

Article

Barriers to Open Access Publishing: Views from the Library Literature

Amy Forrester

School of Information Sciences, College of Communication and Information, University of Tennessee, 1345 Circle Park Drive, Com 451, Knoxville, TN 37996-0341, USA; E-Mail: aforres4@vols.utk.edu; Tel.: 202-288-8504

Academic Editor: Remedios Melero

Received: 30 May 2015 / Accepted: 02 September 2015 / Published: 3 September 2015

Abstract: The library and information science (LIS) community has an active role in supporting access to information and, therefore, is an important stakeholder in the open access conversation. One major discussion involves the barriers that have hindered the complete transition to open access in scientific publications. Building upon a longitudinal study by Bo-Christer Björk that looked at barriers to the open access publishing of scholarly articles, this study evaluates the discussion of those barriers in the LIS literature over the ten year period 2004–2014, and compares this to Björk’s conclusions about gold open access publishing. Content analysis and bibliometrics are used to confirm the growth of the discussion of open access in the past ten years and gain insight into the most prevalent issues hindering the development of open access.

Keywords: open access; information science; libraries; scholarly publishing; barriers

1. Introduction

The discussion around open access publishing—the free and unrestricted online availability of scholarly literature—has been persistent for over two decades (e.g., Harnad [1]). According to the Budapest Open Access Initiative [2], open access publishing should pose no barrier to a reader other than having access to the Internet. However, while many scholarly journals have embraced open access and the number of open access articles published has grown, there is still much deliberation around the barriers to open access [3].

Libraries and information professionals have always been involved in supporting access to information and knowledge and they play a vital and active role in the success of the open access movement [4,5]. One way to gain an understanding of the discourse around open access is to study the views of the library and information science (LIS) community. The comprehensive role of the LIS professional in open access (*i.e.*, as creator, advocate, consumer, educator, developer) makes them a unique and comprehensive model to measure the overall climate of open access.

In a 2004 study, Bo-Christer Björk [6] explored the barriers that have hindered the complete transition to open access in scientific publications. He then revisited the analysis ten years later to assess the shift over time [7]. He identified six main types of barriers: legal framework, IT-infrastructure, business models, indexing services and standards, academic reward system, and marketing and critical mass. Björk used anecdotes and secondary sources to illustrate open access conditions in 2003 and leveraged data from published studies to report the update.

Borrowing Björk's [6] six types of barriers to open access, this study will evaluate the discussion of barriers in the LIS literature over the ten year period 2004–2014 and compare this to Björk's conclusions. As a proxy for the LIS community dialogue of open access, the research set will examine journals articles from an established LIS database that are indexed with the subject term "open access". This study will first describe characteristics of the research sample such as, publishing models and author traits. Using bibliometrics will enable the detection of trends by measuring changes over time [8,9].

The second phase of this study will specifically investigate the subject of open access barriers within this dataset. Content analysis, a research method that has been used to understand a document's content and make inferences from the data about its context, can be used to gain insight into the development of the six barriers within the literature [10]. This can provide knowledge about the focus of a discipline over a period of time, as it indicates subject trends and major issues that occupy the discourse [11]. Previous LIS research has typically questioned what topics were being discussed within the literature to identify emerging patterns [12,13]. This study, however, will employ directed content analysis, which is a deductive method based on prior research to support or extend an existing theory [14]. This type of content analysis will utilize Björk's [6] existing barrier types as the initial coding categories and employ a coder's interpretations (software algorithms) of the meaning of the content set by identifying words and phrases in the abstracts that define the categories [15]. Björk's [7] conclusion states that the majority of the barriers are lower today than ten years ago. This study builds upon Björk's research to analyze the LIS literature and answer the following research questions: How can the attention to barriers to gold open access be explored using LIS literature; How does this discussion compare to Björk's results; How has the focus on these barriers to open access among the authors of the LIS community changed over time?

2. Previous Research

There have been many articles studying the development of open access in the LIS journal literature. The majority use descriptive statistics or bibliometrics to examine publishing characteristics of LIS-related journal publications by analyzing entire journal title contents. Way [16] and Singh, Shah, and Gul [17] report on the availability and growth of open access journals among all of the LIS identified journals from Ulrichsweb: Global Serials Directory (Ulrichs). Many more studies analyze only open-access

LIS-related journals by aggregating appropriate titles from periodical directories, e.g., Directory of Open Access Journals (DOAJ), based on the LIS subject classification [3,18–22]. They describe data such as, the publication's language, distribution, indexing coverage, country of origin, publishing models and licensing, and authorship patterns. Singh and Chikate [23] limit their open access-LIS study to a particular geographic region (Asia) and Yuan and Hua [24] use similar methodology to only research scholarly impact of LIS open access journals demonstrating examples of more narrowly focused open access-LIS studies.

Another method has been to use bibliometrics to examine open access development by drawing random samples of articles from bibliographic databases over time. There are two papers that describe using this method to study the issue of “open access” within LIS literature, not geographically limited. Liu and Wan [25] were the first to survey publication trends of scholarly journal articles on open access in the LIS literature from 2000 to 2005. This study used open access related search terms to extract articles from databases, such as Library and Information Science Abstracts (LISA) and Social Sciences Citation Index, as well as from bibliography lists. The authors analyzed the content by journal type, article type, author type, country type, and content category. Grandbois and Beheshti [26] searched the LISA database for the term “open access” in the title of articles from 2003 to 2011. This study additionally limited their search to English language and peer reviewed journals and reported on availability of the articles, characteristics of the articles and authors, publication trends, and correlations between these attributes.

3. Data and Methods

In this study, EBSCO's Library and Information Science Source (LISS) was used to retrieve data from 1994–2014. To get a thorough view of open access within the LIS community, the data (literature or articles) needed to be collected across a wide breadth of journal titles. LISS was selected because it is a comprehensive bibliographic database in the field of LIS that indexes more than 1,700 publications, including Library, Information Science and Technology Abstracts and H.W. Wilson's Library Literature and Information Science Index, a long time a key resource in LIS [27,28]. Previous studies that examined open access in the LIS literature also used subject databases to collect data, but they extracted articles from smaller databases, such as LISA or broader indexes, such as Social Sciences Citation Index [16,25,26,29].

To investigate the express issue of open access, the search term “open access” was used to retrieve all relevant literature by limiting the term to only search the subject field (SU) in the LISS database. Rather than searching by thesaurus term, SU was chosen because according to EBSCO [30], SU is one of the search fields that is common to every database and LISS is actually a combination previously existing databases. In addition, not all SU terms are listed in the database's thesaurus authority file [31]. To illustrate, there are only two open access related terms available in the thesaurus: “Open access publishing” and “Open access publishing—Finance”. By more broadly searching in the subject field, the results were not limited and included SU terms, such as, “Open access publishing—Evaluation” and “Open access publishing—Research”.

There is much variation in results across other search field options in this database (see Figure 1). Grandbois and Beheshti [26] chose to search for the term “open access” in the article title in the LISA database. However, subject indexing can add search precision to results by providing control for

synonyms, homographs, and related terms [32]. Using this strategy assumes an accurate retrieval of papers on the subject of open access, eliminating articles that use the term “open access” in a different context. For example, the following article has “open access” in the title and abstract but does not discuss open access publishing:

Article Title: *Open access* for ill and carers¹

Abstract: The article reports on a 2013 decision which the British journal publisher, Wiley, made to join a multi-partner program that allows patients and their families free access to *open access* articles on medical conditions and their treatment.

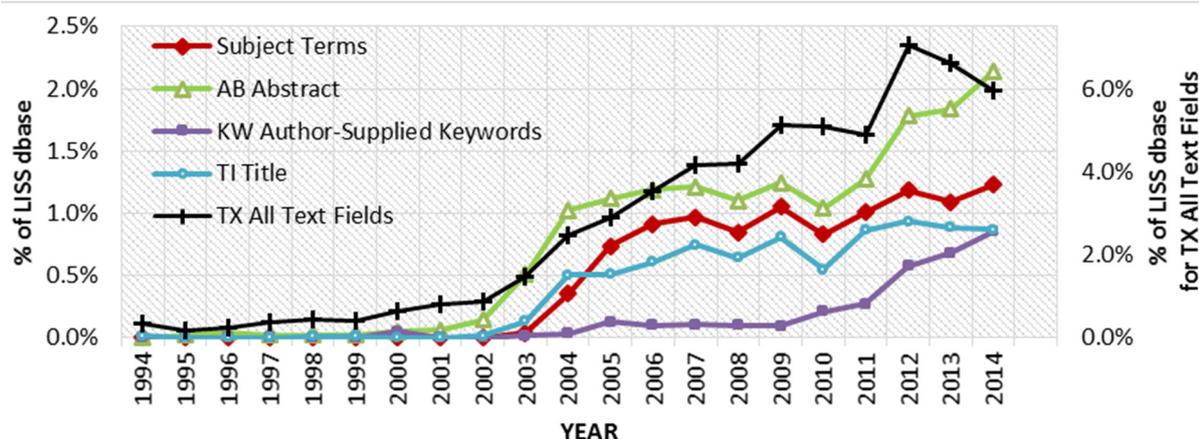


Figure 1. Appearance of search term “open access” across different search fields as a percent of the Library and Information Science Source (LISS) database. Limited to journal articles only.

However, it is important to note that there are inherent problems with the subject indexing process which can result in missed indexing [33]. Using subject indexing to generate the sample data does not represent an absolutely complete corpus of open access publishing related papers. For example, the article, “Publication fees for open access journals: Different disciplines-different methods”² does not have “open access” as a subject term, but the article examines the percentage of articles in DOAJ that charge authors to publish.

The searches in LISS also limited publication type to academic journal. Per the LISS database coverage list from EBSCO, academic journals represent 50% of the database title coverage and of those, 50% are listed as peer-review. Unlike previous studies, searches were not restricted to peer-review only nor was language limited to English only [26].

To do longitudinal text analysis, abstracts of the entire search results were exported into Excel for each of the years: 2004, 2006, 2008, 2010, 2012, 2014 and were downloaded using the LISS interface record manager tool. Although the LISS database contains full text records, not all records in the result set included the full text and represented only half of the available search results. In addition to establishing a large enough sample size, it is generally accepted that the abstract of a journal article states

¹ Jardim, C.; Pakenham, K.I. Open access for ill and carers. *CILIP Update* **2013**, *12*, 6.

² Kozak, M.; Hartley, J. Publication fees for open access journals: Different disciplines-different methods. *J. Am. Soc. Inf. Sci. Technol.* **2013**, *64*, 2591–2594.

important ideas found in the body of the article and are an accepted surrogate for the content of a research paper [34]. Table 1 illustrates the distribution of abstract counts used in the content analysis.

Table 1. Sequential distribution of abstract counts in the content analysis dataset.

YEAR	# Abstract
2004	33
2006	108
2008	122
2010	120
2012	172
2014	154
Total	709

This study began in 2004 because the appearance of the term “open access” as a subject term did not occur until 2003, which only generated three articles. After removing duplicates and non-English abstracts, 709 cases (abstracts or records) were imported into QDA Miner software and subsequently evaluated in the built-in WordStat Content Analysis program. Applying the extraction tool, 1019 two to five word phrases with a minimum frequency of three and a significant list of keywords with frequency greater than twenty (Appendix 1) were generated. Using Keywords-in-Context³ as a guide, phrases were selected to characterize each of Björk’s barrier types to create a code dictionary (Appendix 2). Text classification was run on the entire dataset to tag each abstract with the corresponding code. Records were examined to ensure the context of the code was correct and manually code additional records using single keyword searches. Number and percentage of cases for each barrier type were calculated.

For a temporal comparison of article characteristics, full citations for the entire search results were exported into Excel for each of the years: 2004, 2009, and 2014. To compare publishing models, open access or subscription publisher information was added using Ulrichs. As this information was collected in 2015, discrepancies could exist with earlier data (2004 or 2009) if a title shifted from subscription to open access since it would be recorded as open access. The extent of this was not investigated.

To assess author characteristics, author affiliation information was collected directly from individual articles. To maximize the data set, but maintain a uniform sampling size, 37, 35, and 36 records from 2004, 2009, and 2014 were examined respectively as 37 was the total number of records in 2004. Geographic location of the author was identified and author’s professional position or affiliation was categorized into sectors. Authors affiliated with a library or who maintained an information science position were tagged as LIS Community. Academic (non-LIS) included professors, administrative, and researchers working in any other discipline (e.g., engineering, computer science). Any author associated with a publisher or society was labelled Publisher/Society and those identified as students were also categorized. Excel software was used to describe the data.

³ A tool to utilize the surrounding words to understand the underlying meaning of the identified keyword.

4. Results

4.1. Bibliometric Analysis

Figure 2 shows the difference in publishing models for the representative sample of articles in the LIS literature discussing open access. While there appears to be a slight increase (9%) in publishing in open access journals over the past ten years, this is still a very small percentage (25%) of the articles examining open access as a topic overall. Indeed, over the last five years, there has been virtually no increase in “open access” articles being published in open access journals. These results compliment Grandbois and Beheshti’s [26] analysis of 203 “open access” articles from 2003–2011 in which they reported 25% were published in open access journals.

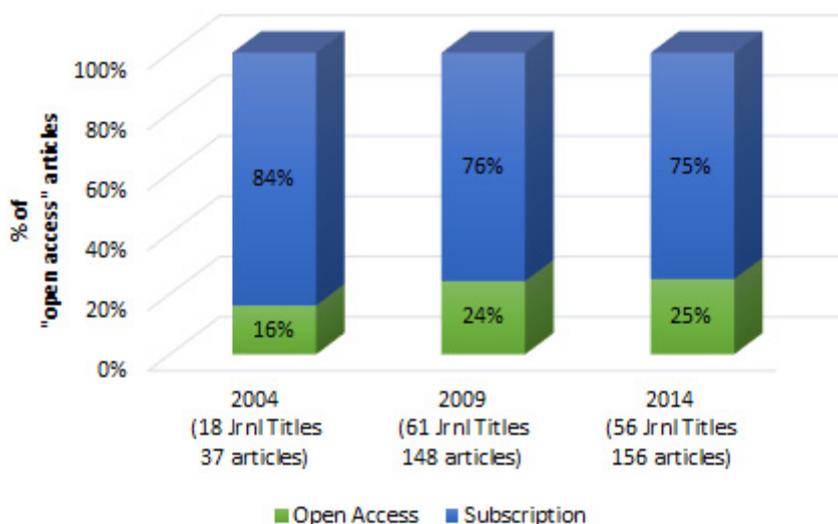


Figure 2. Percentage of journal articles with subject term “open access” by publishing model.

The geographic distribution of authors discussing open access can be seen in Figure 3. In 2004, 98% of the authors of articles in LISS with the subject term “open access” were from North America and Europe. North American authors (56%) only had a slight advantage over European authors (42%). Over the next five and ten years, these two regions still comprised the majority of authors, but the overall percentage dropped to 73% and 74% respectively. The remaining approximately 25% of authors represented thirteen countries in 2009 and nine countries in 2014. While the authors are predominantly from North America and Europe, there is an interesting positive trend of Indian authors (Figure 4). This correlates to an overall increase in open access initiatives and publishing channels in India [35]. For example, over the period 2007–2011, the number of Indian open access journals increased by nearly 180% while the total number of all open access journals only increased by 58% [36,37]. However, Singh *et al.* [17] demonstrate that Indian journals in 2014 only comprise 5% of all LIS journals.

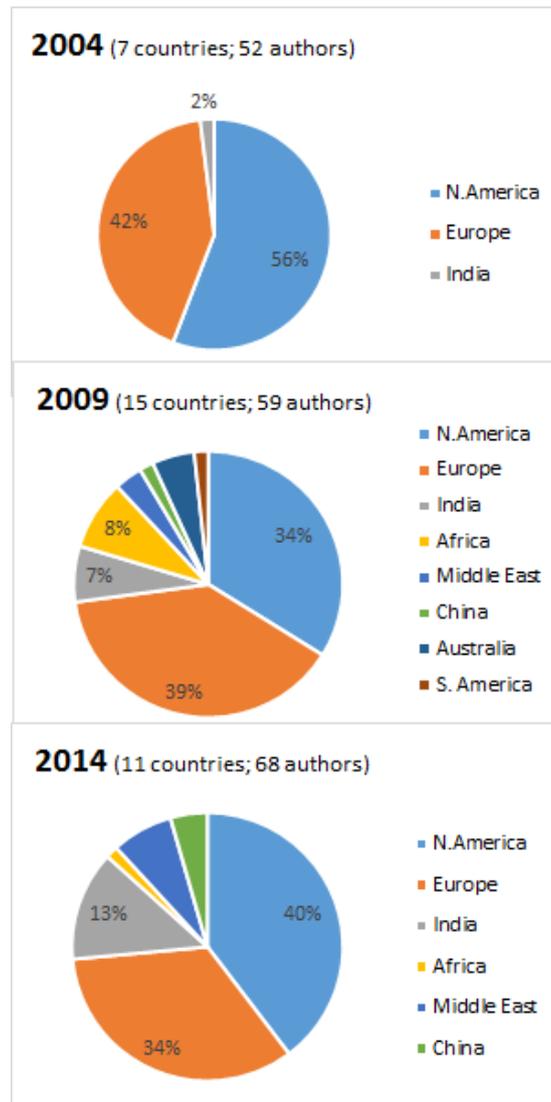


Figure 3. Geographic distribution of authors of journal articles with subject term “open access”.

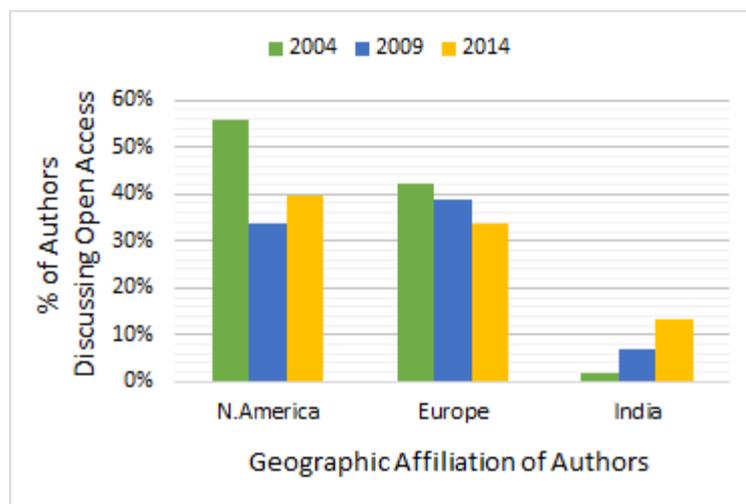


Figure 4. Growth trend of top three geographic regions of authors of journal articles with subject term “open access”.

Figure 5 presents a view of the author’s affiliation sector among the open access literature across the three years. The majority of author’s discussing open access are from the LIS community and the percentage of authors from the LIS community has not changed over the past five years. At the same time, there appears to be a decrease in the number of non-LIS academic authors publishing about open access in the LIS literature. Although Liu and Wan [25] used slightly different parameters to classify their author types, for 2004 they reported similar percentages for LIS community (37%) and Academics (31%). The results in Figure 5 show 23% for publisher/society affiliation which is higher than Liu and Wan’s [25] 16% for publishing professionals, but their study did not include author’s affiliated with societies in this category.

Is it also important to note that the goal of the bibliometrics analysis was to describe the overall characteristics of the data, the process of which was quite labor intensive. The reported data only represents analysis of a subset of the total dataset, however the results did corroborate with previous studies.

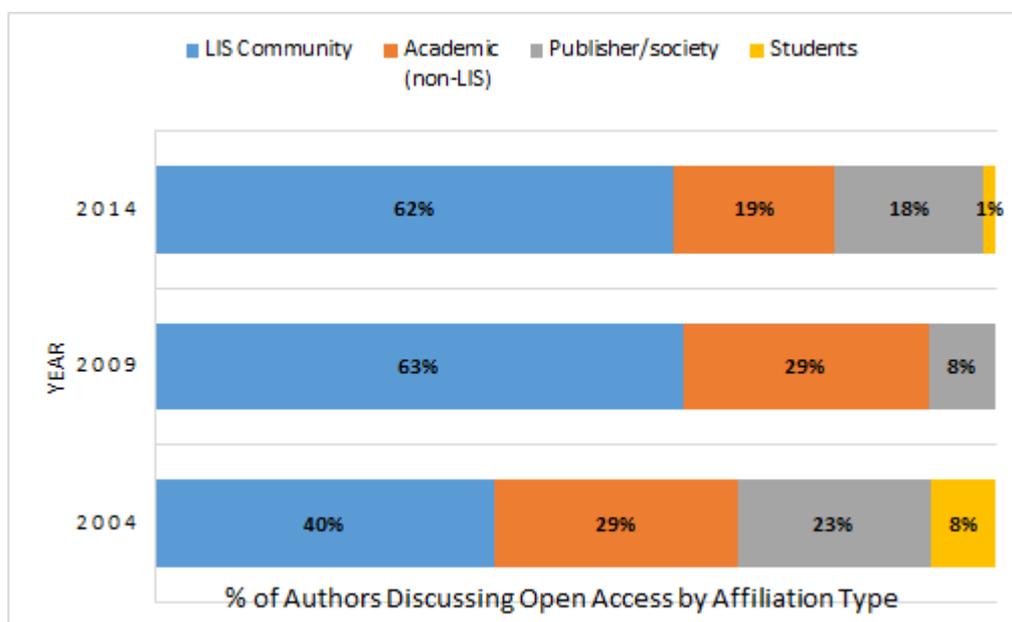


Figure 5. Percentage of journal articles with subject term “open access” by author affiliation type.

4.2. Content Analysis

Out of 709 article abstracts with subject term “open access,” 72% were classified with “barrier” codes. Figure 6 shows the percentage of abstracts coded for each barrier type over all the years combined (2004–2014). Almost 30% of all articles were classified as business models, which is more than two times greater than all other barrier types. Marketing and critical mass, IT infrastructure, and legal framework each classified only 8% of the abstracts.

The percent change in the number of abstracts classified by each barrier type over the entire ten year period can be seen in Figure 7. There is a decrease in the percentage of abstracts per barrier type, except for the increase in legal framework. However, by dividing the ten year period into two year intervals and visualizing the percentage of abstracts for each barrier type illustrates much variability among the abstract classification over the time frame (Figure 8).

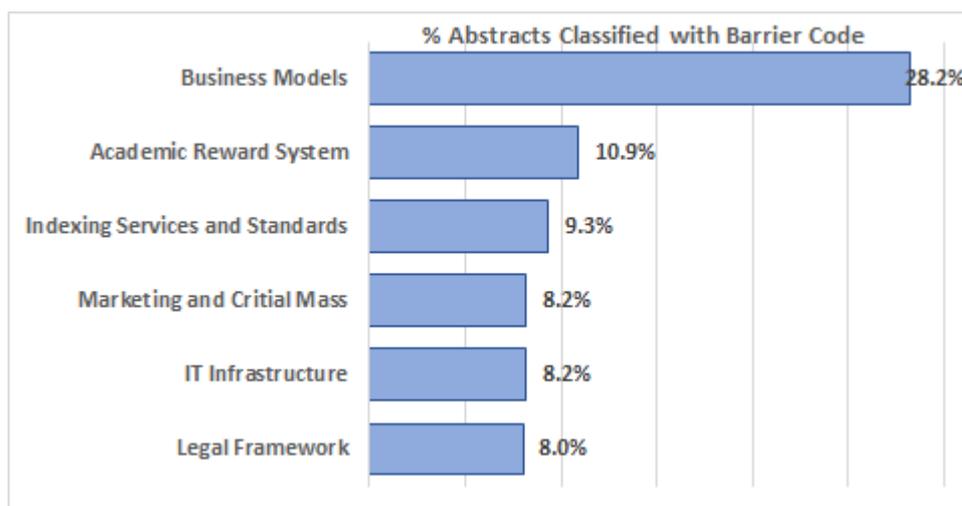


Figure 6. Percentage of “open access” abstracts coded for each barrier type over all the years combined (2004–2014) (*n* = 709).

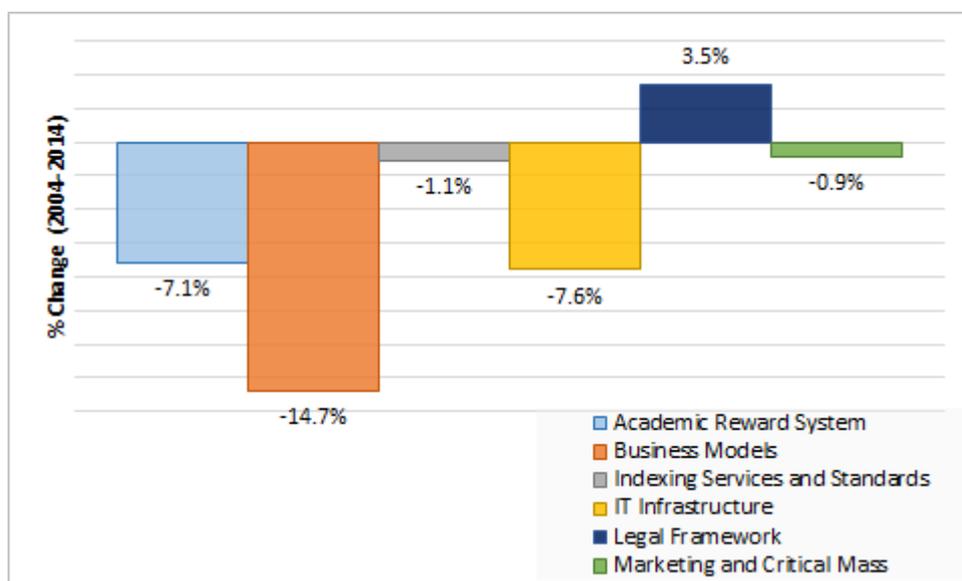


Figure 7. Percentage change for each barrier type over the ten year period, 2004–2014.

5. Discussion

Björk’s [7] discussion of barriers to open access details how each barrier can impact open access publishing parsed into broad publishing channels—journals that make content freely available (gold open access) and authors, or third parties, making their content available by depositing into repositories (green open access). Within this division, he also subjectively ranks how much the barrier impedes the development of open access over time based on anecdotal evidence and other research. This content analysis, however, is only considered through Björk’s lens of journal publishing channels or gold open access. The bibliometric data collected in this study is used to describe the data sample and also gauge how the LIS community compares to some of Björk’s barriers dialog. To begin, the majority of authors in this this study, those discussing “open access,” are LIS professionals either affiliated with a library or

maintain an information science position (see Figure 5). Their geographic affiliation is predominantly North American and European (see Figure 3).

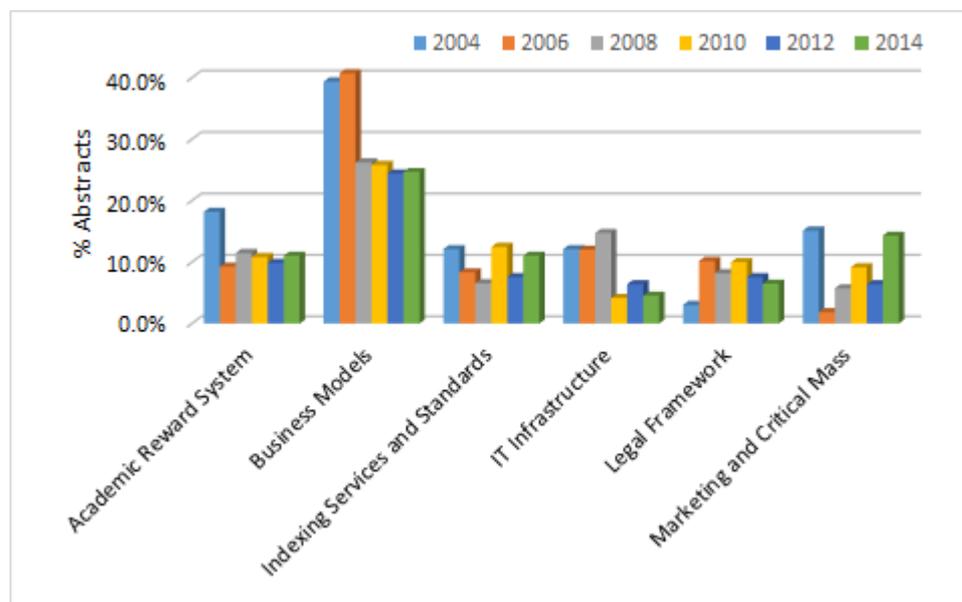


Figure 8. Percentage of “open access” abstracts for time period 2004–2014 by two year intervals for each barrier type.

In addition to describing the data, this study uses text mining to specifically explore the barrier types within the LIS literature. This analysis of the LIS literature assesses the appearance over time of the topics that depict the barrier within the discussion of open access; it does not evaluate the impact of that barrier on open access. In other words, a negative percentage value in an individual barrier from Figure 7 does not imply that topic is no longer a barrier to open access initiatives. Instead, it does indicate a decrease in the percentage of articles being published that contained subject matter related to the barrier type. This, however, could infer that the interest of the LIS community in that topic decreased.

Björk’s [6] study also included his interpretation of importance for each barrier by ranking how much a barrier might disrupt the rapid transition to open access. A mashup of the two datasets can be seen in Figure 9, where the bars represent the occurrence of a barrier as a topic in the LIS literature (LIS interest) and the stars denote Björk’s ranking system (three=high). This comparison shows similarities, for example in 2004⁴ the academic reward system, business models, and marketing are assigned the highest rank by Björk and concurrently show the highest percentage of articles (interest) for that year. It certainly stands to reason that if a topic is considered disruptive to an existing system, the professionals in the field would be discussing the topic. And following that reasoning, a barrier that no longer imposes constraints to open access would be less prevalent in the literature.

The discussion around IT infrastructure illustrates this supposition. By 2004, the technology for electronic publishing of scholarly literature was established and the subsequent development of new technologies only facilitated further publishing opportunities and initiatives [38]. Björk’s [7] assessment that IT infrastructure is no longer a barrier to gold open access is similar to the decline in the percentage

⁴ Note: Björk’s 2013 data is aligned with 2014 data

of IT-infrastructure related articles. While IT and infrastructure are still important to open access, there is little controversy around the need, which is generally an impetus for the intensification of a topic in the literature. The articles that do appear in 2014 report on specific software and technology integration by organizations, not necessarily dynamic debates.

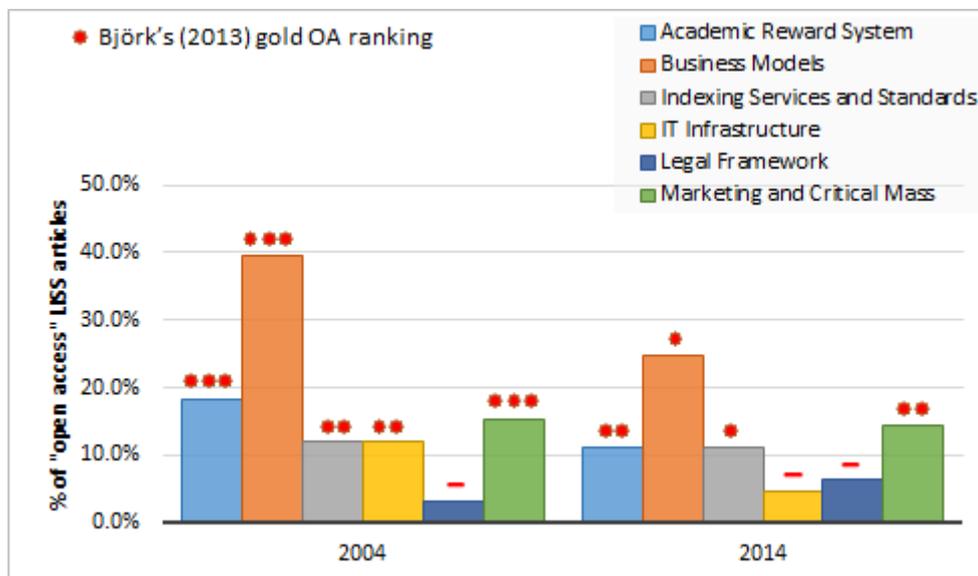


Figure 9. A mashup of Björk’s [7] ranking system for gold open access (open access journals) with content analysis data. Note: Björk’s data is aligned per his article publication dates.

This interesting parallel continues as the decrease in the topics surrounding the barriers (Figure 7) corresponds with Björk’s [7] conclusion that the barriers have indeed decreased in the past ten years. One disparity is that while there is an increase in the discussion of legal framework, Björk [7] argues that it had no change to the impact of open access over ten years (Figure 9).

Björk [6] assigns no rank to legal framework stating that the copyright agreements for open access journals do not hinder the development of open access; ten years later he does not alter the assessment. The content analysis data likewise indicates that legal framework issues are not prevalent in the literature in 2004 and although there is an increase in the percentage of literature published in 2014, it is still low compared to the other topics. Björk [7] pointed to the rising popularity of the Creative Commons licenses to further support his conclusion. While copyright exists to protect the rights of an owner of an original work by imposing restrictions on re-use, Creative Commons licenses “maximizes digital creativity, sharing, and innovation” by enabling a license holder to grant specific permission terms for using, modifying, and repurposing their work [39] (para 10). By facilitating sharing and re-use in an open access environment, Creative Commons’ licenses would certainly reduce the legal framework barrier to open access and accordingly, there has been massive uptake. In 2009, the estimate number of works with a Creative Commons licenses was 350 million [40]. However, this is a legal tool, not a law, and it is not always clear how to apply the licenses to specific situations and some argue it can be manipulated to clash with open access goals [41,42]. Therefore, as the LIS community endeavors to understand the issues, it follows that there would be an increase in the extent of articles about legal framework topics in the literature.

Björk's [6,7] description of the academic reward system as another barrier to open access points to the academic tenure system as a driver. He explains that the perception of open access journals lacking quality and citation impact effect an author's decision on where to publish for career promotion. He states that the situation is improving, for example open access journals now have traditional impact factors [7]. Recent studies have confirmed that there is no difference in the scientific impact of open access vs non-open access journals and that an open access article is more likely to be used and cited than one behind subscription paywalls [43,44]. In addition, surveys are showing that researchers do not believe publishing in open access journals would be considered a disadvantage by tenure and promotion committees [45]. Yet, in 2014, only 25% of all the LIS articles about open access were published in open access journals (Figure 2). Although this represents an overall increase over the ten years, it is still a small percentage considering the content of these articles includes some discourse regarding the unrestricted online access to scholarly information. It is reasonable assumption that as the prestige of open access journals increases, the academic reward system barrier would decrease. However, it is quite possible that author behavior is lagging behind attitude and the barrier is still present. Figure 2 clearly shows there has been little increase in the percentage of these articles being published in open access journals in the past five years.

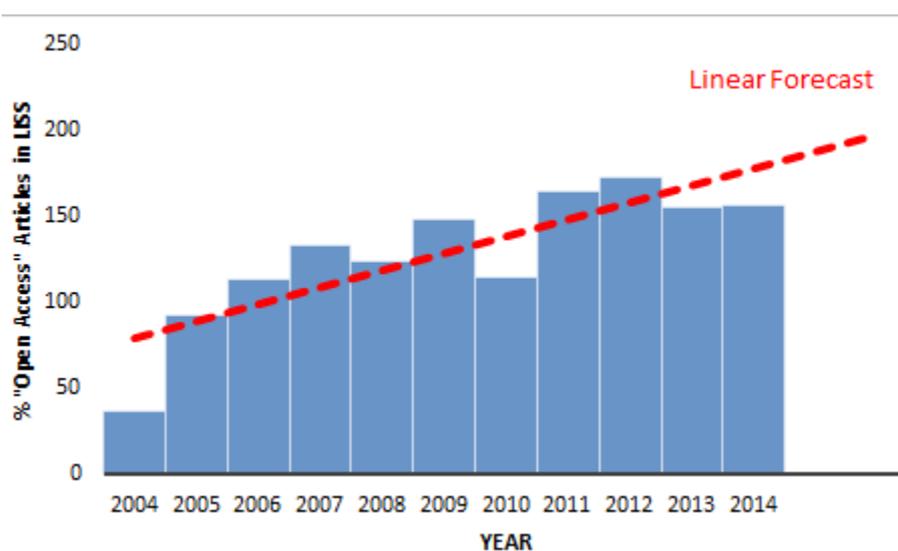


Figure 10. Growth and predictive growth of articles in LISS database with subject field (SU) “open access”.

Björk [6] discussed that marketing and branding are critical to the viability of scholarly journals as they are dependent upon getting authors to submit their best papers to remain in the market. He used the longitudinal growth of open access journals and articles to support his claim that the marketing and critical mass barrier to open access has decreased as more and more open access articles are published [7]. Singh *et al.* [17] reported that the growth of LIS journals in DOAJ increased from 3% in 2004 to 23% in 2014. Figure 10 demonstrates the growth of LIS articles specifically discussing open access. The volume of articles that contained the subject term “open access” tripled from 0.4% in 2004 to 1.2% in 2014 and the trend is to continue. This represents an increased discussion of “open access” by the LIS community via increased amount of articles about open access published. This does support Björk's

view, however, this is still a very low percentage of the total output from the LIS community. At the same time, the content discussion of open access journal marketing and critical mass did show a decrease, albeit very slight. As more published articles about open access continue to grow, the discussion about the volume of open access journals continues.

Business models is another barrier type that Björk [7] explains has decreased and although the content analysis data corresponds, mechanisms to keep an open access journal operating still remains an important topic in the LIS literature. Many open access business models have emerged and are becoming accepted by publishers, such as, author publishing charges [46]. However, until the situation stabilizes, the continued discussion and interest of the LIS community—as seen as a high percentage of articles about business models in 2014—is reasonable.

Björk [6] describes the extent to which a journal is indexed in commercial indexing services as the indexing services and standards barrier to open access. These services assist the visibility and access to journals and often lend prestige to a title [47]. Per Björk [7], after ten years the increase in open access journals appearing in newly developed indexes (e.g., DOAJ, Scopus) supports his claim that this barrier has decreased. The content analysis data shows only a slight decrease, however, implying that the topic has not decreased among the LIS community. In fact, while the availability of open access articles in commercial indexing services is still low, research is showing that the influence of the literature is increasing [48]. Instead of the discussion decreasing, it has shifted from quantifying the open access journals in commercial services to new ways of discovering open access content and new methods of measuring journal impact.

6. Final Remarks

This study adds to the dialog of barriers to gold open access by exploring the voice of the LIS community and illustrating changes in interest over time. As LIS professionals are major stakeholders in all things open access, this can represent the most prevalent views in that scholarship. The bibliometric data confirmed that this was an appropriate sample set and additionally verified the growth of the discussion of open access in the past ten years and beyond. This study additionally complemented Björk's results that the majority of the barriers to gold open access are lower today than ten years ago. Analyzing the dataset specifically for the factors that impede open access showed a correlation between what previous studies have quantified and what is considered a prevalent topic in the LIS literature, thus an important issue to open access.

Although it would result in a smaller sample size, further research might consider analyzing full text instead of article abstracts. The articles selected for this analysis were collected from multiple sources and this had an effect on the consistency of the sample. Some journals contained very structured abstracts while others only provided a single sentence or did not state the purpose and/or conclusion of the study. Other investigations have also shown that when using text mining methods, abstracts have different structural and content characteristics from article bodies even when the abstracts are similarly structured [49,50].

This study produced more data that can be further investigated to increase the understanding of the LIS dialogue around open access. While this project specifically compared 2004 to 2014 to represent the change over the ten years, the two year incremental data (Figure 8) shows much variation within the

time frame. Additional research into this temporal change could further shed light on what factors (e.g., political, cultural) are enmeshed in the prevalent barriers to open access as well as illuminate emerging conversations to identify new obstacles impeding an open access transition. Recognizing and studying the vital role of LIS in the open access discussion (*i.e.*, strategies and best practices) is critical to the continued growth and development of this scholarly communication.

Acknowledgments

The author would like to thank Dr. Carol Tenopir for her endless support and guidance and the opportunity to participate in Team Science at the School of Information Science, University of Tennessee-Knoxville.

Conflicts of Interest

The author declares no conflict of interest.

Appendix

Table A1. List of keywords with frequency >20. Includes the number of cases (abstracts) in which each keyword appears.

Keyword	FREQ	NO. CASES	Keyword	FREQ	NO. CASES	Keyword	FREQ	NO. CASES	Keyword	FREQ	NO. CASES	Keyword	FREQ	NO. CASES
ACCESS	1566	639	WORK	88	70	INITIATIVE	54	44	SOCIETY	41	31	PROBLEMS	31	25
OPEN	1476	635	INCLUDING	87	83	BENEFITS	52	43	HUMANITIES	40	30	PRODUCTION	31	20
RESEARCH	627	297	PEER	87	66	ISSUE	52	42	PRINT	40	35	SUPPLY	31	13
JOURNALS	588	241	WORLD	87	70	SEARCH	52	31	SIGNIFICANT	40	36	AWARENESS	30	25
OA	511	149	BUSINESS	85	55	SERVICE	52	31	VISIBILITY	40	30	CENT	30	10
PUBLISHING	487	270	FREE	85	68	COMMERCIAL	51	41	DISCIPLINES	39	30	COMMUNICATIONS	30	19
ARTICLE	429	345	DESIGN	84	82	STATUS	51	32	RECENT	39	36	DISSERTATIONS	30	16
INFORMATION	413	252	COMMUNITY	83	69	CHALLENGES	50	43	LAW	38	24	EDITORS	30	23
SCHOLARLY	341	213	PROVIDE	83	72	COSTS	50	40	SOFTWARE	38	28	LIBRARIAN	30	21
REPOSITORIES	289	123	SURVEY	83	45	FUNDED	50	43	TYPES	38	27	MEMBERS	30	25
JOURNAL	265	159	ARCHIVING	81	45	OFFERS	50	48	ACTIVITIES	37	31	PROGRAM	30	17
PAPER	265	159	RELATED	81	69	INDIA	49	24	AVAILABILITY	37	27	READERS	30	27
LIBRARY	250	168	INCLUDE	80	73	LEVEL	49	42	FOCUS	37	34	SCHOLAR	30	20
LIBRARIES	237	154	PROJECT	77	46	MAJOR	49	44	INTRODUCTION	37	36	ADVANTAGE	29	14
ARTICLES	229	112	SUPPORT	77	63	SYSTEMS	49	39	MAIN	37	35	CENTRAL	29	25
ACADEMIC	226	137	CURRENT	76	65	GROUP	48	38	OPPORTUNITIES	37	34	CITED	29	19
SCIENTIFIC	222	120	FUTURE	76	66	HEALTH	48	30	PRACTICES	37	29	CONSIDERED	29	26
STUDY	212	132	PRESENTED	76	74	DIRECTORY	47	35	BOOK	36	25	CONTEXT	29	27
UNIVERSITY	209	135	PRESENTS	76	75	FIELD	47	39	FREELY	36	31	EXAMINED	29	29
DIGITAL	201	121	UNIVERSITIES	76	55	FOCUSES	47	47	LEGAL	36	23	IMPORTANCE	29	27
INSTITUTIONAL	187	113	MAKE	74	62	FULL	47	35	PROVIDING	36	36	MEANS	29	28
PUBLISHERS	184	120	METHODOLOGY	74	70	GENERAL	47	40	SCIENTISTS	36	25	PARTICIPANTS	29	13

Table A1. Cont.

Keyword	FREQ	NO. CASES	Keyword	FREQ	NO. CASES	Keyword	FREQ	NO. CASES	Keyword	FREQ	NO. CASES	Keyword	FREQ	NO. CASES
DATA	180	89	TOPICS	74	70	HIGH	47	42	DOCUMENTS	35	20	TERMS	29	22
AUTHORS	178	120	ANALYSIS	73	50	COST	46	35	FACTOR	35	21	ARCHIVE	28	21
SCIENCE	168	110	CITATIONS	73	30	MATERIALS	46	29	PART	35	32	CHANGE	28	22
RESEARCHERS	162	96	SOCIAL	73	53	MEDICAL	46	31	PROJECTS	35	30	COLLECTED	28	23
MODEL	159	93	SYSTEM	73	59	SHARING	46	37	RESOURCE	35	32	COMPARED	28	21
IMPACT	158	102	REPORTS	71	69	THESES	46	21	RIGHTS	35	28	GOVERNMENT	28	19
REPOSITORY	151	84	INITIATIVES	70	53	YEARS	46	40	SPECIFIC	35	26	METHODS	28	27
PUBLICATION	149	99	TECHNOLOGY	70	57	ARCHIVES	45	37	ACCESSIBLE	34	28	TITLES	28	15
COMMUNICATION	146	108	PAPERS	69	36	BOOKS	45	34	EXISTING	34	31	TRENDS	28	22
LIBRARIANS	140	91	MANAGEMENT	68	51	DESCRIBES	45	44	GROWING	34	32	ACCESSIBILITY	27	20
BASED	139	115	FUNDING	67	47	DOAJ	45	26	IDENTIFIED	34	26	BARRIERS	27	24
DISCUSSES	138	129	SCHOLARS	67	47	EDUCATION	45	37	INCLUDED	34	24	CREATION	27	20
LITERATURE	138	80	DISCUSSED	66	62	INCREASE	45	37	PRESERVATION	34	24	GREY	27	8
WEB	137	84	INTERNET	66	58	METADATA	45	24	PROVIDED	34	33	LACK	27	24
AUTHOR	134	93	ORIGINALITY	66	66	STUDENTS	45	32	STATE	34	29	NETWORK	27	24
PUBLISHED	132	99	POLICY	66	48	ENVIRONMENT	44	40	WAYS	34	32	PHYSICS	27	21
ELECTRONIC	131	95	ROLE	66	54	POLICIES	44	35	CASE	33	32	SELECTED	27	23
KNOWLEDGE	127	84	USERS	66	49	SUBSCRIPTION	44	38	COUNTRY	33	24	SET	27	22
ONLINE	127	84	COUNTRIES	65	39	DATABASES	43	21	FORM	33	30	SOURCES	27	20
CONTENT	126	80	QUALITY	65	50	FACTORS	43	29	HELD	33	30	USAGE	27	18
FINDINGS	125	105	COPYRIGHT	62	45	GLOBAL	43	34	INFRASTRUCTURE	33	26	USER	27	22
PUBLIC	119	81	NATIONAL	60	46	HIGHER	43	36	LIS	33	14	YEAR	27	21
FACULTY	118	46	TRADITIONAL	60	48	KEY	43	40	OUTPUT	33	25	ADOPTION	26	20
ISSUES	117	89	PUBLISHER	59	46	TECHNICAL	43	36	RATE	33	23	DIFFERENCES	26	21

Table A1. Cont.

Keyword	FREQ	NO. CASES	Keyword	FREQ	NO. CASES	Keyword	FREQ	NO. CASES	Keyword	FREQ	NO. CASES	Keyword	FREQ	NO. CASES
APPROACH	115	97	PROCESS	58	50	TOOLS	43	35	AREAS	32	26	INFLUENCE	26	21
PURPOSE	111	76	GOOGLE	57	31	DISCUSSION	42	35	COLLECTIONS	32	22	MATERIAL	26	21
INSTITUTIONS	108	76	IMPLICATIONS	57	44	DISSEMINATION	42	37	EXAMINES	32	30	PROFESSIONAL	26	19
DEVELOPMENT	107	91	IMPORTANT	57	52	REPORT	42	34	GROWTH	32	30	REGARD	26	19
MOVEMENT	106	80	PUBLISH	57	45	SCHOLARSHIP	42	35	INCREASING	32	29	SAMPLE	26	17
PUBLICATIONS	104	79	SUBJECT	57	41	UK	42	34	MAKING	32	28	AMERICAN	25	18
RESOURCES	103	75	DEVELOPING	56	38	AIMS	41	38	PRESS	32	21	GOLD	25	16
RESULTS	100	82	INTERNATIONAL	56	47	COLLECTION	41	26	STUDIES	32	32	CHINA	24	12
REVIEW	100	70	IR	56	18	DATABASE	41	27	TEXT	32	25	AGE	23	13
SERVICES	98	66	SOURCE	56	40	DOCUMENT	41	21	DEVELOP	31	28	INDIAN	23	9
CITATION	97	37	POTENTIAL	55	52	EUROPEAN	41	29	DEVELOPMENTS	31	22	RESPONDENTS	23	12
MODELS	94	71	TIME	55	46	INSTITUTION	41	32	ECONOMIC	31	25	SKILLS	23	10
SCIENCES	89	62	CONFERENCE	54	45	PRESENT	41	38	EDUCATIONAL	31	21	IRS	22	11
NUMBER	88	64	DEVELOPED	54	44	REVIEWED	41	39	INCREASED	31	29	ETDS	21	7

Table A2. WordStat extracted phrases selected to create code dictionary.

ACADEMIC_REWARD_SYSTEM	
ALTERNATIVES_TO_THE_IMPACT_FACTOR	INCREASED_IMPACT
CHOICE_OF_A_JOURNAL	JOURNAL_CITATION
CITATION_IMPACT	METRICS_FOR
CITATION_IMPACTS	METRICS_IN
CITATION_INDICATORS	PRESTIGE
DECIDING_WHERE_TO_PUBLISH	PUBLISHING_IN_OPEN_ACCESS_JOURNALS
FACTORS_THAT_MOTIVATE	RANKED_JOURNALS
GREATER_RESEARCH_IMPACT	RESEARCH_IMPACT
IMPACT_ADVANTAGE	SCIENTIFIC_PUBLISHING_AND_PEER_REVIEW
IMPACT_FACTOR	JOURNAL_CITATION
CHOICE_OF_A_JOURNAL	TENURE_AND_PROMOTION
IMPACT_FACTORS	
BUSINESS_MODELS	
ALTERNATIVE_MODELS	MODEL_OF_OPEN_ACCESS
APCS	MODELS_FOR_SCIENTIFIC
ARTICLE_PROCESSING_CHARGES	OA_BUSINESS
AUTHOR_CHARGES	OA_MARKET
AUTHOR_PAYS	OA_MODEL
BIG_DEAL	OA_MODELS
BIG_DEALS	OA_MOVEMENT
BUSINESS_MODEL	OPEN_ACCESS_MODEL
BUSINESS_MODELS	PAYS_MODEL
COSTS_OF_PUBLISHING	PUBLICATION_CHARGES
ECONOMIC_REALITIES	PUBLICATION_FEES
FINANCIAL_SUSTAINABILITY_OF	PUBLICATION_MODEL
FREE_OF_CHARGE	PUBLISHING_BUSINESS
FUNDING_AGENCIES	PUBLISHING_FEES
FUNDING_AGENCY	PUBLISHING_MODEL
GOLD_OA	PUBLISHING_MODELS
GOLD_OPEN_ACCESS	RESEARCH_AND_LIBRARY_FUNDING
GREEN_OA	RESEARCH_FUNDING
GREEN_ROAD	SIDED_MARKETS
HYBRID_JOURNALS	SUBSCRIPTION_COSTS
HYBRID_OPEN_ACCESS	SUBSCRIPTION_MODEL
JOURNAL_SUBSCRIPTIONS	SUSTAINABILITY_OF_OPEN_ACCESS
LIBRARY_BUDGETS	TRADITIONAL_SUBSCRIPTION
INDEXING_SERVICES_AND_STANDARDS	
CITATION_ADVANTAGE	INDEXING_SERVICES
DISCOVERABILITY_OF	QUALITY_ASSURANCE
FINDABILITY	SEARCH_ENGINE
GOOGLE_SCHOLAR	SEARCH_ENGINES

Table A2. Cont.

IT_INFRASTRUCTURE	
ACCESS_TO_ELECTRONIC	INFORMATION_TECHNOLOGY
COMMUNICATION_TECHNOLOGY	OPEN_JOURNAL_SYSTEMS
EMERGING_TECHNOLOGIES	OPEN_SOURCE_SOFTWARE
INFORMATION_SYSTEM	PUBLISHING_INITIATIVES
INFORMATION_SYSTEMS	SOURCE_TECHNOLOGY
INFORMATION_TECHNOLOGIES	TECHNOLOGICAL_INNOVATIONS
LEGAL_FRAMEWORK	
AUTHOR_RIGHTS	COPYRIGHT_LAWS
COPYRIGHT_CONCERNS	COPYRIGHT_POLICIES
COPYRIGHT_ISSUES	CREATIVE_COMMONS
COPYRIGHT_LAW	INTELLECTUAL_PROPERTY_RIGHTS
MARKETING_AND_CRITICAL_MASS	
GROWTH_OF_OPEN_ACCESS	
MARKETING	
OPEN_ACCESS_INITIATIVE	

References

1. Harnad, S. Scholarly skywriting and the prepublication continuum of scientific inquiry. *Psychol. Sci.* **1990**, *1*, 342–343.
2. Read the Budapest Open Access Initiative. Available online: <http://www.budapestopenaccessinitiative.org/read> (accessed on 24 April 2015).
3. Pujar, S.M. Open access journals in library and information science: A study. *Annals of Library and Inf. Stud.* **2014**, *61*, 199–202.
4. Mercer, H. Almost halfway there: An analysis of the open access behaviors of academic librarians. *Coll. Res. Libr.* **2011**, *72*, 443–453.
5. Tomaszewski, R.; Poulin, S.; MacDonald, K.I. Publishing in discipline-specific open access journals: Opportunities and outreach for librarians. *J. Acad. Librariansh.* **2013**, *39*, 61–66.
6. Björk, B.-C. Open access to scientific publications: An analysis of the barriers to change. *Inf. Res.: Int. Electron. J.* **2004**, *9*, 170.
7. Björk, B.-C. Open access—Are the barriers to change receding? *Publications* **2013**, *1*, 5–15.
8. Borgman, C.L.; Furner, J. Scholarly communication and bibliometrics. *Annu. Rev. Inf. Sci. Technol.* **2002**, *36*, 1550–8382.
9. Miguel, S.; Chinchilla-Rodriguez, Z.; de Moya-Anegón, F. Open access and Scopus: A new approach to scientific visibility from the standpoint of access. *J. Am. Soc. Inf. Sci. Technol.* **2011**, *62*, 1130–1145.
10. Krippendorff, K. *Content Analysis: An Introduction to Its Methodology*; Sage: Newbury Park, CA, USA, 1980.
11. Blessinger, K.; Frasier, M. Analysis of a decade in library literature: 1994–2004. *Coll. Res. Libr.* **2007**, *68*, 155–170.

12. Atkins, S.E. Subject trends in library and information science research, 1975–1984. *Libr. Trends* **1988**, *36*, 633–658.
13. Buttlar, L. Analyzing the library periodical literature: Content and authorship. *Coll. Res. Libr.* **1991**, *52*, 39–53.
14. Hsieh, H.-F.; Shannon, S. Three approaches to qualitative content analysis. *Qual. Health Res.* **2005**, *15*, 1277–1288.
15. Potter, W.J.; Levine-Donnerstein, D. Rethinking validity and reliability in content analysis. *J. Appl. Commun. Res.* **1999**, *27*, 258–284.
16. Way, D. The open access availability of library and information science literature. *Coll. Res. Libr.* **2010**, *71*, 302–309.
17. Singh, J.; Shah, T.A.; Gul, S. Growth and visibility of LIS journals: An analytical study. *Ann. Libr. Inf. Stud.* **2014**, *61*, 193–198.
18. Mukherjee, B. Open access scholarly publishing in library and information science: *Ann. Libr. Inf. Stud.* **2008**, *55*, 212–223.
19. Mukherjee, B. Scholarly research in LIS open access electronic journals: A bibliometric study. *Scientometrics* **2009**, *80*, 169–196.
20. Rufai, R.; Gul, S.; Shah, T. Open access journals in library and information science: The story so far. *Trends Inf. Manag.* **2011**, *7*, 218–228.
21. Parveen, K. A bibliometric study on open access journals in library science discipline in DOAJ. *Int. J. Inf. Libr. Soc.* **2013**, *2*, 21–29.
22. Thavamani, K. Directory of Open Access Journals: A bibliometric study of library and information science. *Collab. Librariansh.* **2014**, *5*.
23. Singh, N.; Chikate, A. Open access LIS periodicals and digital archives: An evaluation with reference to Asian countries. *Electron. Libr.* **2014**, *32*, 710–725.
24. Yuan, S.; Hua, W. Scholarly impact measurements of LIS open access journals: Based on citations and links. *The Electronic Library* **2011**, *29*, 682–697.
25. Liu, Z.; Wan, G. Scholarly journal articles on open access in LIS literature: A Content analysis. *Chinese Librarianship* **2007**, *23*.
26. Grandbois, J.; Beheshti, J. A bibliometric study of scholarly articles published by library and information science authors about open access. *Inf. Res.* **2014**, *19*, paper 648.
27. Pagell, R. Curb your expectations. *Online Search.* **2013**, *37*, 52–58.
28. EBSCO and H.W. Wilson Source Databases. Available online: <http://www.ebscohost.com/superdatabases> (accessed on 24 April 2015).
29. Mittal, R. Library and information science research trends in India. *Annals of Library and Information Studies* **2011**, *58*, 319–325.
30. What Field Codes Are Available When Searching EBSCO Discovery Service (EDS)? Available online: http://support.ebsco.com/knowledge_base/detail.php?id=3198 (accessed 24 April 2015).
31. What Is the Difference between Subject Facets and Subject: Thesaurus Terms Facets? Available online: http://support.ebscohost.com/knowledge_base/detail.php?topic=996&id=4045 (accessed on 24 April 2015).
32. Golub, K. *Subject Access to Information: An Interdisciplinary Approach*; ABC-CLIO: Oxford, UK, 2014.

33. Mai, J.-E. Semiotics and indexing: an analysis of the subject indexing process. *J. Doc.* **2001**, *57*, 591–623.
34. Braam, R.R.; Moed, H.F.; Raan, A. Mapping of science by combined co-citation and word analysis. *J. Am. Soc. Inf. Sci.* **1991**, *42*, 233–266.
35. Hemantha Kumar, G.H.; Srinivasa, V.; Bhaskara Reddy, M.; Chandra, B.T. India's Contribution to Agriculture and Food Sciences through Open Access Literature. *DESIDOC J. Libr. Inf. Technol.* **2012**, *32*, 53–58.
36. Sahu, S.K.; Arya, S.K. Open access practices in India. *Libr. Hi Tech News* **2013**, *30*, 6–12.
37. Laakso, M.; Björk, B.-C. Anatomy of open access publishing: A study of longitudinal development and internal structure. *BMC Med.* **2012**, *10*, 124.
38. Yiotis, K. The open access initiative: A new paradigm for scholarly communications. *Inf. Technol. Libr.* **2005**, *24*, 157–162.
39. About the Licenses. Available online: <http://creativecommons.org/about> (accessed on 24 April 2015).
40. History. Available online: <http://creativecommons.org/about/history> (accessed on 24 April 2015).
41. Carroll, M.W. Why full open access matters. *PLoS Biol.* **2011**, *9*, e1001210.
42. Gulley, N. Creative Commons: challenges and solutions for researchers; a publisher's perspective of copyright in an open access environment. *Insights: UKSG J.* **2013**, *26*, 168–173.
43. Björk, B.-C.; Solomon, D. Open access *versus* subscription journals: A comparison of scientific impact. *BMC Med.* **2012**, *10*, 73.
44. Van Noorden, R. Open access: The true cost of science publishing. *Nature* **2013**, *495*, 426–429.
45. Nariani, R.; Fernandez, L. Open access publishing: What authors want. *Coll. Res. Libr.* **2012**, *73*, 182–195.
46. Kozak, M.; Hartley, J. Publication fees for open access journals: Different disciplines-different methods. *J. Am. Soc. Inf. Sci. Technol.* **2013**, *64*, 2591–2594.
47. Getting your journal indexed: A SPARC guide. Available online: <http://www.sparc.arl.org/resources/papers-guides/journal-indexing> (accessed on 24 April 2015).
48. Cummings, J. Open access journal content found in commercial full-text aggregation databases and journal citation reports. *New Libr. World* **2013**, *114*, 166–178.
49. Cohen, K.B.; Johnson, H.L.; Verspoor, K.; Roeder, C.; Hunter, L.E. The structural and content aspects of abstracts *versus* bodies of full text journal articles are different. *BMC Bioinformatics* **2010**, *11*, 492–501.
50. Cohen, K.B.; Hunter, L.E. Chapter 16: Text mining for translational bioinformatics. *PLoS Comput. Biol.* **2013**, *9*, e1003044.