the wider discussion on online information brokerage/gatekeeping behavior (DiMaggio et al. 2001; Vos and Heinderyckx 2015). Second, we believe that by investigating Wikipedia's scholarly publications, we add depth to the knowledge base on online-based research. In fact, Levitt (2012) argued that researchers investigating bibliometric research should focus on linkages and networks rather than bounded offline sites. Finally, by focusing solely on analyzing Wikipedia's scholarly publications data, rather than on traditional offline encyclopedias, we enrich the knowledge base of this under-represented area. More specifically, in this investigation we aim to find an answer to the following major research questions:

RQ1 How Wikipedia's scholarly research evolved since its emergence in 2001?
RQ2 Which Wikipedia authors and articles are most influential?
RQ3 What are the current thematic trends/hotspots of Wikipedia's scholarly research?
RQ4 What are the major countries/regions forming the geographic atlas of Wikipedia's scholarly research?
RQ5 What type of collaborative relationship does exist among authors, institutions and nations producing Wikipedia's scholarly research?

To find answers to our research questions, we apply a novel method based on data mining and social network analysis (SNA) in order to examine Wikipedia's scholarly publications, its network structure, its relationship properties, its topics pattern of communications and its different communities. Several scholars have recently advocated the need for adopting a "relational turn" in investigating online user-generated content (Wallace 2018, p. 218). The imperative to adopt a relational turn is theoretically based on the SNA paradigm, which argues that there is a need to move from methodologies that focus on attributive categorization into socio-economic status, race, etc. to focus rather on social ties among social actors (Wasche et al. 2017). In response to this call, we apply bibliometric and SNA techniques, which are reliable and flexible quantitative methodologies that have been found to be useful in uncovering and mapping out structural attributes of communication networks as well as illuminating how information is exchanged (Scott 2013). By implementing such approach in this study, we emphasize the connectivity-based formation of the Wikipedia network as we investigate the overall structure of its bibliometric network and its major communities.

This paper is organized as follows. Section two reviews briefly the literature on Wikipedia. The following section deals with the method used to conduct the analysis. In this section issues related to research design and documents' selection are presented. In Section four results of the analysis are presented. Section five discusses research findings and presents major research implications. This section also presents research limitations and explores avenues for future research.

2 Related work

Although Wikipedia is the largest user-contributed online encyclopedia in the world (Kim et al. 2019), academic research on Wikipedia "has been predominantly technical, focusing, for instance, on the size, reliability, and creation process of encyclopedic content" (Gobel and Munzert 2018, p. 157). An extensive scholarly literature review on Wikipedia has revealed three distinct streams of research. One stream of research has investigated the encyclopedia's usage as an indicator or predictor of offline phenomena. A second stream of research has questioned the credibility of Wikipedia, while a third line of research has investigated the different navigational strategies followed by Wikipedia users. In the following subsections we discuss in some detail the different research areas identified.

2.1 Wikipedia as a proxy of offline phenomena

Wikipedia was extensively used as an indicator of a plethora of offline phenomena. For example, Yang and Lai (2011) have explored how attitudes towards Wikipedia influence knowledge-sharing behavior. Yu and Gloor found a link between a firm's centrality position in Wikipedia and its innovation capability. Mestyan et al. (2013) used Wikipedia data to predict box office revenues. Public anxiety related to swine flu was also linked to Wikipedia visits (Tausczik et al. (2012). Zant et al. (2018a) analyzed major world terrorist networks using a reduced Google matrix of Wikipedia. Park et al. (2015) used Wikipedia to classify cultural values among nations based on the editors' collaboration system. Shi et al. (2019) analyzed millions of Wikipedia edits to study the effect of team composition on performance. In a similar vein, Gobel and Munzert (2018) explored how politicians tend to revise their biographies on Wikipedia as a means to appeal to their constituencies. Yasseri and Bight (2015) argued that the election outcomes can be predicted using Wikipedia data. The mass salience of the American Supreme Court decisions was also measured using Wikipedia (Wilkerson 2015). Behrendt et al. (2020) presented empirical evidence demonstrating that large-scale Wikipedia search data for individual-level stocks can be used to infer collective investor behavior. In a similar vein, Moat et al. (2013) investigated the link between Wikipedia's large-scale information search and stock market movements. Finally, Wikipedia has been used to investigate painters, universities and cancerous disease networks (Coquide et al. 2019; Rollin et al. 2019; Zant et al. 2018b).

2.2 Wikipedia credibility

Since Wikipedia allows anyone to edit and create articles, a critical issue related to authoritative reference source has emerged (Korfiatis et al. 2006). The debate on whether Wikipedia is a credible source of information has led to a plethora of research examining the encyclopedia's reliability (Koppen et al. 2015), quality (Arazy et al. 2011), accuracy (Brown 2011), content validity (Wallace and Fleet 2005), impact (Xu and Zhang 2013) and coverage (Keegan et al. 2013). However, research evaluating the Wikipedia's credibility compared to expert-produced encyclopedias suggests that the actual difference is trivial. For example, Tomaszewski and MacDonald (2016) found that the use of Wikipedia citations in refereed journals has been increasing. A *Nature* research team also found that Wikipedia's scientific information is reliable and accurate (Giles 2006). In a similar vein, several studies have shown that user-generated entries in Wikipedia to be as accurate and reputable as expert-created encyclopedias such as the Encyclopedia Britannica (Flanagin and Metzger

2011; Dewald 2014). In a study investigating citations to Wikipedia in chemistry journals, Brazzeal (2011) found that although "a small percentage of all articles contained a citation to Wikipedia, it is in fact being cited as a credible information source in articles in major chemistry journals." In a similar vein, Thompson and Hanley (2019) conducted a large-scale controlled-trial experiment on the scientific content of Wikipedia's articles and have given a robust support to Wikipedia's validity argument.

2.3 Wikipedia navigational strategies

Computational cognitive models have been extensively used to model Wikipedia navigation. For example, using the information foraging concept (Pirolli and Card 1999), Chi (2001) argued that "information scent", borrowed from the optimal animal foraging theory in biology, can be used to guide users' search for patches of information. Pirolli and Fu (2003) developed a cognitive model known in the literature as the SNIF-ACT model. This model aims at analyzing different navigation choices selected by users in their navigation between webpages. Rodi et al. (2017) modeled Wikipedia users' navigation using a stochastic memoryless process and argued that such statistical models are useful in simulating human navigation at the page level. Ratkiewicz et al. (2010) investigated Wikipedia users browsing activities over several months. The study was based on traffic flow gathered locally. Similarly, some studies have examined Wikipedia's navigational strategies via specially-designed games which track the navigation path followed by players under some spatial and temporal constraints (West et al. 2009, 2015). Studies comparing navigation models against real human search process found that such models do indeed mimic human navigation patterns (Lamprecht et al. 2015).

From this brief literature review, we find that although several studies have investigated Wikipedia using plethora of methodologies, no previous studies have implemented bibliometric network analysis to investigate Wikipedia's scholarly publications. In this research we fill this research gap by examining Wikipedia's scholarly publications over two decades.

3 Method

In this study we adapt the methodology suggested by Wamba and Mishra (2017) in which a four-stage bibliometric process was implemented. More specifically, the following steps were applied to guide the analyses reported in the article:

1. Defining the bibliometric database and search terms selection.
2. Performing preliminary data analysis.
3. Conducting network analysis.
4. Conducting thematic, conceptual structure and historiographic analyses.

To conduct the analysis, several packages within the R version 4.1 environment (R Development Core Team 2021) were used. The R ecosystem was selected because it provides a flexible and extensible environment to conduct sophisticated research since it is updated almost on a daily basis by researchers and industry practitioners. The *VOS viewer* software (van Eck and Waltman 2019) was also used to enhance graphs presentation. In the following subsections we discuss in some detail the methodological process mentioned above.

3.1 Database and search terms selection

Published documents on Wikipedia were retrieved from the Clarivate Analytics (previously known as Thomson Reuters) WoS database. With indexing coverage dating back to 1898, the WoS covers around 60 million records and produces several reliable impact metrics such as the journals' impact factors and the Journal Citation Report, among others (Merediz-Sola and Bariviera 2019). Although some studies have used the Scopus database to conduct bibliometric analyses (e.g., Jayantha and Oladinrin, Forthcoming), recent bibliometric studies found that around 70 percent of the researchers rely on the WoS to conduct scientometric research (Cuccurullo et al. 2016; Zupic and Cater 2015). To avoid false-positive results (Aleixandre et al. 2015), only article titles, abstracts and keywords were searched using the search terms "Wikipedia" OR "Wiki." We limited selection to documents written in English and we chose 2001 as the date of reference because Wikipedia was launched on January 15, 2001. Thus, the study covers twenty years of research in the field of Wikipedia's scholarly publications (January 15, 2001 till January 14, 2021). Figure 1 depicts the search strategy followed in this study.

3.2 Preliminary documents analysis

Documents' full bibliographic records were extracted, including information on authors, titles, keywords and other publication information. The retrieved documents were then transformed to a BibTeX format for further filtering and analysis. Table 1 shows the main attributes of the documents extracted from the WoS. From the table, we see that a total of 3790 documents written by 10,636 authors were extracted using the "Full Record and Cited References" option in the WoS database. The average Wikipedia document was cited 20.88 times, while the collaboration index was 3.23. Suggested by Ajiferuke et al. (1988), the collaboration index is a single measure based on the counting of fractional productivity defined by Price and Beaver (1966), whereas a value near zero indicates dominance of single-authored papers. It should be noted that selecting a specific type of document for analysis is debatable in the meta-analysis literature (Mostafa 2015, 2016). For example, while some authors focus only on journal articles (e.g., Corte et al. 2018), others concentrate on both journal articles and books (e.g., Aryadoust and Ang, Forthcoming), yet others

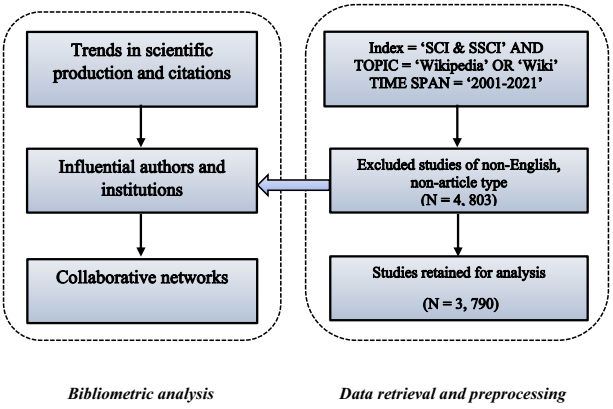

Fig. 1 Schematic flowchart of data acquisition methodology (Adapted from Chen et al. 2020a, b)

Table 1 Main document information

Description	Results
Timespan	2003:2021
Documents	3790
Average citations per documents	20.88
Average citations per year per doc	2.598
References	110,456
Document types	
Article	3790
Document contents	
Keywords plus (ID)	3677
Author's keywords (DE)	8951
Authors	
Authors	10,636
Authors of single-authored documents	602
Authors of multi-authored documents	10,034
Authors collaboration	
Documents per author	0.356
Authors per document	2.81
Co-Authors per documents	3.58
Collaboration index	3.23

eliminate only editorial material, meeting abstracts and corrections (e.g., Al-Khalifa 2014). In this study we select only peer-reviewed academic articles to ascertain the inclusion of only high-quality research articles. Chen et al. (2021, p. 206) argued that such articles "usually undergo a meticulous peer-review process and are generally of high quality."

3.3 Network analysis

A network is "a structure composed of a set of actors, some of whose members are connected by a set of one or more relationships" (Knoke and Yang 2010, p. 8). Thus, in a SNA, a relationship is represented by an edge, which is basically an arc connecting two interacting nodes or actors. As data are added, the network graph grows from a dyad, a one connection between two nodes, into a fully-fledged social network. Khan and Wood (2016, p. 388) argued that "when used to synthesize the existing literature from a network perspective, the SNA technique can reveal valuable invisible patterns that can certainly facilitate theory development and uncover areas for future research." From a social networks' perspective, there are various statistical tools that can be used to describe a network (Gruzd et al. 2011). In this study we use node size (number of nodes or actors/users in the network), density (proportion of actual links to the total possible links/relationships in the network) and diameter (the largest geodesic distance from one node to another in the network) to describe all the networks implemented. Network analysis techniques have been used extensively in bibliometric analysis. For example, in a co-citation network, papers A and B are said to be co-cited if both are cited by paper C. When two papers or authors are frequently co-cited, it is highly likely that their ideas are related to each other and hence belong to a similar research area and form a research cluster (Banckendorff 2009; Hjorland

2013). Thus, co-citation can be regarded as a measure of semantic similarity among authors or documents as the more co-citations articles receive, the more likely they are semantically related (Shiau et al. 2017). Yang et al. (2015) argued that co-citation analysis can be used to map the knowledge structure of specialized research areas and to identify key authors and influential journals in a particular field. Moreover, monitoring co-citation patterns over time can help in understanding the dynamic evolution of a research field and in detecting a paradigm shift in schools of thought (Fang et al. 2018; Yang et al. 2016). When mapping a co-citation network, an edge is formed between two nodes of authors or articles when a third author or article cites them together. Thus, conceptual clusters are created as certain groups of authors or articles are repeatedly cited in conjunction with each other.

Wang et al. (Forthcoming) argued that in a journal co-citation analysis, the more frequently two journals are co-cited, the stronger is their linkage. Chen and Liu (2020) found that a journal co-citation analysis can reveal the structure of a research field in which academic journals can be regarded as cardinal means of communication. This might be true since a journal co-citation analysis is basically "a visual illustration of knowledge capital flows that may reflect commonalities between journals such as research focus or methodology" (Wakefield 2008, p. 237).

Scientific collaboration networks can be analyzed at the micro (authors), meso (institutions) or macro (countries) level (Chen and Liu 2020). Co-authoring publications plays a critical role in disseminating research innovation, sharing knowledge as well as improving research quality (Zou et al. 2018). Evidence also shows that co-authored works are generally published in higher impact journals and receive more citations (Glänzel and Schubert 2005). Ding (2011) found that although collaboration among individual authors can play an important role in knowledge sharing and information dissemination, collaboration among research institutions can benefit the scientific field concerned in terms of creating research partnerships that feed into policy making. Countries collaboration networks can show the spread of research in a specific area among various nations. Finally, keywords co-occurrence networks can result in creating "a tapestry of themes" discussed in a research area (Banckendorff 2009, p. 6). By so doing, a set of "signal-words" that mark the literature and reflect its core contents can be detected (Space and Owens 1990).

3.4 Thematic and conceptual structure

Initially proposed by Law et al. (1988), a thematic map, also known as a strategic diagram, is used to examine the dynamics and evolution of the clusters derived from the keywords or co-word occurrences' analysis (Gonzales-Valiente 2019). The map is constructed based on the Callon's (1991) density and centrality metrics and it merges approaches related to co-word networks and portfolio analysis (Ávila-Robinson and Wakabayashi 2018). Because of its simplicity and significance, the diagram has been extensively used in the literature (Khasseh et al. 2017; Lee and Chen 2012; Zong et al. 2013). Conceptual structure of a research area can be mapped using conceptual structure maps, in which a domain is decomposed into distinct knowledge clusters in order to derive new insights in the data associated with each cluster (Wetzstein et al. Forthcoming). Temporal analysis can be used in bibliometric analysis to analyze the evolution of a research field within a longitudinal framework (Cobo et al 2011b). Emerging trends in a scientific field can be captured by articles receiving a "citation burst" (Kim and Chen 2015).

4 Results

4.1 Scientific production and influential authors

To find an answer to the first research question, we follow Corte et al. (2018) in tracing the evolution of Wikipedia literature. Figure 2 depicts the Wikipedia annual scientific production trends. The graph shows an exponential yearly growth rate. However, this growth is not evenly distributed. In the early years (2003–2007) the literature appears rather scarce with a few papers per year. This period may be termed the initial stage in the Wikipedia research. The second period (2008–2015) witnessed more attention towards Wikipedia topics as evidenced by the tremendous increase in publications. This stage can be termed the high growth stage. The next five years (2016–2021) can be termed the consolidation and stabilization stage as growth in this stage appears to reach the maturity or saturation stage. Similar results were reported in different scientific fields such as emergency evacuation and road safety studies (Zou et al. 2018). Table 2 shows the annual Wikipedia's annual scientific production. The overall growth of the scientific literature on Wikipedia is also reflected in the increasing number of authors (Table 3), citations per article (Table 4), sources used (Fig. 3) and countries involved (Fig. 4). Although Table 4 shows the most productive Wikipedia authors, analyzing most-productive authors' performance over time can be done using an author dominance map. Figure 5 shows a graph depicting the top twenty authors over time. Circles of varying sizes and colors indicate which authors are most influential in the field. From the graph, we see that the most influential authors were Ulrike Cress from 2008 till 2020 and Brendan Luyt from 2008 to 2021. From the graph, it seems also that Yuncheng Jiang's dominance has started only in 2012 with publications investigating sematic similarity between Wikipedia's concepts (Jiang 2017).

Evenness and concentration of authors' contribution is regarded as an important characteristic in bibliometric studies (Mercediz-Sola and Bariviera 2019). Several measures have been developed to measure concentration such as the Shannon entropy (Shannon and Weaver 1949) and the Lotka's law (1926), which argues that authors' productivity can be modeled using a form of Zipf's law. This empirical law implies that the number of authors producing a certain number of articles is a fixed ratio, 2, to single-article authors. In this

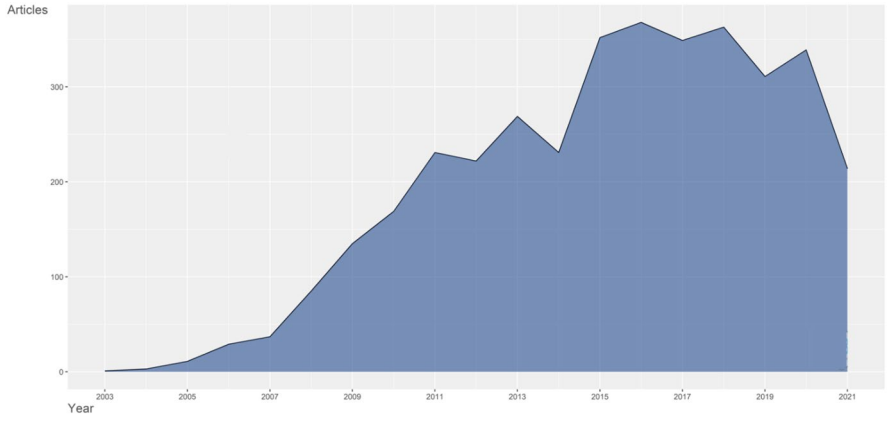

Fig. 2 Wikipedia annual scientific production

Table 2 Wikipedia annual scholarly publications

Year	Articles
2003	1
2004	3
2005	11
2006	29
2007	37
2008	85
2009	135
2010	169
2011	231
2012	222
2013	269
2014	231
2015	352
2016	368
2017	349
2018	363
2019	311
2020	339
2021	214

Table 3 Most productive authors

Authors	Articles	Articles fractionalized
Shepelyansky DL	24	8.08
Cress U	21	5.79
Fritz S	21	2.18
See L	20	2.02
KImmerle J	18	5.48
Chu SKW	17	6.49
Luyt B	15	12.25
Perger C	15	1.36
Yasseri T	15	4.82
Brigo F	14	2.70
Frahm KM	13	4.33
Mccallum I	13	0.91
Arazy O	12	3.84
Jiang YC	12	4.53
Oeberst A	12	3.15
Schepaschenko D	12	1.01
Moskaliuk J	11	3.48
Bragazzi NL	10	1.25
Lesiv M	10	0.95
Wagner C	10	4.03

Table 4 Top manuscripts per citations

Paper	Total citations	TC per year
Kaplan AM, 2010, Busuness Horizons	5225	435.4167
Kozomara A, 2014, Nucleic Acids Research	3281	410.125
Michie S, 2011, Implementation Science	1424	129.4545
Bizer C, 2009, Journal of web Semantics	879	67.6154
GRosdidier A, 2011, Nucleic Acids Research	815	74.0909
Lehmann J, 2015, Semant Web	758	108.2857
Hoffman MD, 2013, Journal of Mach Learning Research	733	81.4444
Nawrocki EP, 2015, Nucleic Acids Research	623	89
Wang XW, 2008, RNA	514	36.7143
Navigli R, 2012, Artificial Intelligence	483	48.3
SKINNER ME, 2009, Genome Research	477	36.6923
Rhead B, 2010, Nucleic Acids Research	449	37.4167
Hoffart J, 2013, Artificial Intelligence	436	48.4444
NETELER M, 2012, Environmental Modelling & Software	434	43.4
Arrowsmith CH, 2015, Nature Chemical Bioliology	432	61.7143
Delorme A, 2012, PLOS One	423	42.3
Karolchik D, 2008, Nucleic Acids Research	414	29.5714
RObinson TP, 2014, PLOS One	382	47.75
Bucher T, 2012, New Media & Society	370	37
van Gemert-Pijnen JEWC, 2011, Journal of Medical Internet Research	356	32.3636

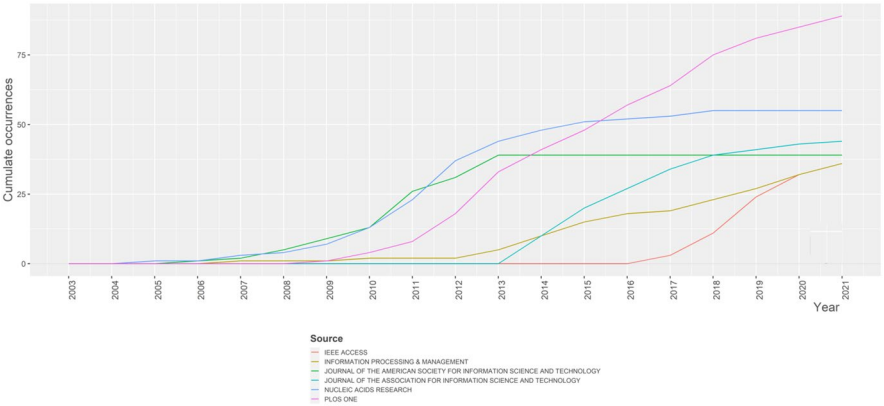

Fig. 3 Source dynamic growth

study we use the Lotka's law to measure Wikipedia authorship concentration because it has been widely used in similar bibliometric studies (Corbet et al. 2019). Figure 5 shows the observed vs the fitted Lotka's distributions. A formal Kolmogorov–Smirnov two-sample test suggests that there is no difference between the theoretical and the empirical distributions (*K-S p value* > 0.05). Thus, we conclude that Lotka's law appears to hold in the Wikipedia literature at the traditional 0.05 significance level. This result is in line with findings

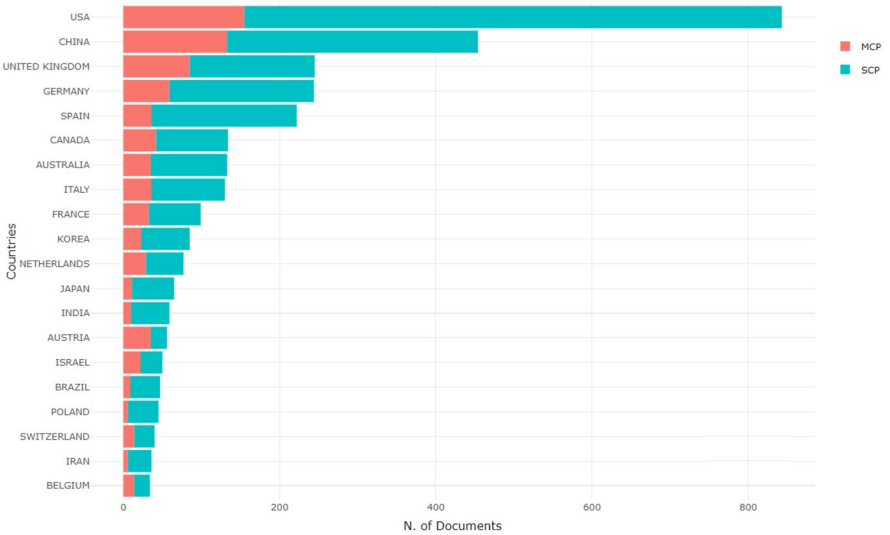

Fig. 4 Most productive countries in Wikipedia scholarly research. *Note* SCP = Single Country Publication; MCP = Multiple Country Publication

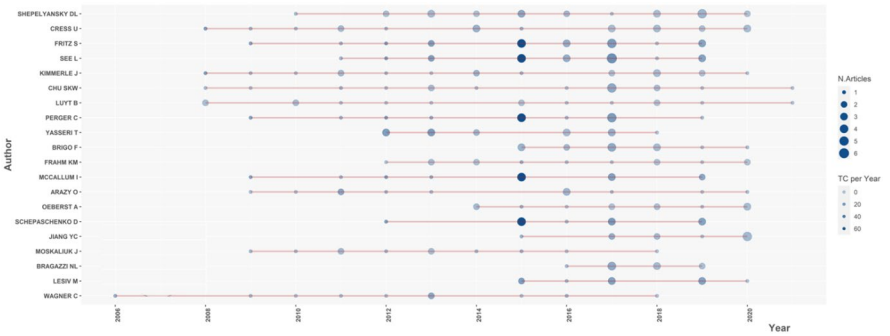

Fig. 5 Wikipedia's top-authors' production

reported in several research fields such as marketing (Vieira and Brito 2015), risk management (Chun-Hao and Jian-Min 2012), accounting (Corbet et al. 2019) and business ethics (Talukdar 2015). However, it is clear from Fig. 6 that single-paper authors represent a large tail compared to what may be expected. This implies that some authors might delve into the Wikipedia research domain, however, they do not focus on it as a main research area.

We also performed a geographical analysis to produce a "geospatial atlas" for Wikipedia research (Moradi, Forthcoming). This kind of analysis is usually performed "to discover where an event occurs and its impact on the neighboring areas" (Aria and Cuccurullo 2017, p. 962). Geospatial analysis has been extensively used in the literature. For example, Zhang et al (2015) used this approach to investigate the geographical distribution of world-wide tourism research. Geospatial results are usually visualized over a choropleth or proportional symbol map (Hu and Zhang 2017). Table 5 displays the geographic location of Wikipedia article citations. From the table, we see that the main regions publishing

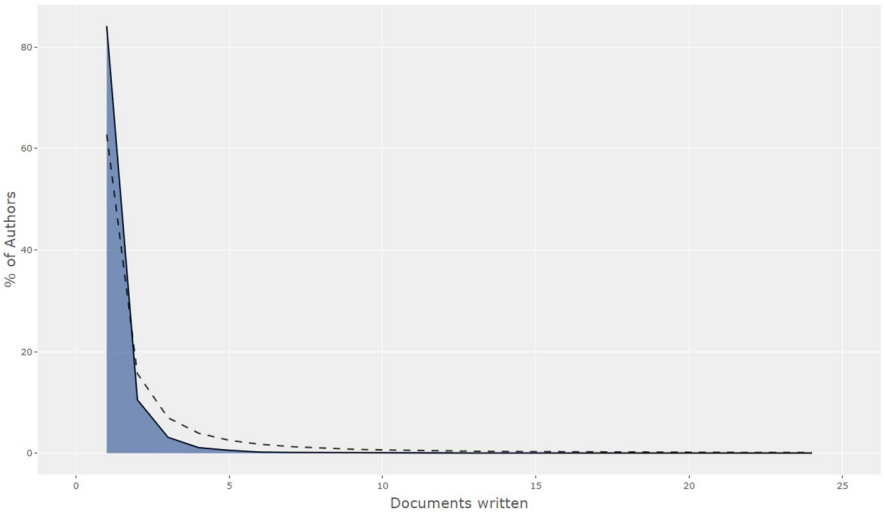

Fig. 6 Frequency distribution of Wikipedia scholarly research (Lotka's law)

Table 5 Total citations per country

Country	Total citations	Average article citations
USA	21,982	26.08
United Kingdom	9666	39.45
Germany	7200	29.51
France	6569	66.35
China	5836	12.85
Canada	2726	20.34
Italy	2618	20.14
Australia	2493	18.74
Spain	2260	10.18
Netherlands	2059	26.74
Austria	1804	32.21
Switzerland	1280	32.00
Israel	1116	22.32
Belgium	874	25.71
Korea	834	9.81
Sweden	714	23.03
Singapore	693	20.38
Norway	612	24.48
Japan	577	8.88
New Zealand	527	18.82

the most cited Wikipedia research are located in the USA, the UK, Germany, France, and China, with citations ranging between 527 in countries like New Zealand to 21,982 in the USA. It should be noted that we opted for the total number of citations instead of the

absolute numbers of published articles in our geospatial analysis since citations are better indicators of activity within a research area (Garfield et al. 1978; Osareh 1996), signifying a relatively more influential role for the more frequently cited articles in the scientific field (Nisonger 1994). In a similar vein, Pieters et al. (1999) argued that citations are valuable means in tracing the genealogical development in a research domain.

4.2 Network analysis

4.2.1 Co-citation networks

Figure 7 shows a Wikipedia's author co-citation network. A disambiguation algorithm was used to detect and merge duplicate profiles. A number of meaningful insights can be gained from examining the graph. For example, by focusing on the node sizes, we see that J Giles occupies a central position in the network. Thus, this author can be regarded as an influential in the field of Wikipedia research: an actor who is central in the network. Influential authors usually play an important role in defining interaction boundaries and information flow in a research network (Lin and Himelboim 2018). According to Bakshy et al. (2011), influential authors tend to have a disproportionate impact on the diffusion of information through a network because they are well-positioned to spark conversations and trigger feedback. Thus, influential authors tend to anchor each community and they have a large impact on other communities as they control and stimulate information diffusion through research activities. From the graph, we also see that some nodes are quite close to each other, indicating a strong "homophily effect". Homophily refers to the fact that "similarity breeds connection" (McPherson et al. 2001, p. 415). This effect occurs when actors in an environment similar to a virtual room deal with topics of common interest or discuss a common agenda (Findlay and Rensburg 2018). Jiang et al. (2019, p. 1940) found that homophily in bibliometric networks "is often determined by underlying disciplinary or thematic similarity." For example, we see from the graph that the node representing S Fritz is very close to the node representing M Kaktay, indicating possible homophily effect as both authors tend to work on topics related to Wikipedia's collaborative knowledge building. The graph also shows that the cluster at the far right (green-colored) is virtually disconnected sub cluster separated from the other clusters by white space, indicating the existence of structural holes (Burt 1992). Haythornthwaite (1996) argued that researchers can capitalize on structural holes by producing research papers linking the disjoint clusters. Such papers may act as a bridge as they can provide a link between otherwise disconnected clusters. Authors who link such disjoint clusters can receive citations from other authors in both clusters and they can play an important role in bridging structural holes in the network. Thus, such

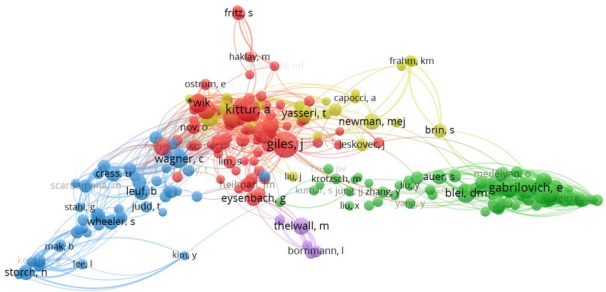

Fig. 7 Wikipedia author co-citation network (> = 20 articles)

authors might be regarded as information brokers who connect different groups in the network, which give them a "structural advantage" over other actors (Burt 1999).

Figure 8 depicts Wikipedia's journal co-citation network. From the graph, it is clear that there are seven distinct journal co-citation clusters. For example, the green cluster represent journals such as *Journal of the American Society for Information Science and Technology, MIS Quarterly, Knowledge Management* and *Social Networks Analysis*. This cluster of journals seems to focus on Wikipedia's research from a knowledge-based perspective. The blue cluster seems to focus on journals dealing with using Wikipedia in teaching and higher learning and includes journals such as *Contemporary Educational Psychology, Computers and Education, Language Learning Technology,* and *the Australasian Journal of Educational Technology*. With 154 items, the red cluster is the largest one and it includes journals such as *Lecture Notes on Computer Science, Neural Computing* and *Neurocomputing*. This cluster seems to focus on Wikipedia's computer science issues. The smallest cluster is the orange one and it seems to focus on Wikipedia's research within infometrics and comprises journals such as *Sientometrics* and *Informetrics*. From the graph, we also see that "core journals" occupy a central position within each cluster, while nodes that are dissimilar tend to drift further apart. The graph also reveals that scholarly interactions among the seven clusters are minimal, indicating an "orthodox core-heterodox periphery" phenomenon (Glötzl and Aigner 2018), in which each cluster is dominated by a few heavily-cited "orthodox journals", while "heterodox journals" are located in the periphery (Dobusch and Kapeller 2012).

4.2.2 Collaboration networks

Figure 9 displays the Wikipedia's author collaboration network. In this graph, the node size is proportional to each author's publications, while the link thickness is proportional to the total number of joint publications. The graph shows that the network is very sparsely connected, indicating the absence of close collaboration among the authors in the Wikipedia research field. From the graph, we see that the authors are grouped into three major research communities based on the color used. For example, the largest node in the red cluster is occupied by Steffen Fritz, Christoph Perger and Linda See. The green sub-network is dominated by Florian Kraxner and Ian McCallum, while the blue cluster is dominated by Dmitry Schepaschenko. The fragmented network indicates that the scale of cooperation in the field is nevertheless negligible. This implies that prolific authors in the field tend to be embedded in dense clusters with sparse ties outside their scientific neighborhoods. Overall, the Wikipedia network structure is quite similar to co-authorship networks in other research fields (Trier and Molka-Danielsen 2013; Vidgen et al. 2007).

Figure 10 displays the Wikipedia's research institutions' collaboration network. In this graph, the node size is proportional to each institution's publications, while the link

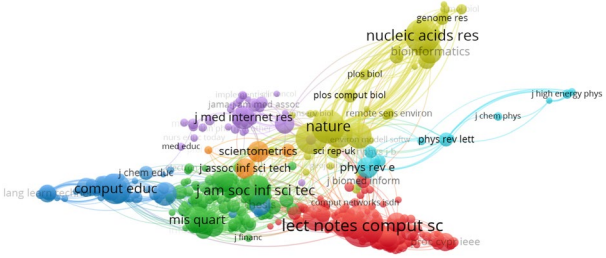

Fig. 8 Wikipedia journal co-citation network (>=40 articles)

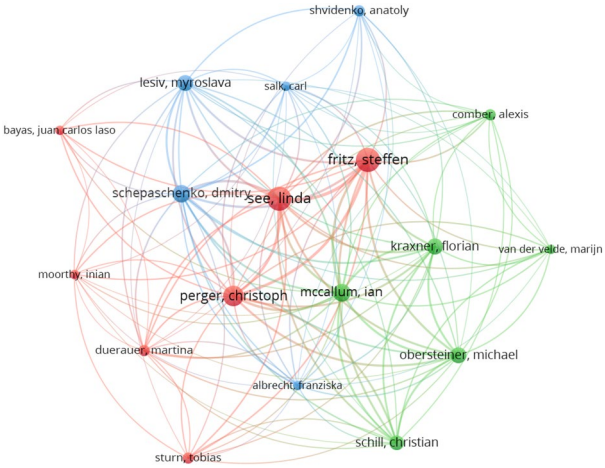

Fig. 9 Wikipedia author collaboration network (>=3 articles)

thickness is proportional to closer collaboration. From the graph, we see that the institutions are grouped into twelve major research communities based on the color used. From the graph, we also see that the green cluster subgroup includes universities such as the Chinese Academy of Science, the City University of Hong Kong, and Peking University. The red cluster includes universities such as the University of Toulouse in France, the University of Tubingen in Germany, and the Stanford University. The orange cluster includes universities such as the University of Limerick in Ireland, the University of Groningen in the Netherlands, and the University of North Carolina. The brown cluster includes universities such as the University of Helsinki in Finland, the University of Oslo in Norway, and the University of Exeter in the UK. This type of research institutions cooperation reflects a locally-centralized-globally-discrete type of cooperation (Zou et al. 2018). For example, although we see from the graph a strong collaboration among Chinese universities, we see only a weak cooperation among other countries. This type of cooperation may also reflect

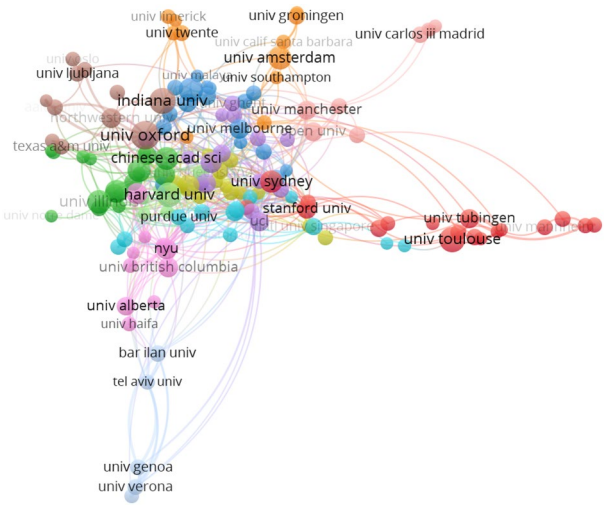

Fig. 10 Wikipedia institutions' collaboration network (>=10 articles)

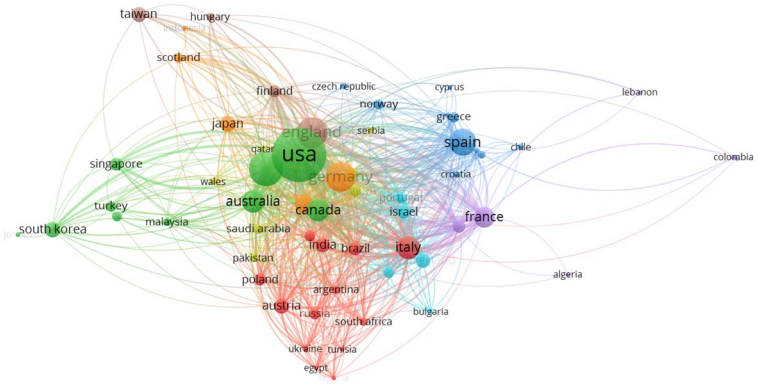

Fig. 11 Wikipedia countries collaboration network (>=5 articles)

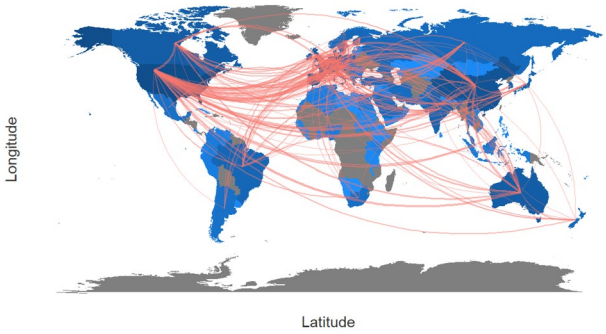

Fig. 12 Wikipedia world collaboration map

a North–South divide within the Wikipedia research domain as a very limited cooperation occurs between developed and developing countries' institutions. Notable exceptions include the blue cluster with cooperation among universities such as the King Saud University in Saudi Arabia, Nanyang Technological University in Singapore, the University of Tokyo, and Wuhan University. In terms of link strength, most of the close cooperation occur among institutions located in close geographic proximity. Overall, the type of collaboration among research institutions involved in Wikipedia research might reflect the fragmented nature of this research field, indicating the dearth of knowledge sharing among individual researchers.

Figure 11 displays the Wikipedia's research countries' collaboration network. In this graph, the node size represents a country's publications, while the link represents the level of cooperation among countries. From the graph, we see that the most productive country is the USA, with a strong cooperation with the Canada and South Korea. Lesser cooperation is also found between France and other countries such as Algeria and Lebanon. Although this network is denser than both individual and institutional collaboration networks, the density metric signifies that only around 9% of all possible cooperation relationships in the network is attained. This pattern is also evident from Fig. 12, which depicts the Wikipedia's research collaboration patterns on a world map. Color intensity in this map is

proportional to the number of publications, while the pink line represents the connection between cooperating nations.

5 Keywords and co-occurrence network analysis

Because keywords are generally characterized by a high conceptual level of abstraction, they can be used to identify the core content of scholarly publications, the thematic trends in a research area and the main research activity in a specific domain (Chen et al. 2008; Su and Lee 2010). Figure 13 displays a wordcloud of Wikipedia author-provided keywords. A wordcloud, also known as a tag cloud, can be regarded as a visual graph depicting the frequency of occurrence of a specific word in a document, whereby the higher the frequency of a word, the larger will its presence in the graph and vice versa. From the graph, we see that the most repetitive keywords were "Wikipedia", "information" and "knowledge". The sudden evolution of specific keywords' usage can indicate emerging research trends (Chen et al. 2012, 2014). Qian et al. (2019) argued that keywords that experience a recent abrupt surge and continue until the time of study might reveal future directions in a research area. Figure 14 displays keywords that have experienced a sharp recent rise in frequency, indicating "potential fronts" of Wikipedia's research since "the body of knowledge in a certain discipline can be seen as a sequence of topics that appear, grow in importance for a particular period and then disappear" (Colicchia et al. 2019). The keywords were extracted from the author-provided keywords of the studied corpus of the Wikipedia publications. From the graph, we see that areas such as information, knowledge, web and technology might be considered as important research fronts or "hotspots" in Wikipedia research.

We also used a Sankey diagram, also known as a three-field plot, to analyze the "flow" among keywords (left side), authors (middle) and cited documents (right side). A Sankey diagram shows the flow trend between two or more entities (Sankey 1898). The boxes and arrows widths in a Sankey diagram are proportional to the related quantity/the number of publications (Soundararajan et al. 2014). Figure 15 shows the Wikipedia's Sankey diagram in which the theme is named based on the keyword with the maximum degree. From the graph, we see that the edge widths emanating from keywords such as "Wikipedia", "Wiki" and "crowdsourcing" are the largest, indicating that several authors have used such keywords in their publications. We also see that authors such as Ulrike Cress and Ofer Arazy have used relatively more keywords compared to other authors, indicating that the authors

Fig. 13 Wikipedia keyword wordcloud

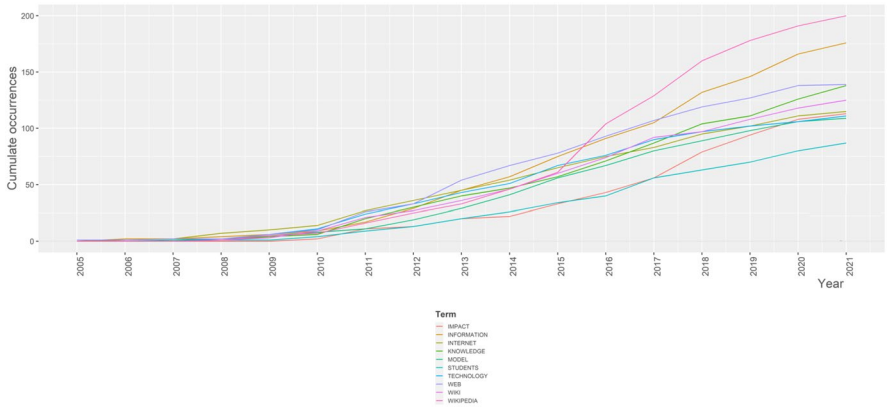

Fig. 14 Wikipedia keywords growth dynamics

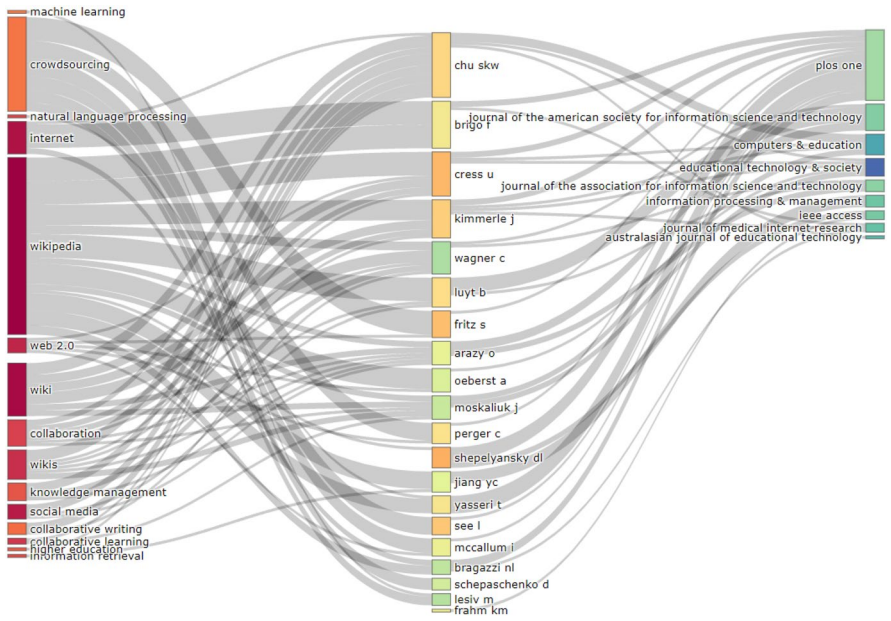

Fig. 15 Wikipedia Sankey diagram (keyword-author-source flow)

have covered several themes within the field of Wikipedia's research. This is also indicated by the diversity of the outlets used by the two authors to publish their works.

Furthermore, we used author-provided keywords to build the keyword co-occurrence network as "authors of a paper should be the ones that have the best feel as to what areas are spoken to by the paper" (Corbet et al. 2019). The network was constructed to further probe how often keyword co-occur in the same document. Figure 16 depicts the Wikipedia keyword co-occurrence network, where the size of the node provides a relative indication of the frequency of each keyword and the width of the edges or links shows the frequency

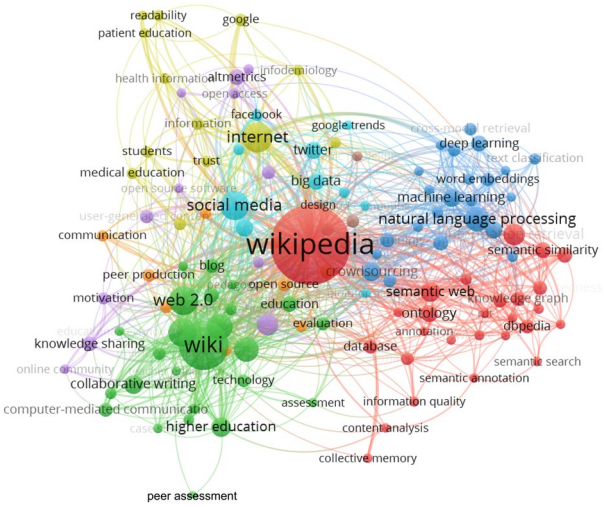

Fig. 16 Wikipedia keyword co-occurrence network

each pair of keywords were mentioned together. From the graph, we see that distinct clusters are formed dealing mainly with Wikipedia's usage in medical education, semantic relatedness, knowledge discovery, collaborative learning and Web 2.0/online communities. These topics can be taken as indicators for priorities and trends in Wikipedia research since large nodes near the center of the network tend to represent important research paradigms (Neff and Corley 2009) or research "hotspots" (van Eck and Waltman 2014).

5.1 Conceptual structure map

Following Demiroz and Haase (2019), we conducted a multiple correspondence analysis (MCA) on the keywords obtained. Figure 17 presents the conceptual structure of the keywords associated with Wikipedia's research over a period of twenty years. From the graph, we see that the best dimension reduction achieved for the first two dimensions of the MCA account for roughly 42% of the total variability. In this graph, the closer the dots, the similar the profile they represent, whereas each cluster of dots represents discriminating profiles (Wong et al. 2021). The resulting conceptual map organizes the intellectual structure oh Wikipedia research in three distinct clusters. A closer look at the graph reveals the breadth, diversity and intellectual thrust of the field. For example, the biggest cluster in red includes keywords that emphasize Wikipedia's information content. Examples include "knowledge", "information", "quality" and "web." Representative articles include Arazy et al. (2011) who constructed a framework aiming at investigating information quality in Wikipedia and Cress and Kimmerle (2008) who investigated Wikipedia's collaborative knowledge building. The second cluster in blue is comprised of keywords, including "communication", "education" and "technology." This cluster deals with Wikipedia use in education. Authors within this cluster discuss issues such as cognitive learning and the evolution of Wikipedia's medical content (Oeberst et. al, 2014; Shafee et al. 2017). The third cluster in green is the smallest cluster and seems to deal with esoteric topics such as ontology, tool and database. Examples include Forte et al. (2009) who investigated the relationship between Wikipedia's decentralization and satisfaction.

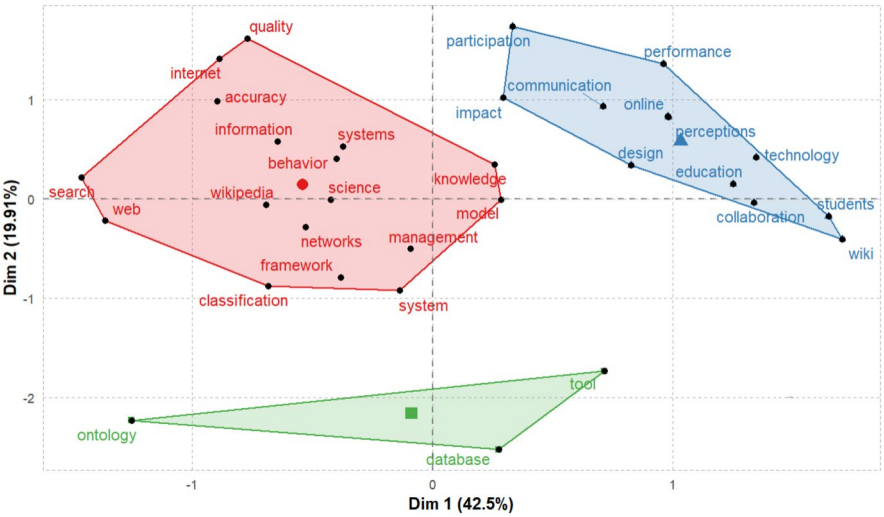

Fig. 17 Wikipedia conceptual structure map

5.2 Thematic maps

Figure 18 shows the Wikipedia thematic/strategic map. The dotted line in the middle of the diagram represents the average values of both axes and divide the diagram into four quadrants representing different types of themes, while the size of the bubbles is proportionate

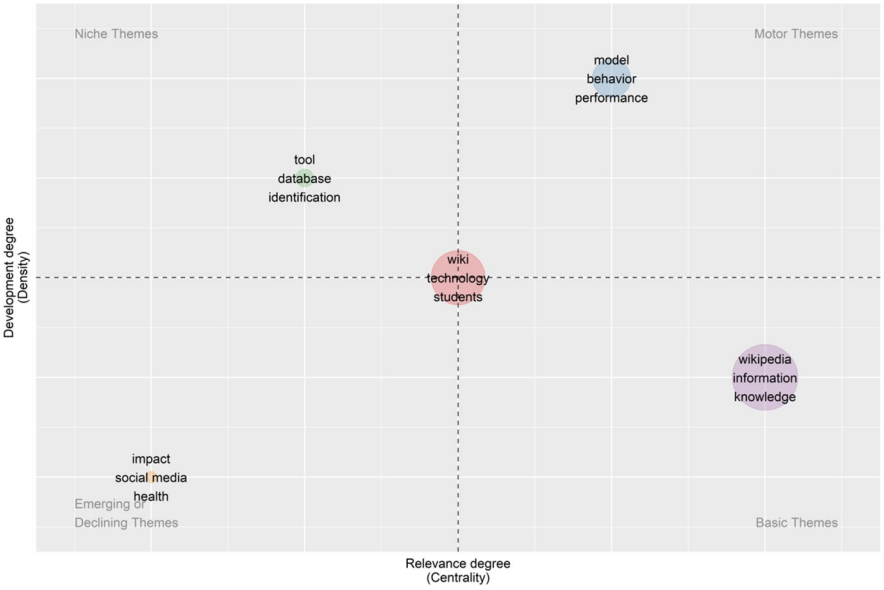

Fig. 18 Wikipedia thematic map

to the number of publications in which the keywords appear. According to Cobo et al. (2011a), the first quadrant is characterized by high density and high centrality, signifying well-developed themes both internally and externally. Such themes are referred to as "motor themes." The second quadrant is characterized by high density and low centrality, signifying that, although the detected themes have well-developed internal ties, their external ties are marginally important. The themes falling at this quadrant are known as "highly-developed-and-isolated themes." The third quadrant is characterized by both low density and centrality, signifying themes with weak ties both internally and externally. Such themes are referred to in the literature as "emerging-or-declining themes." Finally, the fourth quadrant comprises themes with low density and high centrality, signifying themes characterized by weakly developed internal ties, however, they have important external ties. Themes belonging to this quadrant are known as "basic-and-transversal themes."

In this sense, the first quadrant on the top right contains themes with both high centrality and density. Thus, themes in this quadrant are both well developed and can influence the research domain. It can therefore be observed that themes such as "model", "behavior" and "performance" have remained seminal and transversal during the two decades of Wikipedia's scholarly publications. Upper left quadrant keywords such as "databases", "tool" and "identification" represent a niche theme, which is well developed in terms of internal links, however, such keywords have weak external ties and is of marginal importance. Lower left quadrant keywords such as "social media" and "health" represent possible emerging themes/hotspots in the field of Wikipedia scholarly research. Finally, the lower-right quadrant represents basic themes, and it includes terms such as "information" and "knowledge." This indicates that such components are essential for the development of scholarly research on Wikipedia.

5.3 Historiographic analysis

We conducted a historiographic analysis to map the most relevant chronological citations resulting from a bibliographic collection in the form of a network. Historiographic analysis identifies the most influential works on a given topic and traces its historical development on a year-by-year basis. By so doing, a historiography can be used to draw the intellectual linkages in historical order. Historiographic mapping has been used in several bibliometric studies, including mapping sex selection technologies literature, problematic weeds (Diekmann et al. 2013) and service networks (da Silva et al. 2017). Figure 19 displays the Wikipedia research historiography based on the direct citation linkages. The graph reveals a citation network with three distinct clusters. The main cluster, which lasted from 2006 to 2015, is comprised primarily of six influential authors like Roy Rosensweig (2006), Sabina Niederer and John Jones. This group of widely-cited authors have investigated issues such as wisdom of the crowd, patterns of revision and networks. It should be noted that the historiography starts with Rosensweig's (2006) seminal work on the use of Wikipedia in historical studies. In this article, Rosensweig (2006, p. 117–118; 137) termed Wikipedia's success "the great democratic triumph" arguing that Wikipedia represents "the most important application of the principles of the free and open-source software movement." The smaller cluster started in 2009 with the influential works of Denise Anthony and Andrea Forte, with works on Wikipedia collaborative work, governance and reputation and reliability. This result suggests that this publication had been highly influential in the Wikipedia's research domain, especially during the beginning of its burst. Another notable finding from the historiography is that this article achieved its impact over a relatively long

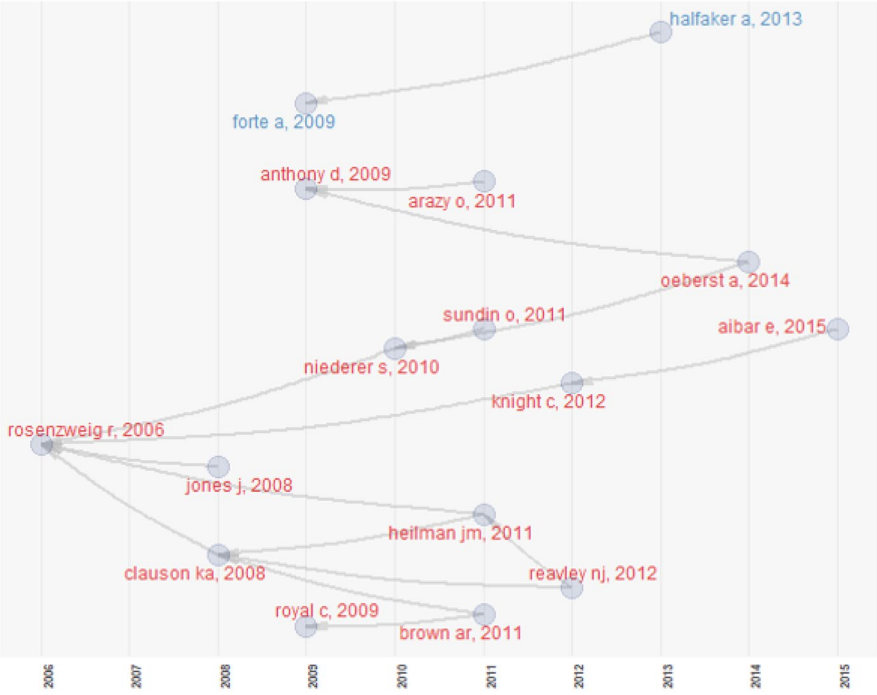

Fig. 19 Wikipedia historical direct citation network

period of time as although it was published in 2006, it took two years for the publication to become a focal point of Wikipedia research.

6 Discussion, limitations and future research

This study was driven by an interest in visualizing and mapping the structure and scope of Wikipedia research by simultaneously tracing the development of its scholarly publications and revealing its intellectual core. Using bibliometric networks, a novel and rigorous methodological approach, we analyzed 3790 WoS Wikipedia documents spanning two decades and written by 10,636 authors from 100 countries. Indeed, Gheisari and Esmaeili (2019) argued that given the growth of a research field, and given yet that current knowledge about the field is limited in focus, mapping the state of this knowledge is timely. Bibliometric methodology objectively presents results that represent a joint view expressed by real data as opposed to subjective techniques since "an arbitrary selection of evidence is often not fully representative of the state of existing knowledge, and the selection of some studies over others ultimately leads to what is known in statistical analysis as a sample selection bias" (Linnenluecke et al. Forthcoming, p. 2). By borrowing concepts from SNA, we present a "big picture" of Wikipedia's research and we remedy the inherent problems in traditional methodological approaches that often overlook the dynamics and connectedness between authors, articles and the journals in which the studies are published. Chen and Leydesdorff (2014) noted that this type of interconnectedness is crucial in understanding

the intellectual turning points and the progress of knowledge in a research domain over time.

By analyzing all Wikipedia documents indexed in the WoS since 2001, the paper constitutes the first comprehensive bibliometric analysis of the field. The large amount of data made it possible to identify the most influential scholars, core journals and major trends and themes of Wikipedia's multidisciplinary research domain. For example, a closer inspection of the co-citation networks (author and journal/source) reveals that the typology of the Wikipedia research networks reflects a hub-and-spoke structure typical of small world networks (Milgram 1967; Watts and Strogatz 1998). This structure indicates the existence of few key actors who dominate the network and maybe regarded as influentials/academic leaders in the Wikipedia research network. Such actors may also play an important role in information diffusion/brokerage through the network, forming bridges between other key clusters in the network (Park et al. 2015). It has been argued that small-world networks are characterized by small groups of nodes that are well-connected and highly clustered, which "let messages move through the network effectively" (Shirky 2008, p. 215). Smith and Graham (2019, p. 1318) argued that small-world networks are also robust since "randomly removing nodes from the network will not significantly impact the effectiveness and dynamics of the network." The Wikipedia research networks reflect also what is sometimes called the Matthew Effect in sociology- a mechanism akin to the preferential attachment in which a small number of hubs have larger number of connections compared to the rest of actors. Newman (2004) argued that new links in a large network are preferably attached to well-connected nodes. Consequently, connectivity in highly connected nodes augments faster than less-connected nodes. This phenomenon is known in the literature as the power law mechanism, which is a variation of the rich-get-richer model in which just a few actors form the core of the cluster in which they are located (Himelboim et al. 2013). Preferential attachment has been found in myriad social networks, including academic work citations, emails, sexual relationships, diseases and World Wide Web hyperlinks (Barabasi et al. 2000; Himelboim and Han 2014).

Another interesting finding relates to the fact that the collaboration between institutions and nations is governed by cultural proximity, geographical distance, and specializations patters, which confirm the resulted reported in Prieto-Gutierrez and Segado-Boj (Forthcoming) who studied collaboration patterns among Asian authors. The finding that the core research forces were universities followed by research institutions was in line with previous research (Chen and Liu 2020). Moreover, the rare alliance in Wikipedia research between developing and developed countries was also reported in similar studies (Vanni et al. 2014).

Notwithstanding the contribution of this study, there are some limitations. First, unlike the study of Holub and Johnson (2018), which combined multiple data sources, we relied only on a single database. As a result, citations and associated connections between documents will be more conservative compared to other indices such as Google Scholar. However, it should be noted that almost all the documents indexed, for example, in Scopus and WoS are dual-indexed in both databases (Gavel and Iselid 2008). Neuhaus et al. (2006) also argued that Google Scholar is less stringent in including citations from blogs, syllabi, unpublished presentations and related material. Future research may add more databases to test whether our results hold. Second, similar to other studies (e.g., Qian et al. 2019), we also limited our search to documents published only in English language. This might limit the scope of coverage, so future research should add articles published in other languages, especially the languages of the most productive nations such as China to test the generalizability of our results across other languages. Third, although we focused on Wikipedia as a knowledge domain, future research may focus on Wikipedia research published

in specific journals such as *the Journal of the American Society for Information Science and Technology, the Journal of the Association for Information Science and Technology,* and *the Australasian Journal of Educational Technology*, among others. In fact, current bibliometric research has conducted bibliometric analyses to various academic journals such as the *Journal of Knowledge Management* (Gaviria-Marin et al. 2018), the *Journal of Infection and Public Health* (Krauskopf 2018), *the International Journal of Computer Integrated Manufacturing* (Laengle et al. 2018), the *Journal of Travel and Tourism Marketing* (Mulet-Forteza et al. 2018), *Sustainability* (Tang et al. 2018a), the *International Journal of Fuzzy Systems* (Tang et al. 2018b), *the Journal of Business-to-Business Marketing* (Valenzuela-Fernandez et al. 2019), *Public Management* Review (Kumar et al. 2020)*, Managerial Finance* (Baker et al. 2020), *Review of Managerial Science (*Mas-Tur et al. 2020), *the Journal of School Health* (Zhang et al. 2017) and *the Journal of Advanced Nursing* (Zeleznik et al. 2017). Fourth, although we used a rigorous bibliometric analysis to examine twenty years of Wikipedia publications, future research might use a topic modeling approach to uncover hidden patterns in Wikipedia's textual data (Chen et al. 2020a, b, 2021). Finally, although we used a co-citation network method widely used in the literature, Skupin (2004, 2009) has argued that this method can sometimes mask important relationships and has shown that alternative techniques such as self-organizing maps and the use of continuous space may be viable alternatives. Thus, future research should compare co-citation methods used in this study with other methods to check the validity of our results.

Funding The authors have not disclosed any funding.

Declarations

Conflict of Interest The authors have not disclosed any competing interests.

References

Ajiferuke, I., Burrel, Q., Tague, J.: Collaborative coefficient: a single measure of the degree of collaboration in research. Scientometrics **14**, 421–433 (1988)

Aleixandre, J., Aleixandre-Tudo, J., Bolanos-Pizarro, M., Aleixandre-Benavent, R.: Mapping the scientific research in organic farming: a bibliometric review. Scientometrics **105**, 295–309 (2015)

Al-Khalifa, H.: Scientometric assessment of Saudi publication productivity in computer science in the period of 1978–2012. Int. J. Web Inf. Syst. **10**, 194–208 (2014)

Arazy, O., Nov, O., Patterson, R., Yeo, L.: Information quality in Wikipedia: the effects of group composition and task conflict. J. Manag. Inf. Syst. **27**, 71–98 (2011)

Aria, M., Cuccurullo, C.: Bibliometrix: An R-Tool for comprehensive science mapping analysis. J. Inform. **11**, 959–975 (2017)

Aryadoust, V., Ang, B: Exploring the frontiers of eye tracking research in language studies: a novel co-citation scientometric review. Comput. Assist. Lang. Learn. (Forthcoming)

Ávila-Robinson, A., Wakabayashi, N.: Changes in the structures and directions of destination management and marketing research: a bibliometric mapping study, 2005–2016. J. Destin. Mark. Manag. **10**, 101–111 (2018)

Azad, H., Deepak, A.: A new approach for query expansion using Wikipedia and WordNet. Inf. Sci. **492**, 147–163 (2019)

Baker, H., Kumar, S., Pandey, N.: Abibliometric analysis of managerial finance: a retrospective. Manag. Finance **46**, 1495–1517 (2020)

Bakshy, E., Hofman, J., Mason, W., Watts, D.: Everyone's an influencer. In: King, I., Nejdl, W., Li, H. (eds.) Proceedings of the 4th ACM International Conference on Web Search and Data Mining – WSDM'11, p. 65. ACM Press, New York (2011)

Banckendorff, P.: Themes and trends in Australian and New Zealand tourism research: a social network analysis of citations in two leading journals (1994–2007). J. Hosp. Tour. Manag. **16**, 1–15 (2009)
Barabasi, A., Jeong, H., Neda, Z., Ravasz, E., Schubert, A., Vicsek, T.: Evolution of the social network of scientific collaboration. Phys. A **311**, 590–614 (2000)
Behrendt, S., Peter, F., Zimmermann, D.: An encyclopedia for stock markets? Wikipedia searches and stock returns. Int. Rev. Financ. Anal. **72**, 101563 (2020)
Block, J., Fisch, C., Rehan, F.: Religion and entrepreneurship: a map of the field and a bibliometric analysis. Manag. Rev. q. **70**, 591–627 (2020)
Bouzembrak, Y., Kluche, M., Gavai, A., Marvin, H.: Internet of Things in food safety: literature review and a bibliometric analysis. Trends Food Sci. Technol. **94**, 54–64 (2019)
-Brazzeal, B.: Citations to Wikipedia in chemistry journals: a preliminary study. Issues in Science and Technology Librarianship, 67 (http://www.istl.org/11-fall/refereed2.html) (2011)
Brown, A.: Wikipedia as a data source for political scientists: Accuracy and completeness of coverage. Polit. Sci. Polit. **44**, 339–343 (2011)
Burt, R.: Structural Holes. Harvard University Press, Cambridge, MA (1992)
Burt, R.: The social capital of opinion leaders. Ann. Am. Acad. Pol. Soc. Sci. **566**, 37–54 (1999)
Cai, K., Spangler, S., Chen, Y., Zhang, L.: Leveraging sentiment analysis for topic detection. Web Intell. Agent Syst. **8**, 291–302 (2010)
Callon, M., Courtial, J., Laville, F.: Co-word analysis as a tool for describing the network of interactions between basic and technological research: the case of polymer chemistry. Scientometrics **22**, 155–205 (1991)
Chen, C.: Visualization of knowledge structures. Handb. Softw. Eng. Knowl. Eng. **2**, 700–744 (2002)
Chen, C., Leydesdorff, L.: Patterns of connections and movements in dual-map overlays: a new method of publication portfolio analysis. J. Am. Soc. Inf. Sci. **65**, 334–351 (2014)
Chen, X., Liu, Y.: Visualization analysis of high-speed railway research based on CiteSpace. Transp. Policy **85**, 1–17 (2020)
Chen, C., Paul, R.: Visualizing a knowledge domain's intellectual structure. Computer **34**, 65–71 (2001)
Chen, C., Song, I., Yuan, X., Zhang, J.: The thematic and citation landscape of data and knowledge engineering (1985–2007). Data Knowl. Eng. **67**, 234–250 (2008)
Chen, C., Hu, Z., Liu, S., Tseng, H.: Emerging trends in regenerative medicine: a scientometric analysis in CiteSpace. Expert Opin. Biol. Ther. **12**, 593–608 (2012)
Chen, C., Dublin, R., Kim, M.: Orphan drugs and rare diseases: a scientometric review (2000–2014). Expert Opin. Orphan Drugs **2**, 709–724 (2014)
Chen, X., Zou, D., Xie, H.: Fifty years of British journal of educational technology: a topic modeling based bibliometric perspective. Br. J. Edu. Technol. **51**, 692–708 (2020a)
Chen, X., Zou, D., Cheng, G., Xie, H.: Detecting latent topics and trends in educational technologies over four decades using structural topic modeling: a retrospective of all volumes of computers & education. Comput. Educ. **151**, 103855 (2020b)
Chen, X., Zou, D., Xie, H., Cheng, G.: Twenty years of personalized language learning: topic modeling and knowledge mapping. Educ. Technol. Soc. **24**, 205–222 (2021)
Chi, E.: Using information scent to model user information needs and actions on the web. In Proceedings of the SIGGHI Conference on Human Factor in Computing Systems. ACM, pp. 490–497 (2001)
Chun-Hao, C., Jian-Min, Y.: A bibliometric study of financial risk literature: a historic approach. Appl. Econ. **44**, 2827–2839 (2012)
Cobo, M., Lopez-Herrera, A., 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. **5**, 146–166 (2011a)
Cobo, M., Lopez-Herrera, A., Herrera-Viedma, E., Herrera, F.: Science mapping software tools: review, analysis, and cooperative study among tools. J. Am. Soc. Inf. Sci. Technol. **62**, 1382–1402 (2011b)
Colicchia, C., Creazza, A., Noe, C., Strozzi, F.: Information sharing in supply chains: a review of risks and opportunities using the systematic literature network analysis (SLNA). Supply Chain Manag. **24**, 5–21 (2019)
Coquide, C., Emann, L., Lages, J., Shepelyansky, D.: World influence and interactions of universities from Wikipedia networks. Eur. Phys. J. B **92**(1), 3 (2019)
Corbet, S., Dowling, M., Gao, X., Huang, S., Lucey, B., Vigne, S.: An analysis of the intellectual structure of research on financial economics of precious metals. Resour. Policy **63**, 101416 (2019)
Corte, V., Gaudio, G., Sepe, F.: Ethical food and the Kosher certification: a literature review. Br. Food J. **120**, 2270–2288 (2018)
Cress, U., Kimmerle, J.: A systemic and cognitive view on collaborative knowledge building with wikis. Int. J. Comput. Support. Collab. Learn. **3**, 105–122 (2008)

Cuccurullo, C., Aria, M., Sarto, F.: Foundations and trends in performance management: a twenty-five years bibliometric analysis in business and public administration domains. Scientometrics **108**, 595–611 (2016)
da Silva, S., Antonio, N., Carvalho, J.: Analysis of the service dominant logic network, authors, and articles. Serv. Ind. J. **37**, 125–152 (2017)
de la Hoz-Correa, A., Munoz-Leiva, F., Bakuca, M.: Past themes and future trends in medical tourism research: a co-word analysis. Tour. Manag. **65**, 200–211 (2018)
Demiroz, F., Haase, T.: The concept of resilience: a bibliographic analysis of the emergency and disaster management literature. Local Gov. Stud. **45**, 308–327 (2019)
Dewald, N.: Future voices in public services. Public Serv. Q. **10**, 245–251 (2014)
Diekmann, F., Ford, R., Harisson, S., Regnier, E., Venkatesh, R.: Bibliometric analysis of the literature on Giant Ragweed (*Ambrosia trifida* L.). J. Agric. Food Inf. **14**, 290–320 (2013)
DiMaggio, P., Hargittai, E., Neuman, W., Robinson, J.: Social implications of the internet. Ann. Rev. Sociol. **27**, 307–336 (2001)
Ding, Y.: Scientific collaboration and endorsement: network analysis of co-authorship and citation networks. J. Informetr. **5**, 187–203 (2011)
Dobusch, L., Kapeller, J.: A guide to paradigmatic self-marginalization: lessons for post-Keynesian economists. Rev. Polit. Econ. **24**, 469–487 (2012)
Ezell, J.: Empathy plasticity: decolonizing and reorganizing wikipedia and other online spaces to address racial equity. Ethn. Racial Stud. (Forthcoming)
Fang, Y., Yin, J., Wu, B.: Climate change and tourism: a scientometric analysis using CiteSpace. J. Sustain. Tour. **26**, 108–126 (2018)
Findlay, K., van Rensburg, O.: Using interaction networks to map communities on Twitter. Int. J. Mark. Res. **60**, 169–189 (2018)
Flanagin, A., Metzger, M.: From encyclopedia britannica to wikipedia. Inf. Commun. Soc. **14**, 355–374 (2011)
Forte, A., Larco, V., Bruckman, A.: Decentralization in Wikipedia governance. J MIS **26**(1), 49–72 (2009)
Francisco, G., Enrique, C., Bartolome, M., Mercedes, U.: Identifying the 'knowledge base' or 'intellectual structure' of research on international business, 2000–2015: a citation/co-citation analysis of JIBS. Int. Bus. Rev. **28**, 713–726 (2019)
Gaede, J., Rowlands, I.: Visualizing social acceptance research: a bibliometric review of the social acceptance literature for energy technology and fuels. Energy Res. Soc. Sci. **40**, 142–158 (2018)
Garfield, E., Malin, M., & Small, H.: Citation data as science indicators. Eds. Y. Elkana, J. Lederberg, R. Merton, R. Thackray, A., & H. Zuckerman, pp. 179–208. Wiley, New York (1978)
Gavel, Y., Iselid, L.: Web of science and scopus: a journal title overlap study. Online Inf. Rev. **32**, 8–21 (2008)
Gaviria-Marin, M., Merigo, J., Popa, S.: Twenty years of the journal of knowledge management: a bibliometric analysis. J. Knowl. Manag. **22**, 1655–1687 (2018)
Gheisari, M., Esmaeili, B.: Applications and requirements of unmanned aerial systems (UASs) for construction safety. Saf. Sci. **118**, 230–240 (2019)
Giles, J.: Internet encyclopedias go head to head. Nature **438**, 900–901 (2006)
Glänzel, W., Schubert, A.: Analyzing scientific networks through co-authorship. In: Moed, H., Glanzel, W., Schmoch, U. (eds.) Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems. Springer, Dordrecht (2005)
Glötzl, F., Aigner, E.: Orthodox core-heterodox periphery? Contrasting citation networks of economics departments in Vienna. Rev. Polit. Econ. **30**, 210–240 (2018)
Gobel, S., Munzert, S.: Political advertising on the Wikipedia marketplace of information. Soc. Sci. Comput. Rev. **36**, 157–175 (2018)
Gonzales-Valiente, C.: Redes de citación de revistas iberoamericanas de bibliotecología y ciencia de la información en Scopus. Bibliotecas Anal. Investig. **15**, 83–98 (2019)
Gruzd, A., Wellman, B., Takhteyev, Y.: Imagining Twitter as an imagined community. Am. Behav. Sci. **55**, 1294–1318 (2011)
Gupta, R., Pandey, R., Sebastian, V.: International entrepreneurial orientation (IEO): a bibliometric overview of scholarly research. J. Bus. Res. **125**, 74–88 (2021)
Haythornthwaite, C.: Social network analysis: An approach and technique for the study of information exchange. Library Inform. Sci. Res. **18**, 323–342 (1996)
Himelboim, I., Han, J.: Cancer talk on Twitter: community structure and information sources in breast and prostate cancer social networks. J. Health Commun. **19**, 210–225 (2014)

Himelboim, I., Smith, M., Shneiderman, B.: Tweeting apart: applying network analysis to detect selective exposure clusters in Twitter. Commun. Methods Meas. **7**, 195–223 (2013)
Hjorland, B.: Citation analysis: a social and dynamic approach to knowledge organization. Inf. Process. Manage **49**, 1313–1325 (2013)
Holub, M., Johnson, J.: Bitcoin research across disciplines. Inf. Sci. **34**, 114–126 (2018)
Hu, J., Zhang, Y.: Structure and patterns of cross-national big data research collaborations. J. Doc. **73**, 1119–1136 (2017)
Jayantha, W., Oladinrin, O.: Knowledge mapping of office workspace: a scientometric review of studies. Facilities (Forthcoming)
Jiang, Y., Bai, W., Zhang, X., Hu, J.: Wikipedia-based information content and semantic similarity computation. Inf. Process. Manag. **53**, 248–265 (2017)
Jiang, Y., Ritchie, B., Benckendorff, P.: Bibliometric visualization: an application to tourism crisis and disaster research. Curr. Issue Tour. **22**, 1925–1957 (2019)
Jullien, N.: What we know about Wikipedia: a review of the literature analyzing the project(s). 86 (2012)
Keegan, B., Gergle, D., Contractor, N.: Hot off the Wiki: structures and dynamics of Wikipedia's coverage of breaking news events. Am. Behav. Sci. **57**, 595–622 (2013)
Khan, G., Wood, J.: Knowledge networks of the information technology management domain: A social network analysis approach. Commun. Assoc. Inf. Syst. **39**, 367–397 (2016)
Khasseh, A., Soheili, F., Moghaddam, N., Chelak, A.: Intellectual structure of knowledge in iMetrivs: a co-word analysis. Inf. Process. Manag. **53**, 705–720 (2017)
Kim, M., Chen, C.: A scientometric review of emerging trends and new developments in recommendation systems. Scientometrics **104**, 239–263 (2015)
Kim, J., Kim, S., Lee, C.: Anticipating technological convergence: link prediction using Wikipedia hyperlinks. Technovation **79**, 25–34 (2019)
Knoke, D., Yang, S.: Social Network Analysis. SAGE, Los Angeles, CA (2010)
Koppen, L., Phillips, J., Papageorgiou, R.: Analysis of reference sources used in drug-related Wikipedia articles. J. Med. Libr. Assoc. **103**, 140–144 (2015)
Korfiatis, N., Poulos, M., Bokos, G.: Evaluating authoritative sources using social networks: an insight from Wikipedia. Online Inf. Rev. **30**, 252–262 (2006)
Kosterich, A., Weber, M.: Transformation of a modern newsroom workforce: a case study of NYC journalist network histories from 2011 to 2015. J. Pract. **13**, 431–457 (2019)
Krauskopf, E.: A bibliometric analysis of the Journal of Infection and Public Health: 2008–2016. J. Infect. Public Health **11**, 224–229 (2018)
Kumar, S., Pandey, N., Haldar, A.: Twenty years of public management review (PMR): a bibliometric overview. Public Manag. Rev. **22**, 1876–1896 (2020)
Laengle, S., Modak, N., Merigo, J., de la Sotta, C.: Thirty years of the international journal of computer integrated manufacturing: a bibliometric analysis. Int. J. Comput. Integr. Manuf. **31**, 1247–1268 (2018)
Lamprecht, D., Strohmaier, M., Helic, D., Nyulas, C., Tudorache, T., Noy, N., Mark, A.: Using ontologies to model human navigation behavior in information networks: a study based on Wikipedia. Semant. Web **6**, 403–422 (2015)
Law, J., Bauin, S., Courtial, J., Wittaker, J.: Policy and the mapping of scientific change: a co-word analysis of research into environmental acidification. Scientometrics **14**, 251–264 (1988)
Lee, M., Chen, T.: Revealing research themes and trends in knowledge management: from 1995 to 2010. Knowl.-Based Syst. **28**, 47–58 (2012)
Leong, C., Lee, Y., Mak, W.: Mining sentiments in SMS texts for teaching evaluation. Expert Syst. Appl. **39**, 2584–2589 (2012)
Levitt, P.: Religion on the move: mapping global cultural production and consumption. In: Bender, C., et al. (eds.) Religion on the Edge: De-Centering and Re-Centering the Sociology of Religion, pp. 159–176. Oxford University Press, Oxford (2012)
Lin, J., Himelboim, I.: Political brand communities as social network clusters: winning and training candidates in the GOP 2016 primary elections. J. Polit. Mark. **16**, 1–29 (2018)
Linnenluecke M.K., Singh, A.: Conducting systematic literature reviews and bibliometric analyses. Aust. J. Manag. (Forthcoming)
Lotka, A.: The frequency distribution of scientific productivity (1926)
Mas-Tur, A., Brandtner, M., Ewert, R., Kursten, W.: Advances in management research: a bibliometric overview of the review of managerial science. RMS **14**, 933–958 (2020)
McPherson, M., Smith-Lovin, L., Cook, J.: Birds of a feather: homophily in social networks. Ann. Rev. Sociol. **27**, 415–444 (2001)

Merediz-Sola, I., Bariviera, A.: A bibliometric analysis of bitcoin scientific production. Res. Int. Bus. Finance **50**, 294–305 (2019)
Mestyan, M., Yasseri, T., Kertesz, J.: Early prediction of movie box office success based on Wikipedia activity big data. PLoS ONE **8**(8), e71226 (2013)
Milgram, S.: The small-world problem. Psychol. Today **2**, 60–67 (1967)
Moat, H., Curme, C., Avakian, A., Kenett, D., Stanley, H., Preis, T.: Quantifying Wikipedia usage patterns before stock market moves. Sci. Rep. **3**, 1801 (2013)
Moradi, S.: The scientometrics of literature on smart cities. Libr. Hi Tech (Forthcoming)
Mostafa, M.: Do products' warning labels affect consumer safe behavior? A meta-analysis of the empirical evidence. J. Bus. Econ. Stud. **22**, 24–39 (2015)
Mostafa, M.: Do consumers recall products' warning labels? a meta-analysis. Int. J. Manag. Mark. Res. **9**, 81–96 (2016)
Mulet-Forteza, C., Martorell-Cunill, O., Merigo, J., Genovart-Balaguer, J., Mauleon-Mendez, E.: Twenty-five years of the journal of travel & tourism marketing: a bibliometric ranking. J. Travel Tour. Mark. **35**, 1201–1221 (2018)
Navarrete, T., Borowiecki, K.: Changes in cultural consumption: ethnographic collections in Wikipedia. Cult. Trends **25**, 233–248 (2016)
Neff, M., Corley, E.: 35 years and 160,000 articles: a bibliometric exploration of the evolution of ecology. Scientometrics **80**, 657–682 (2009)
Neuhaus, C., Neuhaus, E., Asher, A., Wrede, C.: The depth and breadth of Google Scholar: an empirical study. Libr. Acad. **6**, 127–141 (2006)
Newman, M., Girvan, M.: Finding and evaluating community structure in networks. Phys. Rev. E **69**, 1–15 (2004)
Nisonger, T.: A methodological issue concerning the use of social sciences citation index journal citation reports impact factor data for journal ranking. Libr. Acquis. Pract. Theory **18**, 447–458 (1994)
Okoli, C.: A brief review of studies of Wikipedia in peer-reviewed journals. In: 2009 Third International Conference on Digital Society, pp. 155–160. IEEE (2009)
Osareh, F.: Bibliometrics, citation analysis and co-citation analysis: a review of literature. Libri **46**, 149–158 (1996)
Park, S., Lim, Y., Park, H.: Comparing Twitter and YouTube networks in information diffusion: The case of the "Occupy Wall Street" movement. Technol. Forecast. Soc. Chang. **95**, 208–217 (2015)
Pieters, P., Baumgartner, H., Vermunt, J., Bijmolt, T.: Importance and similarity in the evolving citation network of the international journal of research in marketing. Int. J. Res. Mark. **16**, 113–127 (1999)
Pirolli, P., Fu, W.: SNIF-ACT: a model of information foraging on the World Wide Web. In: The 9th International Conference on User Modeling. Springer, New York (2003)
Pirolli, P., Card, S.: Information Foraging. Princeton University Press, Princeton (1999)
Portugal-Ferreira, M., Frias-Pinto, C., Ribeiro-Serra, F.: The transaction costs theory in international business research: a bibliometric study over three decades. Scientometrics **98**, 1899–1922 (2014)
Price, D., Beaver, D.: Collaboration in an invisible college. Am. Psychol. **21**, 1011–1018 (1966)
Prieto-Gutierrez, J., Segado-Boj, F.: Annals of library and information studies: a bibliometric analysis of the journal and a comparison with the top library and information studies in Asia and worldwide. Ser. Libr. (Forthcoming)
Qi, T., Wang, T., Ma, Y., Zhang, W., Zhu, Y.: A scientometric analysis of e-participation research. Int. J. Crowd Sci. **2**, 136–148 (2018)
Qian, J., Law, R., Wei, J.: Knowledge mapping in travel website studies: a scientometric review. Scand. J. Hosp. Tour. **19**, 192–209 (2019)
Qin, H., Prastyo, Y., Bass, M., Sanders, C., Prentice, E., Nguyen, Q.: Seeing the forest for the tree: a bibliometric analysis of environmental and resource sociology. Soc. Nat. Resour. **3**, 1131–1148 (2020)
R Development Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (www.R-project.org) (2021).
Ratkiewicz, J., Menczer, F., Fortunato, S., Flammini, A.: Traffic in social media II: modeling bursty popularity. In: IEEE Second International Conference on Social Computing, pp. 393–400 (2010)
Rivera, M., Pizam, A.: Advances in hospitality research: "from Rodney Dangerfield to Aretha Franklin." Int. J. Contemp. Hosp. Manag. **27**, 362–378 (2015)
Rodi, G., Loreto, V., Tria, F.: Search strategies of Wikipedia readers. PLoS ONE **12**(2), e0170746 (2017)
Rollin, G., Lages, J., Shepelyansky, D.: Wikipedia network analysis of cancer interactions and world influence. PLoS ONE **14**(9), e0222508 (2019)

Rosensweig, R.: Can history be open source? Wikipedia and the future of the past. J. Am. History **93**, 117–146 (2006)
Ruiz-Alba, J., Guzman-Parra, V., Oblitas, J., Mediano, J.: Entrepreneurial intentions: a bibliometric analysis. J. Small Bus. Enterp. Dev. **2**, 121–213 (2021)
Sankey, H.: Introductory note on the thermal efficiency of stream-engines. In: Proceedings of the Institution of Civil Engineers, pp. 278–283 (1898)
Scott, J.: Social Network Analysis. Sage, London (2013)
Shafee, T., Masukume, G., Kipersztok, L., Das, D., Häggström, M., Heilman, J.: Evolution of Wikipedia's medical content: past, present and future. J. Epidemiol. Community Health **71**, 1122–1129 (2017)
Shannon, C., Weaver, W.: The Mathematical Theory of Communication. University of Illinois Press, Champaign, IL (1949)
Shi, F., Teplitsky, M., Duede, E., Evans, J.: The wisdom of polarized crowds. Nat. Hum. Behav. **3**, 329–336 (2019)
Shiau, W., Dwivedi, Y., Yang, H.: Co-citation and cluster analysis of extant literature on social networks. Int. J. Inf. Manag. **37**, 390–399 (2017)
Shirky, C.: The Power of Organizing Without Organizations. Penguin Press, New York (2008)
Skupin, A.: The world of geography: visualizing a knowledge domain with cartographic means. Proc. Natl Acad Sci. U. S. a. **101**, 5274–5278 (2004)
Skupin, A.: Discrete and continuous conceptualizations of science: Implications for knowledge domain visualization. J. Informetr. **3**, 233–245 (2009)
Smith, N., Graham, T.: Mapping the anti-vaccination movement on Facebook. Inf. Commun. Soc. **22**, 1310–1327 (2019)
Soundararajan, K., Ho, H., Su, B.: Sankey diagram framework for energy and exergy flows. Appl. Energy **136**, 1035–1042 (2014)
Space, D., Owens, K.: Lexical co-occurrence and association strength. J. Psycholinguist. Res. **19**, 317–330 (1990)
Su, H., Lee, P.: Mapping knowledge structure by keyword co-occurrence: a first look at journal papers in technology foresight. Scientometrics **85**, 65–79 (2010)
Surowiecki, J., Silverman, M.: The wisdom of the crowds. Am. J. Phys. **75**, 190 (2007)
Talukdar, D.: Research productivity patterns in the organizational behavior and human resource management literature. Int. J. Hum. Resour. Manag. **26**, 467–484 (2015)
Tang, M., Liao, H., Wan, Z., Herrera-Viedma, E., Rosen, M.: Ten years of sustainability (2009 to 2018): a bibliometric overview. Sustainability **10**, 1655 (2018a)
Tang, M., Liao, H., Rosen, M.: A bibliometric overview and visualization of the international journal of fuzzy systems between 2007 and 2017. Int. J. Fuzzy Syst. **20**, 1403–1422 (2018b)
Tausczik, Y., Faasse, K., Pennebaker, J., Petrie, K.: Public anxiety and information seeking following the H1N1 outbreak: blogs, newspaper articles, and Wikipedia visits. Health Commun. **27**, 179–185 (2012)
The Economist (January 9th 2921). The other giant, pp. 49–50
Thompson, N., Hanley, D.: Science is shaped by Wikipedia: evidence from a randomized control trial. MIT Sloan Research Paper No. 5238-17. SSRN-ID 3039505 (2019)
Tomaszewski, R., MacDonald, K.: A study of citations to Wikipedia in scholarly publications. Sci. Technol. Libr. **35**, 246–261 (2016)
Trier, M., Molka-Danielsen, J.: Sympathy or strategy: social capital drivers for collaborative contributions to the IS community. Eur. J. Inf. Syst. **22**, 317–335 (2013)
Valenzuela-Fernandez, L., Merigo, J., Lichtenthal, J., Nicolas, C.: A bibliometric analysis of the first 25 years of the journal of business-to-business marketing. J. Bus. Bus. Mark. **26**, 75–94 (2019)
van Eck, N., Waltman, L.: Visualizing bibliometric networks. In: Ding, Y., Rousseau, R., Wolfram, D. (eds.) Measuring Scholarly Impact: Methods and Practice. Springer, New York (2014)
van Eck, N., Waltman, L.: VOSviewer, Version 1.6.13 (2019)
Vanni, T., Mesa-Frias, M., Sanchez-Garcia, R., Roesler, R., et al.: International scientific cooperation in HIV and HPV: a network analysis. PLoS ONE **9**(3), e93376 (2014)
Vidgen, R., Henneberg, S., Naude, P.: What sort of community is the European conference on information systems? A social network analysis 1993–2005. Eur. J. Inf. Syst. **22**, 317–335 (2007)
Vieira, F., Brito, C.: Science mapping in industrial marketing. J. Bus. Ind. Mark. **16**, 5–19 (2015)
Vila-Lopez, N., Kuster-Boluda, I.: A bibliometric analysis on packaging research: towards sustainable and healthy packages. Br. Food J. **123**, 684–701 (2021)
Vos, T., Heinderyckx, F.: Gatekeeping in Transition. Routledge, New York (2015)
Wakefield, R.: Networks of accounting research: a citation-based structural and network analysis. Br. Account. Rev. **40**, 228–244 (2008)
Wallace, J.: Modeling contemporary gatekeeping: the rise of individuals, algorithms and platforms in digital new dissemination. Digit. Journal. **6**, 274–293 (2018)
Wallace, D., Fleet, C.: The democratization of information? Ref. User Serv. Q. **45**, 10–102 (2005)

Wamba, S., Mishra, D.: Big data integration with business processes: a literature review. Bus. Process. Manag. J. **23**, 477–492 (2017)
Wang, C., Lim, M., Zhao, L., Tseng, M., Chien, C., Lev, B.: The evolution of omega-the international journal of management science over the past 40 years: a bibliometric overview. Omega
Wäsche, H., Dickson, G., Woll, A., Brandes, U.: Social network analysis in sport research: an emerging paradigm. Eur. J. Sport Soc. **14**, 138–165 (2017)
Watts, D., Strogatz, S.: Collective dynamics of 'small-world' networks. Nature **393**, 440–442 (1998)
West, R., Pineau, J., Precup, D.: Wikipedia: an online game for inferring semantic distances between concepts. In: Proceedings of the 21st International Conference on Artificial Intelligence, Morgan Kaufmann Publishers, pp. 1598–1603 (2009)
-West, R., Paranjape, A., Leskovec, J.: Mining missing hyperlinks from human navigation traces: a case study of Wikipedia. In: Proceedings of the 24th International Conference on World Wide Web, pp. 1242–1252. ACM (2015)
-Wetzstein, A., Feisel, E., Hartmann, E., Benton, W.: Uncovering the supplier selection knowledge structure: a systematic citation network analysis from 1991 to 2017. J. Purch. Supply Manag. (Forthcoming)
Wilkerson, B.: Using Wikipedia page views to measure the mass salience of US Supreme Court decisions. In: The Annual Meeting of the Midwest Political Science Association, Chicago, IL, April 16 (2015)
Wong, W., Mittas, N., Arvanitou, E., Li, Y.: A bibliometric assessment of software engineering themes. Schools and institutions (2013–2020). J. Syst. Softw. **180**, 111029 (2021)
Wu, G., Cunningham, P.: Integration of multiple network views in Wikipedia. Knowl. Inf. Syst. **45**, 473–490 (2015)
Xu, S., Zhang, X.: Impact of Wikipedia on market information environment: evidence on management disclosure and investor reaction. MIS Q. **37**, 1043–1068 (2013)
Yang, H., Lai, C.: Understanding knowledge-sharing behavior in Wikipedia. Behavi. Inf. Technol. **30**, 131–142 (2011)
Yang, G., Li, G., Li, C., Zhao, Y., Zhang, T., Liu, T., Chen, D., Huang, M.: Using the comprehensive patent citation network (CPC) to evaluate patent value. Scientometrics **105**, 1319–1346 (2015)
Yang, S., Han, R., Wolfram, D., Zhao, Y.: Visualizing the intellectual structure of information science (2006–2015): introducing author-keyword coupling analysis. J. Informetr **10**, 132–150 (2016)
Yasseri, T., Bright, J.: Predicting elections from online information flows: towards theoretically informed models. arXiv:1505–01818 (2015)
Zant, S., Frahm, K., Jaffres-Runser, K., Shepelyansky, D.: Interactions and influence of world painters from reduced Google matrix of Wikipedia networks. IEEE Access **6**, 47735–47750 (2018b)
Zant, S., Frahm, K., Jaffres-Runser, K., Shepelyansky, D.: Analysis of world terror networks from the reduced Google matrix of Wikipedia. Eur. Phys. J. B 91 (2018a)
Zeleznik, D., Blazun, H., Kokol, P.: A bibliometric analysis of the journal of advanced nursing: 1976–2015. J. Adv. Nurs. **73**, 2407–2419 (2017)
Zhang, S., Lyu, P., Yan, Y.: Global geographical and scientometric analysis of tourism-themed research. Scientometrics **105**, 385–401 (2015)
Zhang, C., Zheng, X., Su, C., Huang, H., Yan, F., et al.: A bibliometric study of the journal of school health: 1965–2014. Chin. Nurs. Res. **4**, 75–83 (2017)
Zhu, J., Hua, W.: Visualizing the knowledge domain of sustainable development research between 1987 and 2015: a bibliometric analysis. Scientometrics **110**, 893–914 (2017)
Zong, Q., Shen, H., Yuan, Q., Hu, X., Hou, Z., Deng, S.: Doctoral dissertations of library and information science in China: a co-word analysis. Scientometrics **94**, 781–799 (2013)
Zou, X., Yue, W., Vu, H.: Visualization and analysis of mapping knowledge domain of road safety. Accid. Anal. Prev. **118**, 131–145 (2018)
Zupic, I., Cater, T.: Bibliometric methods in management and organization. Organ. Res. Methods **18**, 429–472 (2015)