Journal of Advances and Scholarly Researches in Allied Education
Vol. 17, Issue No. 2, October-2020, ISSN 2230-7540

A Study on Content Analysis and Information Science Research in Online Libraries

Pradeep Kumar Singh[1]*, Dr. Mohan Lal Kaushal[2]

[1] Research Scholar, Shri Krishna University, Chhatarpur M.P.

[2] Associate Professor, Shri Krishna University, Chhatarpur M.P.

Abstract - The goal of this research is to analyze the papers published in the chosen journals, focusing on their content and bibliometric metrics. The study of academic books and the development of library cataloging software are two of Library and Information Science's most important facets. Scientometrics examines all facets of the study of scientific literature and its approaches, tactics, and models, D-Lib Magazine, on the other hand, is dedicated to digital innovation and the digital dissemination of library materials. The journal D-Lib is dedicated to digital libraries, whereas the periodical Scientometrics includes scientometrics, bibliometrics, and webometrics.

Keywords - content analysis, information science, research, online libraries, etc.

----------X----------

INTRODUCTION

While we know the diversity of journal publications, we do not have a figure on how much of the collection is classified as research in the Library and the Information Science (LIS). The published discipline literature, such as journals, publications, monographs, conferences etc. presents the basis of knowledge and reflects the history, trends, research standards and social communication systems of the subjects "Windsor and Windsor are the knowledge bank in any area" (1973). In the same vein, 'disciplinary concerns are nowhere more clearly represented in its literature than in its research.' Most library and information science studies, like many other disciplines, are published as newspaper articles. (1)

LIBRARY AND INFORMATION SCIENCE

Library and informationscience is an area that helps people to discover and organize important information sources. Library and ICS is an interdisciplinary topic which relates to library practices, views and management instruments, information technology, education and other sectors; collecting, organizing, maintaining and diffusing information sources and the information political economy. Dr. Melvil Dewey of Columbia University established the first school for library science in 1887. (2)

Content Analysis

Content Analysis is referred to as the scientific communication content research. The analysis of content has been born as quantitative methods D.Laswell Harold. De solaPool (1952) was stated in these terms by the parent creators of the method. The primary activity of content analysis was counting frequencies. The study of content is based on the meanings, contexts and purposes of the communications. (3) The word content analysis is 75, and since 1961 the Webster English Language Dictionary has been listing it. During World War II, the development of content analyzes was undertaken as a complete scientific technique when the US administration funded well a project to assess enemy propaganda under the guidance of Harold Lass. The resources for research as well as the methodological progress achieved on the issues investigated throughout the project provided a major contribution to the development of the technique in the analysis of the material. (4) Bernard Berelson published in 1952 the content analysis of communications research, announcing technology identification as a flexible instrument for scientists in social sciences and the media. Some scientists have been adopted to study history & politics.

Meaning of Content Analysis

“Analysis of the manifest and latent content of a body of communicated material (as a book or film) through a classification, tabulation, and evaluation of its key symbols and themes in order to ascertain its meaning and probable effect.” (5) Content analyzes are basically a systematic study of the occurrence in books, movies and other materials of word, sentence, ideas and so on.

CITATION ANALYSIS

One of the most common uses of citation analysis is to track out influential writers and journals for a specific purpose, such estimating how long a piece of literature will remain relevant. A citation analysis looks at how often a certain author is referenced in a certain publication and how those references are shown visually. It does this by following the trail of references laid forth in scholarly literature. Bibilometrics use a method called citation analysis. To better grasp the topic at hand, we use this vital research tool to dissect its structure and chart its future development. Document usefulness and author-document connections are evaluated. It is a method for determining the most influential periodicals in a field by tallying the number of citations in each article published in a set of referenced periodicals. After that, a slew of related research appeared in secondary sources such books, dissertations, original journals, and review publications. (6)

Basic approaches to analyzing of Content Analysis

Then, the analysis of focus group data consists of two fundamental approaches: ethnographic summary and content analysis. The content analysis method generates numerical data descriptions. The analysis of content involves the identification of extremely precise variables. The references may be short or extremely comprehensive and weighted. (7)

Ethnography includes reporting, study participant selection, recording observations, discussions, and maintaining journals, but Geertz thinks that the undertaking is not sufficiently defined by any of those methods or processes. He thinks that ethnography is characterized as "a complex attempt in the 'thick description' by the kind of intellectual efforts.

The reality is portrayed in such depth that the reader may experience the whole event as if he or she had been engaged in it. The data should be interpreted to preserve this wealth. This rich description, including group interactions, is one of the benefits of the interview technique for the focus group. The ethnographic summary and the approach to content analysis are no contradictory method of analysis. The combination of these methods gives the study a further strength. (8)

Journal

An offer of scholarship or academic journals produced by a company or an organization that collects the papers authored by and for researchers/academics in an area. Journals are specialized on certain research areas. Some periodicals are characterized as reviewed by peers.

Annals of Library & Information Studies

The journal Annals is chosen for the research of the library & information studies. Library Annals & Information Studies is one of the major libraries and information science publications every quarter. Published by NISCAIR, New Delhi, India and has a great deal of popularity worldwide. (9) It is the main source of information and communication. Annals of library and information studies are a publication of the Indian National Scientific documentary, published since 1954. (INSDOC) Library and information studies annals provide articles, documentation notes and research reviews on library, documentation and science, IT systems, services and products, IT users, bibliometrics, scientmetrics, science and information, education and training, and the associated fields. (10)

SIX STEPS OF CONTENT ANALYSIS

There are typically six stages to a content analysis, and they are as follows:

Design: A mental process in which one analyses their situation, their goals, and their limitations when it comes to direct observation.

Utilizing: The process of defining and, eventually, recognizing units of analysis in the data volume available.

Sampling: Since selecting representative samples is not a natural part of Content analysis, there is a need: Assuring that the frequently conditional hierarchy of chosen sample a unit is correct and correcting for statistical bias.

Coding: This can be done either by giving specific instructions to skilled human coders or by using preexisting computer code to duplicate a rudimentary idea of meaning. (11)

PURPOSE OF CONTENT ANALYSIS

"Content analysis" refers to the process of evaluating a communication based on the information it conveys. In a broad sense, content analysis may be seen as a technique wherein inferences and conclusions are drawn about the message based on the message itself. The field of content analysis is in between document analysis and in-depth observation. Definition: an observational technique wherein information is gathered via observation rather than through questioning participants. It solicits quotes from the communications that individuals have made. To describe the overt content of a communication in an objective, methodical, and quantifiable manner is the goal of content analysis, a method of study. (12) According to Holsti (1968), it is "any method for drawing conclusions based on the systematic and objective identification of specific features of signals." It's a technique used in academic study

that involves following a series of steps to draw reliable conclusions from written material.

OBJECTIVES OF THE STUDY

- To study about the intellectual structure of journals (the subject structure).
- To find the citation pattern in journals.

RESEARCH METHODOLOGY

The data collection for the research consists of 1644 full-text journal articles published in Scientometrics and D-Lib Magazine between 2011 and 2020. A journal article is defined as an article that appears in a journal, excluding editorials, letters to the editor, and letters to the editor's comments, reminiscences, feature reports, news items, columns, historical notes, book reviews, book lists, bibliographies, short communication, world flash, opinion, conference report, correction, and obituaries, among other things. The information for the research was gathered from various parts of the publications (title, abstract, keywords, content, references, authors name and affiliated institution and country). The program 'R' was utilized in this investigation. It is open source statistics analysis software that may be downloaded for free from the internet. The following are the methodologies for this research, organized by objective:

Methodology for Objective One (To study about the intellectual structure of journals (the subject structure))

Co-word analysis is a technique used by numerous academics for the study of content in a variety of disciplines. Co-word analysis helps to organize data in several analysis levels: as link and node networks; as interaction network distributions and as network change across time periods (Ding et al., 2001). It employs patterns of words and phrases in a body or in a text. Figure 1 illustrates the procedures used in this research for co-word analysis

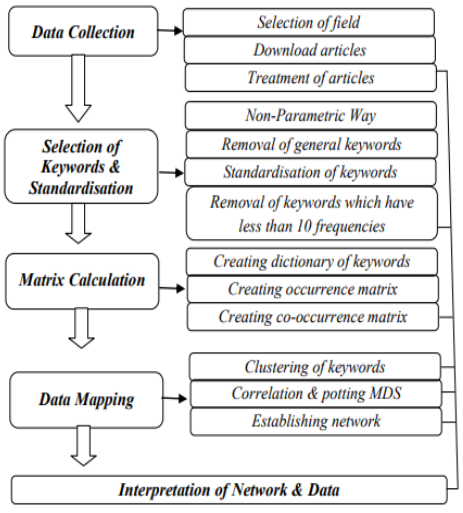

Figure 1: Steps of co-word analysis used for the study

Methodology for Objective Two (To find the citation pattern in journals)

Unlike Scientometrics, only 5442 items that have been considered have been included in the chosen articles of D-Lib Magazine. The aforementioned sample method for Scientometrics cannot be used to D-Lib magazines since fewer than two thousand references will emerge from this sampling strategy.

Data from references added to this goal have been gathered in the chosen research paper. The database was created using Open Office to capture the chosen item of each reference: the authorship(s), article title, article language, publishing nation type, publication country, volume, publishing date and issue, Article subject. A single identifying number was provided for each quotation. A second database with information on the sources utilized for cross-checking had been established.

Result analysis

All study results from the examination and analysis of collected data. It paints a vivid picture of the topic at hand and allows researchers to round out their findings with insightful commentary and recommendations. This chapter analyzes and evaluates the bibliometric features and content of the journals Scientometrics and D-Lib Magazine using quantitative data, bibliographical components, and keywords derived from articles published in both publications between 2011 and 2020.

INTELLECTUAL STRUCTURE (SUBJECT STRUCTURE)

Many researchers have been interested in their areas' subject structures (intellectual architecture) because of Thomas Kuhn's idea of scientific revolutions and paradigm shift. The development of a field of study through time and the impact of emerging subfields and topic specializations may be traced back to an analysis of the field's intellectual structure.

Scientometrics (Journal)

From a sample of 1241 papers published in Scientometrics journals, we extracted a total of 889 raw keywords. These terms were then standardized, resulting in a total of 243 distinct terms that were put through a co-word analysis. On average, each article contains 15.38 keyword occurrences. An article might include anywhere from one to fifty-one keywords. Fifty-two articles, or 4.19 percent, have more than thirty keywords, and two have as many as fifty-one. Articles with 21–30 keywords make up 257 (20.71%), those with 11–20

keywords make up 577 (46.49%), and those with 10 keywords or less account for 355 (28.61%). This means that more than 712 million articles (or 7139%) have more than 10 keywords. Co-occurrence analysis of 243 terms from 2011-2020 was performed to get insight into the journal Scientometrics's intellectual architecture. The years 2011–2015 and 2016–2020 have been separated from each other for the same reason; to track the subject's evolution over the course of the study period, which spans 2011–2020. Complete linkage cluster analysis was used to categorize the keywords for each time period into one of five groups.

Intellectual Structure based on Co-word Analysis (2011-2020)

Using MDS, we were able to create a high-level framework from which to analyze the situation of several Scientometrics subfields from 2011 to 2020. An MDS graph will be constructed using the specified clusters as input variables.

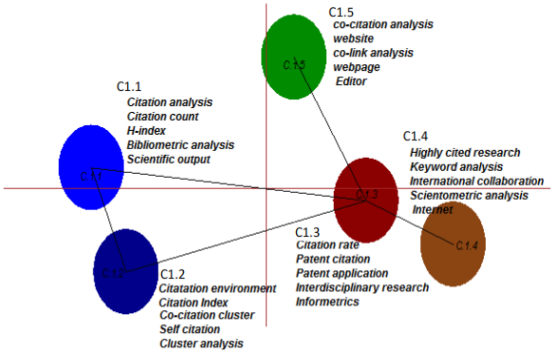

Figure 2: Scientometrics' 2011-2020 organizational frameworks

The five most common terms for each cluster are displayed in Figure. Where a cluster (sub-domain) appears in the graph in respect to other clusters shows how its keyword is related to keywords in other clusters. The figure's thin lines connecting clusters demonstrate a significant connection between them; this is justified by the fact that their Slaton indices are larger than 0.35.

Co-word Structure of Clusters

Using the correlation matrix of 245 keywords, a mapping of the sub-domain based on the cluster keywords as the input variable makes it possible to display the intricate co-word structure of each cluster. Scientometrics uses a five-cluster system, with each cluster's structure shown in detail to highlight a different set of characteristics (sub-domain). A thin line will connect the two terms in the diagram if the Salton Index is greater than 0.2. In the network of related words, the phrase with the scatological sign represents a link to an outside source. Among the 46 phrases found in the most prevalent cluster C1.1 are the terms "citation analysis," "scientific output," "university rating," and "case study." C1.2 is made up of 45 different concepts, including ones like "cluster analysis," "self-citation," "publication output," "citation pattern," "search strategy," and "statistical analysis."

Table 1: Cluster co-word structure characteristics from 2011-2020 (Scientometrics)

Characteristics	Clusters					Average
	C1.1 No.	C1.2 No.	C1.3 No.	C1.4No.	C1.5 No.	Value
1 Inner Link	52	40	102	60	70	64.80
2 Outer Link	39	39	40	13	23	41.20
3 Total Link	91	79	142	73	93	95.60
4 Inner Link %	57	51	72	82	75	67.40
5 Outer Link %	43	49	28	18	25	32.60
6 Inner link Key	24	19	34	29	29	27.00
7 Outer Link Key	14	15	17	7	14	13.40
8 Total Key	44	46	55	51	47	48.60
9 Inner link Key %	56	42.22	62.96	56.86	61.70	55.95
10 Outer Link key %	33	33.33	31.48	13.73	29.79	35.00
11 Average Link Per key	2.12	1.76	2.63	1.43	1.98	1.98
12 Density	0.261	0.301	0.284	0.285	0.273	-
13 Centrality	0.254	0.267	0.250	0.300	0.275	-

Table 1 shows that within clusters C1.1, C1.3, C1.4, and C1.5, the strength of the intra-connection is much greater than that of the outside connection. While C1.2's inner link is just marginally higher than its outer link. Sixty-seven percent of all links are internal linkages, which show the deep partnership between terms in clusters. In contrast to outside links, the number of keywords with internal ones is substantially larger. Approximately 56% of cluster terms have an inner link, whereas 35% have simply an exterior link. The average number of linkages per key is greater than 2 in C1.1 and C1.3, whereas it is fewer than 2 in the other clusters.

D-Lib Magazine

We analyzed 403 D-Lib Magazine articles and found a total of 1251 unique keywords. The previous standardization technique was used to these keywords, resulting in 412 unique keywords for use in the co-word analysis.

The Co-Word Analysis-Based Intellectual Structure (2011-2020)

Using methods developed for Scientometrics magazine, a high-level framework was developed to aid in comprehending the placement idea (relative position) of clusters keywords in D-Lib

Magazine from 2011-2020. Each cluster's (sub-domains') location in the graph represents the relationship of its term to keywords in other clusters. The five most-used terms for each group served as a symbol for that group. When the Slaton index is more than 0.35, a connection is established by thin lines linking clusters that have a strong association.

Co-word Structure of Clusters

Each cluster's features may be seen graphically through the displayed co-word structure (cluster). When the Salton Index value of a pair of keywords is more than 0.2, a thin line connects them. In a co-word structure network, a keyword marked with a cross denotes an outer link keyword (one that connects to other clusters).

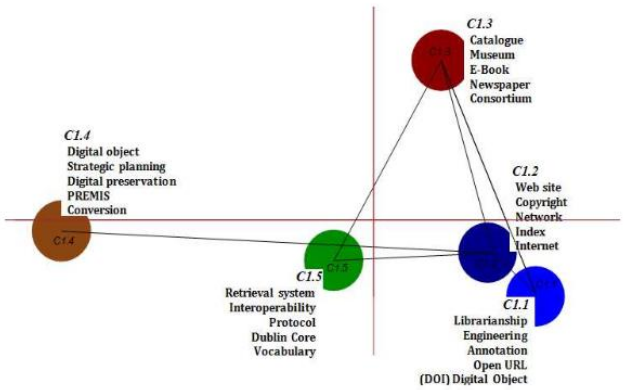

Figure 3: Overview of D-Lib Magazine's 2011-2020 Organizational Structure

Librarianship, open URLs, digital object identification, scientific publishing, electronic journals, innovations, electronic resources, the online, and academic communication are only some of the topics covered by the 62 terms in Custer C1.1. Website, copyright, network, index, Internet, citation, DSpace, journals, infrastructure, and search engine are just a few of the 117 distinct words found in Cluster C1.2. The 132 keywords found in cluster C1.3 were culled from the following diverse sources: a catalogue, museum, electronic book, newspaper, workshops, NSDL, visualization, digital collections, consortium, pictures, and OCLC. There are 24 words in Cluster 1.4, and they are mostly associated with things like digital objects, strategic planning, digital preservation, conversion, digital repositories, digital archives, and file formats. Cluster 1.5 covers 77 different topics, including search engines, metadata, APIs, vocabularies, licensing, automation, semantic analysis, MARC, and DCMI. Examining the network of detailed co-words more closely reveals that not all of the keywords have the same number of links or the same intensity of ties. There is a commonality between the concepts of digital education, digital divide, e-learning, e-entertainment, e-resources, e-journal, information literacy, information society, open URL framework, and print journals.

Table 2: Grouping characteristics based on co-word structure in the years 2011-2020 (D-Lib Magazine)

Clusters						
Characteristics	**C1.1No.**	**C1.2No.**	**C1.3No.**	**C1.4 No.**	**C1.5 No.**	**Average Value**
1 Inner Link	162	358	196	54	222	198.40
2 Outer Link	99	218	140	33	148	127.60
3 Total Link	261	576	336	87	370	326.00
4 Inner Link %	62	62	58	62	60	60.80
5 Outer Link %	38	38	42	38	40	39.20
6 Inner link Key	49	106	105	20	66	69.20
7 Outer Link Key	38	69	67	12	48	46.80
8 Total Key	62	117	132	24	77	82.40
9 Inner link Key %	79	91	80	83	86	83.80
10 Outer Link key %	61	59	51	50	62	56.60
11 Average Link Per key	4.21	4.92	2.55	3.62	4.81	4.022
12 Density	0.285	0.278	0.313	0.276	0.29	-
13 Centrality	0.239	0.238	0.238	0.232	0.239	-

Similarly, more internal links utilize terms in their anchor text than do exterior links. There is an average of 86.40 percent of keys within a cluster and 56.60 percent of keys outside of a cluster. The average number of linkages per key is greater than 4 in clusters C1.1, C1.3, and C1.5, and fewer than 4 in clusters C1.3 and C1.4.

CITATION PATTERN

Whether the source been previously published or not, it must be cited in order to ensure accuracy. Journal citation analysis looks at where experts got the information that shaped a study's authors' own work. The data gathered and organized under several table headers are discussed in this part, along with their analysis, which was driven by the study's secondary aims.

Scientometrics (Journal)

The total number of citations taken into account and analyzed for this report on the Scientometrics journal's citation pattern is 16627.

Types of Cited Resources

In this section, we present data on the sources that were mentioned. Print resources and electronic resources are the two main types of scholarly materials. According to Table 3, almost all citations to the journal Scientometrics originate from print-based sources, while just a small percentage come from online sources. Even in the era of electronic publication, it demonstrates the accessibility and reliability of print-based materials. According to Table, periodicals are the most common type of

resource. Only 744 (4.47%) of the 16627 mentioned works are from Proceedings, compared to 2405 (14.467%) from books and 11627 (60.93%) from journals.

Table 3: Scientometric citation styles

S. No.	Types of Resources	Number of Citations	Percentage of Citation
1	Print Resources	15673	96.93
2	E-Resources	954	3.07
	Total	16627	100

There is a widely held belief in the Scientometrics community that the majority of academic communication takes place through this medium, and the high number of journal citations in Scientometrics Journal supports this idea. Fewer than twelve percent of all citations fall into the other categories combined. There were only 954 citations (3.07%) that referred to an electronic resource, with just 453 (2.72%) referring to Web sites.

D-Lib Magazine

This section of the research looks at the bibliographic data for D-Lib Magazine articles produced between 2011 and 2020. References from 403 publications yielded a total of 5442 citations for this analysis. An examination of the various bibliographic components follows:

Types of Cited Resources

Data on the various sorts of references used are shown here. All of the listed materials have been divided into two main groups: print resources and electronic resources. As can be seen in Table 4, 60% of all citations originate from electronic sources whereas 39.60% come from print-based ones. This demonstrates how digital libraries are making greater use of their electronic resources.

Table 4: Types of cited resources of D-Lib Magazine

S. No.	Types of Resources	Number of Citations	Percentage of Citation
1	Print Resources	2155	39.60
2	E-Resources	3287	60.40
	Total	5442	100

The proportion and number of citations found in both print and online sources are shown in Table 4. It appears from the statistics that websites are the most frequently used resource. The total number of citations in the work is 5442, with 2312 originating from the internet (or 42.48 percent). Journals account for 1450 (28.30%), proceedings for 638 (11.72%), books for 474 (8.17%), and reports, etc. for 193 (3.55%) of the total citations in these instances.

CONCLUSION

The bibliometric study of papers published in both journals gives an overview of the basis, the transfer and the assessment of these papers. The quotation utilized by the writers in both subjects indicates if various types of information resources in the area are available and important. More and more authors in both journals have verified team-based research in these areas. During this time D-Lib Magazine is opposed by the always favorable growth rate for the Scientometrics magazine (2001-2010). English and US institutions and organizations' connections to writers have a dominant influence. However, the techniques used in this study are considered to be helpful in determining the intellectual framework (support structure) and the trend of the topic in any knowledge area for a broad variety of scholars, particularly scientists, researchers and bibliographic professionals.

REFERENCES

1. Agrahari, A, Chaudhary, C. P., & Singh, S. N. (2017) Worldwide growth of Library and Information Science journals, Journal of Social science & Humanities. 4(3), 164-184.

2. Singh, S. N., & Agrahari, A. (2015). Scholar's attitudes towards 'e' resources in School of Earth Science and Natural Resources and Management of Mizoram University: a survey, Asian Academic Research Journal of Social Sciences & Humanities, 2(3), 237-249.

3. Ravikumar, S., Agrahari, A., and Singh, S. N. (2015). Mapping the intellectual structure of scientometrics: A co-word analysis of the journal Scientometrics (2005–2010). Scientometrics, 102(1), 929-955.

4. Chu, H. (2015). Research methods in library and information science: A content analysis. Library & Information Science Research, 37(1), 36-41.

5. Misra, S., and Dutta, B. (2014). Annals of Library and Information Studies: an analysis of citation pattern. Vidyasagar University Journal of Library and Information Science, 18 (1), 61-78.

6. Levine-Clarc, M. & Carter, Toni M. (2014). ALA glaossary of Library and Information Science. New Delhi: DBS Imprints.

7. Garg, K. C. & Bebi (2014). A citation study of Annals of Library and Information Studies (ALIS) and DESIDOC Journal of

Library and Information Technology (DJLIT). Annals of Library and Information Studies, 61 (3), 212-216.

8. Bhat, P. S. (2014). Use of periodicals in Engineering College Libraries under North Maharashtra University, Jalgaon. International Journal of Library and Information Science, 6(3), 35- 39.

9. Zong, Q. J., Shen, H. Z., Yuan, Q. J., Hu, X. W., Hou, Z. P., & Deng, S. G. (2013). Doctoral dissertations of Library and Information Science in China: A co-word analysis. Scientometrics, 94(2), 781-799.

10. Singh, G. (2013). Information source, services and systems (p.8). Delhi: PHI learning.

11. Kim, H. (2013). An Analysis of the Intellectual Structure of the LIS Field: Using Journal Co- citation Analysis. Journal of the Korean BIBLIA Society for library and Information Science, 244, 99-113.

12. Hu, C. P., Hu, (2013). A co-word analysis of library and information science in China. Scientometrics, 97(2), 369-382.

Corresponding Author

Pradeep Kumar Singh*

Research Scholar, Shri Krishna University, Chhatarpur M.P.