

Shadow Libraries: A Bibliometric Analysis of Black Open Access Phenomenon (2011: 2023)

Safinaz Mahmoud Elroukh

Safinaz.mahmoud@nbu.edu.sa

Assistant professor, College of Humanities and Social Sciences, University of Northern Borders, Saudi Arabia
Lecturer, College of Arts, University of Menoufia, Egypt

Abstract

This study analyzes the global literature on the black open-access phenomenon from 2011 to 2023. A bibliometric analysis was conducted using the Scopus database. The search strategy employed advanced queries with multiple synonymous terms to ensure exhaustive retrieval of relevant documents. The VOSviewer software was employed to visualize the co-occurrence networks. The findings reported 90 papers published during the study period. An evolving scholarly landscape was revealed, with heightened attention from 2016 onwards, peaking in 2017, 2021, and 2023. Articles constitute 83.3% of the total published documents. Singh and Srichandan are prolific authors, with 11.2% of the total publications. The United States contributes 18.9% of the papers, followed by India and Spain. Information Development and Scientometrics are pivotal journals in scholarly discussions about this scope, contributing 4.4% of publications. Co-occurrence network visualization revealed "Sci-Hub" and "open access" as the most used keywords in the global literature. The findings underscore the need for additional research to discover innovative business models to safeguard intellectual property rights while meeting researchers' evolving needs. The importance of this paper comes from being the first bibliometric study analyzing international literature related to this phenomenon, which provides a basis for future research efforts and policymaking.

Keywords:

Shadow Libraries; Pirate Libraries; Black open Access; Sci-Hub; Library Genesis; Bibliometric Analysis.

1. Introduction

One of the biggest challenges facing many information institutions is the financial burden of academic database fees. Even Harvard University Library acknowledged for possessing one of the world's most substantial budgets, is confronted with difficulties in meeting the annual subscription costs, estimated at approximately 3.5 million US dollars [1]. This enduring escalation of subscription fees has instigated a crisis. In response to this predicament, the open access (OA) movement has emerged as an alternative to traditional publishing, presenting more

economically viable alternatives to the conventional publishing model [2]. However, the Open Access movement encountered many hurdles, including the pervasive presence of predatory journals and high publication fees [3], [4], which led to the emergence of resistance movements to call for the right to free access to scientific papers to foster scientific development in societies. Some pirate platforms have appeared among researchers containing copyrighted publications accessible for free without permission from the copyright holders [5], known as shadow libraries, pirate libraries, or black open-access platforms.

Perspectives on such platforms vary depending on viewpoints, with opinions ranging from considering it a criminal and unethical practice to viewing it as a justified act of civil disobedience. The ethical and legal implications surrounding using pirate websites underscore the complex landscape of access to academic knowledge and the ongoing discourse regarding open access and intellectual property rights [6]. In this view, the high costs and paywalls associated with traditional publishing are seen as barriers to the democratization of information. On the other hand, this practice is considered criminal and unethical, as it involves the unauthorized duplication of copyrighted content and undermines the traditional publishing model that relies on subscription fees and pay-per-view access. The debate surrounding the use of pirate websites reflects broader discussions about access to information, the sustainability of publishing models, and the balance between intellectual property rights and the pursuit of knowledge.

Sci-hub and Library Genesis are considered the most popular and widely used piracy platforms among researchers worldwide [7]. In September 2011, the software engineer and neurotechnologist from

Kazakhstan, Alexandra Elbakyan, introduced the Sci-Hub platform. She characterized the excessive fees of research articles behind paywalls as academic capitalism and imperialism. Her main goal was "removing all barriers in the way of science" in opposition to these forces. Since its launch, Sci-Hub has played a significant role in enabling open access to scientific knowledge, enabling thousands of people to download over 51,000,000 scientific papers daily [8]. Its popularity has surged, propelled by widespread media coverage and numerous studies highlighting its substantial impact on researchers who have warmly embraced this initiative. The emergence of Sci-Hub has brought forth innovative avenues for accessing scientific information, creating a ripple effect that directly influences academic and research libraries [9]. Intriguingly, although Sci-Hub use in countries like the UK, USA, Malaysia, and China remains limited, it is utilized as much for convenience as necessity [10]. This platform has become a significant force in challenging traditional access models and advocating for the global unrestricted dissemination of scholarly information.

The Library Genesis platform, commonly known as LibGen, hosts and distributes an expansive repository containing 25 million documents, equivalent to an impressive 42 terabytes of data, including research articles, textbooks, and books. This extensive collection is curated through a dual mechanism, with most documents originating from isolated yet substantial uploads. Additionally, LibGen benefits from the collective efforts of users worldwide engaging in daily crowdsourcing activities. The LibGen website features a user-friendly search engine that enables users to explore and download materials directly from its servers. This platform offers files for individual download and the option for bulk download facilitated through peer-to-peer torrent files [11].

Undoubtedly, the emergence of shadow libraries marks a transformative phenomenon in the academic publishing market. However, this surge in popularity has consequences for publishers, who are witnessing significant financial losses and creative harm due to the proliferation of free access to scholarly content [12]. This unprecedented approach challenges the conventional model of paid access to scholarly content, raising profound questions about the accessibility and dissemination of scientific knowledge in the digital age. The impact of these platforms extends beyond

their provision of free access, sparking debates within the academic community and prompting a reevaluation of existing publishing practices and business models [13]. In light of these challenges, there is a call for reconsidering the established rules governing the publishing market and exploring innovative approaches to monetizing scientific output. One proposed solution is a shift towards a subscription-based access model for publishing databases, akin to the structure employed in online music streaming services [14].

2. Purpose of the study

Shadow libraries, exemplified -but not limited- by platforms like Sci-Hub, LibGen, Z-library, Bibliotik, Memory of the World, Monoskop, and more, have transformative implications for scholarly communication. This unprecedented approach challenges the conventional model of paid access to scholarly content, raising profound questions about the accessibility and dissemination of scientific knowledge in the digital age. The study aims to analyze the scholarly landscape surrounding the black open-access phenomenon represented in shadow libraries from 2011 to 2023. It analyzes temporal patterns, document types, authors' contributions, affiliations, geographical representation, linguistic dimensions, and co-occurrence networks. Furthermore, the investigation highlights the key journals contributing to the scholarly discourse and the most cited papers to provide a holistic view of the phenomenon, shedding light on its global scope.

The study aims to address the following questions:

- What types of scholarly contributions dominated the global literature of shadow libraries between 2011 and 2023?
- Is there an apparent temporal pattern in the research output?
- Who are the prominent authors, and what affiliations do they represent?
- Which countries contribute enormously to the debate?
- What languages are used in scholarly contributions?
- Which keywords are most common?
- Which papers have received the most citations?
- Which journals play a crucial role in discussing the phenomenon?

3. Literature Review

Searching the global literature revealed that no previous bibliometric studies analyzed the literature on shadow libraries. However, many studies dealt with the phenomenon regarding researchers' motivations towards using it and its ethical aspects. Most studies focused on the pirate platforms that researchers use, like Sci-hub and LibGen. The most significant previous research is shown below, arranged chronologically from oldest to most recent:

Zimerman (2011) [15] explored the landscape of e-book piracy and its implications for academic libraries. The paper highlighted the views of the publishing industry on piracy and discussed the potential involvement of students in e-book piracy. Despite threats of fines and legal consequences, piracy persists, with individuals, including both hackers and non-hackers, utilizing easily accessible hacking tools. While the technical challenges may deter some, others are drawn to the excitement of overcoming obstacles. Zimerman suggests that the book publishing industry, similar to the film and music industries, must confront the issue of piracy, leading to legal battles and potential changes in copyright laws. The paper emphasizes the need for libraries to assess and address the legal threats of piracy, advocating for measures such as assigning user accounts to prevent copyright breaches. The study underscores the importance of vigilance among libraries regarding e-book readers and piracy, highlighting the need for further investigation.

Bohannon (2016) [16] provided a narrative of Alexandra Elbakyan, the mastermind behind Sci-Hub, the world's largest pirate repository for academic papers. Through Elbakyan's story, Bohannon shed light on the broader issues surrounding scholarly communication and access to academic literature. Bohannon's article elucidated Elbakyan's motivations, challenges, and ethical dilemmas associated with her actions, portraying her as an idealistic advocate for open access and a controversial figure embroiled in legal battles. Bohannon offered valuable insights into the complexities of scholarly publishing, copyright, and the ongoing debate over open access.

Machin-Mastromatteo, Uribe-Tirado, and Romero-Ortiz (2016) [17] investigated the phenomenon of scientific paper piracy in Latin America by analyzing usage data from Sci-Hub. The researchers delved into

the implications of Sci-Hub's presence for information professionals, universities, and libraries, shedding light on the broader impact within the regional academic landscape.

Greshake Bastian (2017) [18] examined the phenomenon of potentially illegal access to scholarly publications facilitated by the web service Sci-Hub amidst the growth of Open Access. Utilizing a recently released corpus of Sci-Hub data, he compared it to approximately 28 million downloads conducted through the service to investigate user motivations and usage patterns. The comparative analysis revealed a preference among users for recently published articles, with a significant proportion (35%) being published after 2013, indicating the circumvention of embargo periods. Additionally, specific scholarly disciplines, notably Chemistry, exhibited a bias in download activity, suggesting barriers to access within these fields. Moreover, the study highlighted the dominance of a few key publishers, with 80% of downloads originating from only nine publishers, underscoring challenges in accessing the scholarly record.

Himmelstein et al. (2018) [19] found that Sci-Hub allows users to download PDF versions of scholarly articles, including those behind paywalls. They revealed that Sci-Hub's database encompasses 68.9% of the 81.6 million scholarly articles registered with Crossref and 85.1% from toll-access journals. Coverage varied across disciplines and publishers, with Sci-Hub preferring popular, paywalled content. Interestingly, Sci-Hub offered more excellent coverage of toll-access articles than the University of Pennsylvania, a major research institution in the United States. Conversely, green open access to toll access articles via legitimate services remained limited. Their findings indicate that nearly all scholarly literature is now accessible to anyone with an internet connection, suggesting potential challenges to the sustainability of the toll access business model.

Bodo (2020) [20] delved into analyzing weblogs from one of the Library Genesis mirrors. The objective was to unravel the social and economic forces influencing the global and European demand for unauthorized scholarly literature. The findings suggested that wealthier regions tend to be the most active users of shadow libraries. However, despite freely accessible knowledge, regions with lower

income levels face inherent structural limitations that impede their ability to leverage these illicit resources fully.

Another study by Luczaj and Holy-Luczaj (2020) [21] examined the strategies employed by academics in Central and Eastern Europe to challenge the limitations in access to scholarly literature. The findings underscored Sci-Hub's role as a form of resistance against academic capitalism and imperialism. Simultaneously, it serves as a compelling call to reform the mechanisms through which global knowledge production is accessed. Paradoxically, while Sci-Hub appears to enhance access to scientific knowledge, it diminishes the overall impact of the OA movement by diminishing the comparative advantage of OA publications in terms of researchers' visibility [5].

Behboudi (2021) [22] provided a comprehensive analysis of the evolving search behavior associated with illegal access to scholarly literature through Sci-Hub over four years, shedding light on the increasing interest in and potential consequences of Sci-Hub usage. By exploring Google search trends for the terms "Sci-Hub," "Plagiarism," and "Plagiarism Checker" globally, the findings indicated a significant surge in global searches for Sci-Hub, with China, Ethiopia, and Tunisia leading in search volumes. The study addressed the ethical implications of accessing research literature illegally, emphasizing the importance of ethics education, the establishment of ethics committees, and the need for copyright literacy training among research stakeholders. The researchers underscored the urgency of national and international efforts to curb illegal access to scholarly content.

In 2021, Singh, Srichandan, and Bhattacharya conducted a meticulous analysis of paper downloads by Indian researchers from Sci-Hub. The daily download requests from India were meticulously tracked, geotagged, and scrutinized based on discipline, publisher, country, and publication year. The findings underscore that blocking Sci-Hub in India could significantly impede the research community. Additionally, the study emphasized the imperative of developing institutional repositories, fostering collaborations with international publishers, and fortifying public-funded Indian journals [23].

Segado-Boj et al. (2022) [24] conducted an international survey to examine researchers' responses

to the persistent prevalence of paywalls and their willingness and motivation to use or refrain from scholarly piracy sites. The results revealed that over 50% of participants have turned to scholarly piracy sites at least once. Respondents who have never used pirate libraries often cite ethical and legal concerns or a lack of awareness.

In a recent study by Correa et al. (2022), the researchers found that articles downloaded from Sci-Hub were cited 1.72 times more than papers not downloaded from Sci-Hub and that the number of downloads from Sci-Hub was a robust predictor of future citations. The results suggest that limited access to publications may limit some scientific research from achieving its full impact [25].

Buehling, Geissler, and Strecker (2022) [26] delved into the impact of free access to the scientific literature on the involvement of underrepresented groups in scientific discourse, focusing on the emergence of Sci-Hub as a black open-access repository. The analysis approaches the introduction of Sci-Hub as an exogenous event that provides unrestricted access to papers usually hidden behind paywalls by combining and synchronizing data monitoring Sci-Hub access with publication data from the Web of Science (WoS). The analysis, particularly in mathematics, suggests that unrestricted access to academic knowledge through platforms like Sci-Hub will likely enhance the representation of authors from developing countries in international journals. This unrestricted access can yield benefits such as generating more original work, replicating empirical findings across diverse settings, and shifting the research focus towards topics overlooked by researchers from more developed countries.

Amin et al. (2022) [27] compared the web traffic ranking, usage, and popularity of ScienceDirect and Emerald Insight subscription databases with the pirated open-access platform Sci-Hub. The findings reveal that while ScienceDirect boasts the highest traffic rank and in-linking sites, Sci-Hub records the most page visits with the fastest download speed. Additionally, users spend less time on ScienceDirect and Emerald Insight than Sci-Hub, which also exhibits the lowest bounce rate. The researchers highlighted that users from developing and developed nations access Sci-Hub, with more visitors from developing nations.

In their recent study, Maddi, A. and Sapinho, D. (2023) [5] delved into how shadow libraries influence researchers' citation practices, particularly regarding the Open Access Citation Advantage (OACA). The study analyzed a large randomized sample and revealed that open access (OA) publications, including those in fully OA journals, tend to receive more citations than subscription-based counterparts. Nevertheless, the OACA has experienced a slight decline over the past seven years. Notably, the study introduces a distinction between publications accessible via the Sci-Hub platform and those that are not among subscription-based articles. The widespread use of Sci-Hub negates the positive impact of OA publishing. The findings indicate that publications in full OA journals suffer from the popularity of Sci-Hub. Thus, despite appearing to enhance access to scientific knowledge, Sci-Hub paradoxically undermines the OA movement by diminishing the comparative advantage of OA publications in terms of visibility to researchers.

Overall, the previous studies provide a comprehensive understanding of shadow libraries' complexities and their impact on scholarly communication, from exploring the motivations behind their usage to assessing their implications for access practices. They emphasize that although shadow libraries are illegal, publishers' general attitude seems unethical. They should reconsider their pricing policies.

4. Methodology

The dataset was collected from the Scopus database on January 2, 2024. The initial investigation revealed that it contains the most significant number of papers relevant to the subject compared to Web of Science and Dimensions databases, ensuring a more exhaustive dataset for analysis. The search strategy involved an advanced query with three limiters using Boolean operators: (title OR abstract OR keywords) AND publication date AND subject area. Multiple synonymous terms were included in the query to ensure the retrieval of all related documents. These terms were chosen as they are the most widely used in global literature. The terms used are: ("Shadow Libraries," "Shadow Library," "Pirate Libraries," "Pirate Library," "Piracy of Libraries," "Violation of

Intellectual Property Rights," "Sci-hub," "SciHub," "Library Genesis," "LibGen," and "Lib-Gen").

The analysis included all types of documents published in all languages worldwide, focusing on documents with a publication year of more than 2011 and less than 2024. This period was chosen as it coincides with the Sci-Hub launch, the most well-known global forum for academic piracy. Subject area limiters were applied to focus on disciplines relevant to publishing practices. Only documents in the fields of arts (ARTS), sociology (SOCI), and multidisciplinary (MULT) were included in the search parameters. Papers from other disciplines were excluded because they discuss the phenomenon from different aspects outside the scope of the current study. The VOSviewer software was used to construct and visualize co-occurrence networks of the terms used to express the phenomenon. The study also conducted a content analysis of the top ten most cited documents, providing a brief overview of the studied phenomenon.

5. Results

The results showed 90 documents published between 2011 and 2023, which may appear modest. This small number of publications could be attributed to the fact that it is a phenomenon rather than a field of study. The distribution of these documents by year, affiliations, authors, types, journals, and the most cited papers are covered in the following section:

5.1 Documents by Year

The bibliometric analysis of documents related to shadow libraries reveals interesting temporal patterns. From 2011 to 2023, the distribution of scholarly outputs demonstrates evolving trends and periods of heightened academic attention to the black open access phenomenon. Figure 1 shows the number of documents:

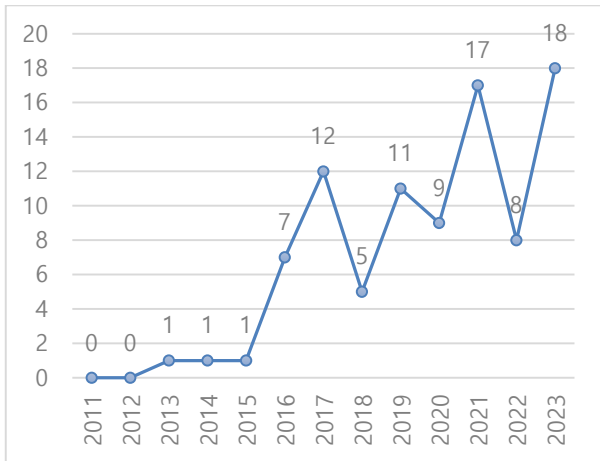


Figure 1. Distributions of documents by years

The findings indicated that the initial years, 2011 and 2012, show no recorded documents on this topic. However, in 2013, the first signs of scholarly interest emerged, with one publication in 2014 and another in 2015. The absence of recorded documents in 2011 and 2012 may suggest a lack of initial academic attention or perhaps a delayed response to the emerging challenges posed by intellectual property infringement within shadow libraries during those years.

A significant uptick in scholarly activity was observed from 2016 onwards, with seven documents in 2016 and a gradual increase in subsequent years. The peak years are 2017, 2021, and 2023, each contributing substantially to the overall dataset with 12, 17, and 18 documents. This surge may be related to pirate platform usage, particularly LibGen, and Sci-hub, which led to international interest in studying this phenomenon and its impact on publishers. According to Bohannon (2016) [16], many researchers worldwide use shadow libraries. In just six months, Sci-Hub made 28 million documents available; more than 2.6 million download requests came from Iran, 3.4 million from India, and 4.4 million from China. It has become the 3,950th most popular website worldwide [28].

The distribution also emphasizes the topic's ongoing relevance, supported by further publications in 2022 and 2023. While these years may lack the pronounced peaks seen in other periods, their consistent contribution underscores the international interest in the piracy phenomenon.

5.2 Documents by Types

The study examined the types of documents published during the study period. As presented in Figure 2, the findings revealed that articles, constituting a noteworthy 83.3% of the total documents, emerge as the principal form of scholarly contribution. This dominance underscores a commitment to comprehensive and in-depth exploration, reflecting scholarly rigor and thorough analysis in addressing the complexities of the black open-access phenomenon. Beyond articles, the dataset showcases a diverse array of document types. Reviews, accounting for 5.6%, offer comprehensive overviews, while notes (4.4%) and book chapters (3.3%) contribute specialized insights. Conference papers (2.2%), books (1.1%), and editorial materials (1.1%) provide additional perspectives, reflecting a well-rounded approach to understanding and addressing the multifaceted challenges.

Specific document types, such as news items, early access publications, and proceeding papers, are conspicuously absent from the dataset. The limited presence of early-access publications aligns with the meticulous review processes associated with intellectual property research, emphasizing a preference for finalized and thoroughly vetted contributions.

In conclusion, the prevalence of articles and the inclusion of diverse document types underscore a commitment to excellence in research and contribute to the richness and depth of the scholarly discourse on this topic.

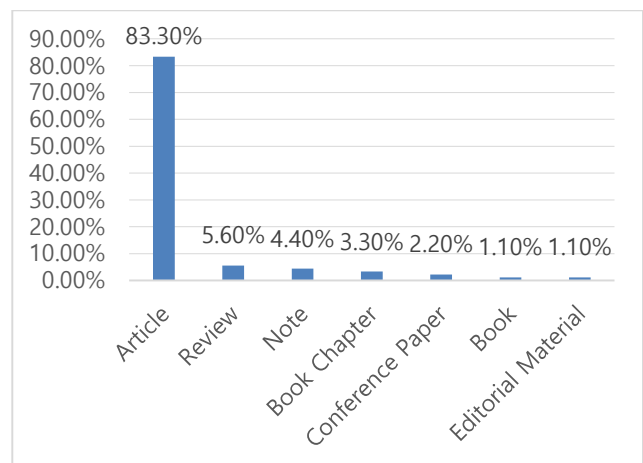


Figure 2. Distributions of documents by Types

5.3 Documents by Authors and their Affiliation

Table 1 provides an insightful examination of the authors contributing to the scholarly landscape of shadow libraries, with their respective affiliations:

Table 1 Documents by Authors and their Affiliation

Authors	Affiliation	N	%
Singh, V.K.	Banaras Hindu University, India	5	5.6
Srichandan, S.S.	Banaras Hindu University, India	5	5.6
Bhattacharya, S.	Banaras Hindu University, India	4	4.4
Abrizah, A.	Banaras Hindu University, India	3	3.3
Boukacem-Zeghmouri, C.	University Claude Bernard Lyon, France	3	3.3
Herman, E.	CIBER Research Ltd., United Kingdom	3	3.3
Nicholas, D.	CIBER Research Ltd., United Kingdom	3	3.3
Piryani, R.	University of Tsukuba, Japan	3	3.3
Rodríguez-Bravo, B.	Universidad de León, Leon, Spain	3	3.3
Xu, J.	Wuhan University, Wuhan, China	3	3.3

The findings show that authors Singh, V.K., and Srichandan, S.S., affiliated with Banaras Hindu University in India, emerge as prolific authors, each having authored five documents, constituting 11.2% of the total publications. Following closely, Bhattacharya, S., also from Banaras Hindu University, contributes with four publications (4.4%), further emphasizing the institution's substantial involvement in this research domain. This presence of multiple authors affiliated with the same institution suggests a collaborative approach within Banaras Hindu University, which can enrich the quality and depth of research on this topic.

The findings highlight the global nature of the black open-access phenomenon. This diversity is reflected in the authors' diverse affiliations. Boukacem-Zeghmouri from Université Claude Bernard Lyon in France, Herman and Nicholas from CIBER Research Ltd. in the United Kingdom, Piryani from the University of Tsukuba in Japan, Rodríguez-Bravo from Universidad de León in Spain, and Xu

from Wuhan University in China represent the international collaboration and interest in this phenomenon. Each contributes 3.3% of the total documents published during the study period.

5.4 Documents by country/regions

The analysis of documents by region, as outlined in Figure 3, provides valuable insights into the geographical distribution of contributions, aiming to delineate the regional dynamics shaping the discourse on this phenomenon.

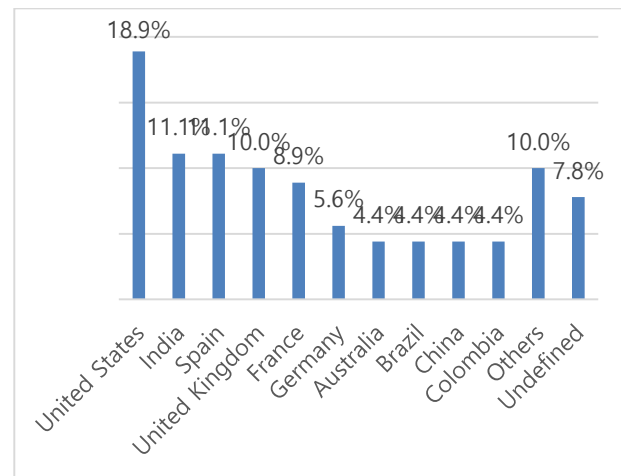


Figure 3. Distributions of documents by Regions

The findings show that the United States is a focal point in this scholarly discourse related to shadow libraries, contributing 18.9% of the total documents. This prominence aligns with the country's robust academic environment and its pivotal role in the global landscape of intellectual property discussions. India and Spain share the second position, contributing 11.1% of the documents. This percentage underscores the global nature of the black open-access phenomenon, with diverse regions actively engaging in scholarly research. The United Kingdom and France follow closely, contributing 10% and 8.9%, respectively. These findings emphasize the significant role of European nations in the academic exploration of intellectual property challenges. Germany, Australia, Brazil, China, and Colombia each contribute 4.4%, reflecting a distributed global engagement with the subject matter. These countries showcase the diverse perspectives and contributions emerging from various regions worldwide.

Other contributions from different countries include Japan, Nigeria, Portugal, Mexico, Peru, Ukraine, Bahrain, Austria, South Africa, Italy, Israel, Iran, and more. This category, collectively constituting 10%, underscores the rich global tapestry of intellectual property research, demonstrating a broad spectrum of participation and perspectives. Additionally, an undefined percentage of 7.8% highlights instances where the specific country or region is not explicitly identified, suggesting potential areas for further refinement in data classification. In conclusion, the geographical distribution of scholarly contributions reveals a diverse and globally distributed research landscape.

5.5 Documents by languages

The analysis of documents by language was examined. Figure 4 shows a nuanced perspective on the linguistic dimensions of scholarly contributions in the realm of the studied phenomenon:

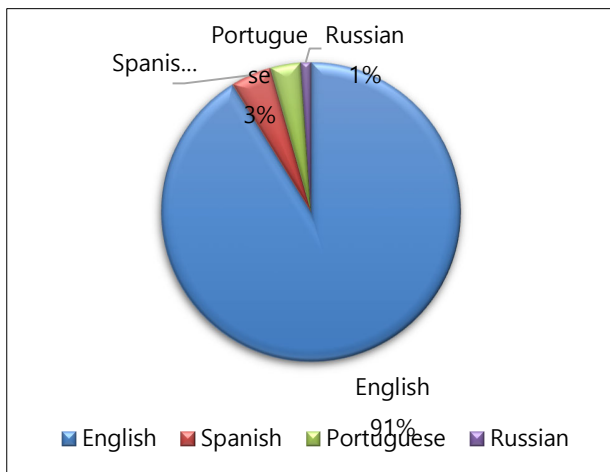


Figure 4. Distributions of documents by Languages

The findings reported that English is the predominant language, constituting a substantial 92.2% of the total documents, aligning with the widespread use of English in global academic communication. The dataset also exhibits a degree of linguistic diversity. Spanish, Portuguese, and Russian Contributions collectively account for 8.8% of the total documents. The presence of documents in Spanish and Portuguese underscores an inclusive approach, ensuring that researchers from diverse linguistic backgrounds contribute to the exploration of

intellectual property infringement across different linguistic and cultural contexts.

The inclusion of documents in multiple languages has implications beyond linguistic diversity. It broadens the scope of intellectual property research and facilitates global dialogue. Researchers utilizing different languages contribute to a richer understanding of the phenomenon.

5.6 The Top keywords

The authors' most commonly used keywords for the shadow library phenomenon were examined. Figure 5 visualizes the prominent keyword clusters based on their occurrences and weights:

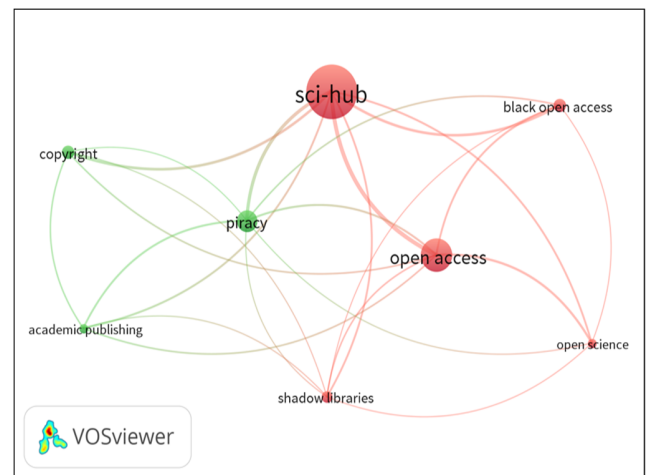


Figure 5. Visualization of top keywords

The cluster weights emphasize the significance of specific keywords, with "sci-hub" and "open access" having significant weights. With the most prominent weight (33), Sci-Hub is the most used keyword in the global literature, indicating that it has received much scholarly attention. Open access follows closely, reflecting a keen interest in the open-access movement (weight = 19). Other keywords to discuss this phenomenon include piracy, black open access, copyright, shadow libraries, and academic publishing. These clusters capture the primary topics and focal points covered in scholarly discussions of the piracy of information resources. The interconnectedness of keywords on the map underscores the interrelated nature of different facets of the shadow libraries phenomenon.

5.7 Top Journals

The study analyzed the key journals contributing to the scholarly discourse on the shadow library phenomenon. Table 2 provides a comprehensive overview of the top ten journals:

Table 2 The top ten journals

Journals	N	%	2022 Rank	2022 Cite Score	2022 SJR
Information Development	4	4.4	#38 of 266	5.2	0.559
Scientometrics	4	4.4	#18 of 262	6.0	1.019
Information Services and Use	3	3.3	#173 of 266	0.7	0.176
Journal of Scientometric Research	3	3.3	#110 of 266	1.7	0.281
Learned Publishing	3	3.3	#72 of 493	4.3	0.758
Plos One	3	3.3	#17 of 134	6.0	0.885
Serials Librarian	3	3.3	#158 of 266	0.8	0.223
College and Research Libraries News	2	2.2	#186 of 266	0.6	0.244
First Monday	2	2.2	#155 of 359	3.4	0.233
Insights the Uksg Journal	2	2.2	#88 of 266	2.2	0.389

The analysis underscores that information development and scientometrics take the lead, contributing 4.4% to publications in shadow libraries. Information development, positioned at #38 out of 266, exhibits a CiteScore of 5.2 and an SJR of 0.559. At the same time, scientometrics, at #18 out of 262, boasts a CiteScore of 6.0 and an SJR of 1.019, signaling its substantial influence in the academic community. Information Services and Use and the Journal of

scientometric research share a 3.3% contribution, with ranks of #173 and #110 out of 266, respectively. Despite a comparatively lower impact, information services and use (CiteScore: 0.7, SJR: 0.176) and the Journal of scientometric research (CiteScore: 1.7, SJR: 0.281) play pivotal roles in disseminating relevant research. Learned publishing and plos one also contributes 3.3% each, holding ranks of #72 and #17 out of 493 and 134, respectively, with significant CiteScores (4.3 and 6.0) and SJRs (0.758 and 0.885). Serials Librarian, with a 3.3% contribution, ranks at #158 out of 266, with a modest CiteScore of 0.8 and an SJR of 0.223, contributing crucially to the field. In conclusion, information development and scientometrics, with their high rankings and impactful metrics, stand out as prominent platforms publishing extensively on shadow libraries, reflecting their pivotal roles in shaping scholarly discussions in this domain.

The top ten journals identified through this analysis serve as vital platforms for disseminating scholarly work on intellectual property infringement within shadow libraries.

5.8 The most cited papers

The high citation counts indicate papers' enduring relevance and influence in shaping scholarly discussions and guiding future research on this phenomenon. Based on the papers with the highest citations, the results shown in Table 3 list the top ten most cited papers discussing shadow libraries, as retrieved from the Scopus database:

Table 3 The top 10 most cited papers

Title	Authors	Year	Cited by
Ethical Decision Making in a Peer-to-Peer File Sharing Situation: The Role of Moral Absolutes and Social Consensus	Bateman C.R.; Valentine S.; Rittenburg T.	2013	31
Piracy of scientific papers in Latin America: An analysis of Sci-Hub usage data	Machin-Mastromatteo J.D.; Uribe-Tirado A.; Romero-Ortiz	2016	25
Bibliogifts in LibGen? A study of a text-sharing platform driven by biblioleaks and crowdsourcing	Cabanac G.	2016	25
My love-hate of sci-hub	McNutt M.	2016	23
The postdigital challenge of redefining academic publishing from the margins	Jandrić P.; Hayes S.	2019	22
Sci-Hub: The new and ultimate disruptor? View from the front	Nicholas D.; etc	2019	22
Use, knowledge, and perception of the scientific contribution of Sci-Hub in medical students: Study in six countries in Latin America	Mejia C.R.; etc.	2017	21
US–Japan Trade Frictions: The Past, the Present, and Implications for the US–China Trade War	Urata S.	2020	17
Sci-Hub: What Librarians Should Know and Do about Article Piracy	Hoy M.B.	2017	16
Sci-hub, a challenge for academic and research libraries	González-Solar L.; Fernández-Marcial V.	2019	15

6. Conclusion

This bibliometric study extensively investigates the scholarly landscape encompassing the black open access phenomenon within shadow libraries from 2011 to 2023, depending on the Scopus database. Future research could analyze the global literature of databases like WOS and Dimensions. The results revealed intriguing temporal patterns, notably increased scholarly activity post-2016, peaking in 2017, 2021, and 2023, indicative of global interest in comprehending the phenomenon's implications for publishers. This increased focus is consistent with the use of pirate sites like Sci-Hub and LibGen, which has been noticed. Most articles highlight a dedication to in-depth investigation, reflecting the critical examination needed to handle the complexity of intellectual property violations.

The author's contributions and affiliations, particularly those of Singh, V.K., and Srichandan, S.S., from Banaras Hindu University, demonstrate the collaborative approaches within the institution. The global nature of intellectual property infringement discourse is evident through diverse author affiliations, showcasing international collaboration. Geographic distribution shows that the United States, Spain, and India are the main contributors. European countries play essential roles in this complex subject, demonstrating worldwide participation. Linguistic characteristics show that English is the most commonly used language, with contributions also made in Spanish, Portuguese, and Russian, reflecting an inclusive language approach.

Top journal analysis underscores the pivotal role of information development and scientometrics, showcasing their substantial influence in shaping scholarly discussions. Examination of the most frequently cited papers reveals extensive social, technological, and ethical investigation. According to co-occurrence network visualization, "Sci-Hub" and "open access" are the most popular keywords worldwide, suggesting scholarly interest in Sci-Hub and the wider ramifications of unauthorized access in the open-access movement.

This study comprehensively overviews the scholarly discourse on intellectual property infringement within shadow libraries. The nuanced analysis of temporal, geographical, linguistic, and authorship dimensions provides a holistic

understanding of the scholarly discourse, providing a foundation for future research in the evolving academic publishing landscape. It underscores the pressing need for publishers to address the challenges posed by shadow libraries, which continue to erode traditional revenue streams. Publishers must adapt by implementing more robust access controls and exploring innovative business models to protect intellectual property rights while meeting the evolving needs of researchers.

7. Recommendation

Based on the results of the study, the following recommendations can be suggested to guide future research and decision-makers in academia and research communities toward a more collaborative and ethically grounded approach to addressing the black open-access phenomenon within shadow libraries:

- Publishers must discover innovative business models to safeguard intellectual property rights while meeting researchers' evolving needs.

- Support and promote open-access initiatives to enhance the accessibility of scholarly publications. Initiatives such as reduced subscription costs, increased access to legal databases, and collaborative programs can reduce reliance on unauthorized platforms.

- Promote awareness of the ethical and legal dimensions of piracy platforms.

- Encourage collaboration between public and private sectors to address access, subscription costs, and intellectual property challenges. Engaging publishers, libraries, and governmental bodies can lead to innovative solutions.

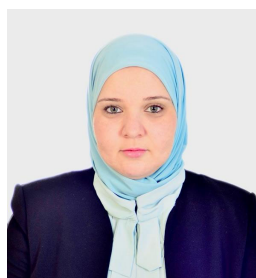
Acknowledgment

The author extends her appreciation to the Deanship of Scientific Research at Northern Border University, Arar, KSA, for funding this research work through the project number "NBU-FFR-2024-2949-01"

References

- [1] Sample I. (2012). Harvard University says it can't afford journal publishers' prices. *The Guardian*. <https://www.theguardian.com/science/2012/apr/24/harvard-university-journal-publishers-prices>
- [2] Albert, K. M. (2006). Open access: Implications for scholarly publishing and medical libraries. National Center for Biotechnology Information. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1525322/>
- [3] Swanberg, S. M., Bulgarelli, N., & Thielen, J. (2020). Faculty knowledge and attitudes regarding predatory open access journals: A needs assessment study. *Journal of the Medical Library Association*, 108(2), 208-218–218. <https://doi-org.sdl.idm.oclc.org/10.5195/jmla.2020.849>
- [4] Sheikh, A., & Richardson, J. (2023). Open access movement in the scholarly world: Pathways for libraries in developing countries. *JOURNAL OF INFORMATION SCIENCE*. <https://doi-org.sdl.idm.oclc.org/10.1177/01655515231202758>
- [5] Maddi, A. and Sapinho, D. (2023) 'On the culture of open access: The Sci-hub paradox', *Scientometrics: An International Journal for all Quantitative Aspects of the Science of Science, Communication in Science and Science Policy*, 128(10), pp. 5647–5658. [doi:10.1007/s11192-023-04792-5](https://doi.org/10.1007/s11192-023-04792-5)
- [6] Lawson, S. (2017). Access, ethics and piracy. *Insights: The UKSG Journal*, 30(1), 25–30. <https://doi-org.sdl.idm.oclc.org/10.1629/uksg.333>
- [7] Bohannon, J. (2016). Who's Downloading Pirated Papers? Everyone. *Science*, 352(6285), 508–512. <https://search-ebscohost-com.sdl.idm.oclc.org/login.aspx?direct=true&db=edsjsr&AN=edsjsr.24744838&site=eds-live>
- [8] Van der Sar, E. (2016b). RuTracker and Sci-Hub nominated for Free Knowledge Award. <http://torrentfreak.com/rutracker-and-sci-hub-nominated-for-free-knowledge-award-160423>
- [9] González-Solar, L., & Fernández-Marcial, V. (2019). Sci-Hub, a challenge for academic and research libraries. 28(1). <https://search-ebscohost-com.sdl.idm.oclc.org/login.aspx?direct=true&db=edseli&AN=edseli.34165&site=eds-live>
- [10] Nicholas, D., Boukacem-Zeghmouri, C., Xu, J., Herman, E., Clark, D., Abrizah, A., Rodríguez-Bravo, B., & Świgoń, M. (2019). Sci-Hub: The new and ultimate disruptor? View from the front. *Learned Publishing*, 32(2), 147-153–153. <https://doi-org.sdl.idm.oclc.org/10.1002/leap.1206>
- [11] Cabanac, G. (2016). Bibliogifts in LibGen? A study of a text-sharing platform driven by biblioleaks and crowdsourcing. *Journal of the Association for Information*

- Science and Technology, 67(4), 874-884-884. <https://doi-org.sdl.idm.oclc.org/10.1002/asi.23445>
- [12] The Guardian. (2023). Four large US publishers sue shadow library for alleged copyright infringement. <https://www.theguardian.com/books/2023/sep/15/four-large-us-publishers-sue-shadow-library-for-alleged-copyright-infringement>
- [13] Novo, L. A. B., & Onishi, V. C. (2017). Could Sci-Hub become a quicksand for authors?. *Information Development*, 33(3), 324-325-325. <https://doi-org.sdl.idm.oclc.org/10.1177/0266666917703638>
- [14] Strielkowski, W. (2017). Will the rise of Sci-Hub pave the road for the subscription-based access to publishing databases?. *Information Development*, 33(5), 540-542-542. <https://doi-org.sdl.idm.oclc.org/10.1177/0266666917728674>
- [15] Zimmerman, M. (2011). E-books and piracy: Implications/issues for academic libraries. *New Library World*, 112(1-2), 67-75-75. <https://doi-org.sdl.idm.oclc.org/10.1108/03074801111100463>
- [16] Bohannon, J. (2016). The frustrated science student behind Sci-Hub. *Science*, 352(6285), 511. <https://doi.org/10.1126/science.352.6285.511>
- [17] Machin-Mastromatteo, J.-D., Uribe-Tirado, A., & Romero-Ortiz, M. E. (2016). Piracy of scientific papers in Latin America: An analysis of Sci-Hub usage data = La piratería de artículos científicos en Latinoamérica: Un análisis de los datos de uso de Sci-Hub. *32(5)*, 1806-1814. <https://doi-org.sdl.idm.oclc.org/10.1177/0266666916671080>
- [18] Greshake, B. (2017). Looking into Pandora's Box: The Content of Sci-Hub and its Usage. *F1000Research*, 6. <https://doi-org.sdl.idm.oclc.org/10.12688/f1000research.11366.1>
- [19] Himmelstein, D. S., Greene, C. S., Romero, A. R., Levernier, J. G., Munro, T. A., McLaughlin, S. R., & Greshake Tzavaras, B. (2018). Sci-Hub provides access to nearly all scholarly literature. *ELife*, 7. <https://doi-org.sdl.idm.oclc.org/10.7554/eLife.32822>
- [20] Bodó, B., Antal, D., & Puha, Z. (2020). Can scholarly pirate libraries bridge the knowledge access gap? An empirical study on the structural conditions of book piracy in global and European academia. *PLoS ONE*, 15. <https://doi-org.sdl.idm.oclc.org/10.1371/journal.pone.0242509>
- [21] Łuczaj, K., & Holy-Łuczaj, M. (2020). How academics outside global research centers challenge limitations in access to scholarly literature. *Portal*, 20(4), 715-731-731. <https://doi-org.sdl.idm.oclc.org/10.1353/PLA.2020.0034>
- [22] Behboudi, E., Shamsi, A., & Bueno de la Fuente, G. (2021). The black crow of science and its impact: analyzing Sci-Hub use with Google Trends. *LIBRARY HI TECH*, 39(4), 970-983. <https://doi-org.sdl.idm.oclc.org/10.1108/LHT-04-2020-0105>
- [23] Singh, V. K., Srichandan, S. S., & Bhattacharya, S. (2021). What do Indian Researchers download from Sci-Hub. <https://doi.org/10.5530/jscires.10.2.4>
- [24] Segado-Boj, F., Martin-Quevedo, J., & Prieto-Gutierrez, J.-J. (2022). Jumping over the paywall: Strategies and motivations for scholarly piracy and other alternatives. <https://doi-org.sdl.idm.oclc.org/10.1177/02666669221144429>
- [25] Correa, J. C., Laverde-Rojas, H., Tejada, J., & Marmolejo-Ramos, F. (2022). The Sci-Hub effect on papers' citations. *Scientometrics: An International Journal for All Quantitative Aspects of the Science of Science, Communication in Science and Science Policy*, 127(1), 99-126. <https://doi-org.sdl.idm.oclc.org/10.1007/s11192-020-03806-w>
- [26] Buehling, K., Geissler, M., & Strecker, D. (2022). Free access to scientific literature and its influence on the publishing activity in developing countries: The effect of Sci-Hub in the field of mathematics. *Journal of the Association for Information Science and Technology*, 73(9), 1336-1355-1355. <https://doi-org.sdl.idm.oclc.org/10.1002/asi.24636>
- [27] Amin, R., Ayoub, A., Amin, S., & Wani, Z. A. (2021). Toll-based access vs pirate access: a webometric study of academic publishers. *Digital Library Perspectives*, 38(1), 33-44. <https://doi-org.sdl.idm.oclc.org/10.1108/DLP-12-2020-0127>
- [28] Gardner, G. J., McLaughlin, S. R., & Asher, A. D. (2017). Shadow Libraries and You: Sci-Hub Usage and the Future of ILL. <https://search-ebcohst-com.sdl.idm.oclc.org/login.aspx?direct=true&db=edseli&AN=edseli.30981&site=eds-live>



Safinaz Elroukh is an assistant professor at Northern Border University, K.S.A., at the Department of General Courses. She is also a lecturer at the Libraries and Information Science Department at Menoufia University, Egypt. Her research interests include artificial intelligence, international publishing, and information behavior. She is a member of the American Library Association (A.L.A.) and the Egyptian Library Association (E.L.A.)