

An Empirical Research of gray Literature Based on Citation Analysis: Taking Journal of Library Science in China as an Example

Fengmei Zhao

Tianjin University of Technology and Education Library, Tianjin

83579778@qq.com

Keywords: citation; literature; electronic documents

Abstract. The paper takes references of all papers published in Journal of Library Science in China between 2005 and 2014 as data source. The paper counted, induced and analyzed the exploitation and utilization of gray literature resources from the perspective of gray literature citations. The influence degree of gray literatures on academic research of scholars in the field of Library and information science were discussed. Proposals for the development and utilization of gray literature resources were proposed.

Introduction

As an important component of information resource, gray literatures often walk in the forefront of research and innovation, for they can provide frontier practical technology and dynamic decision-making, which are important to economic development and knowledge creation. With the development of web technology and database technology, carriers of gray information resources have gradually included printed, online and database categories et al. With the incomparable advantages of fast update frequency, wide range of coverage over printed literatures, gray literatures are playing an increasingly important role in scientific research fields, and becoming an important part of library information literature source for scientific and academic research current technology that researchers must pay attention to and master [1]. According to the discussion of Science Citation Index and corresponding information research organization reports, the number of gray literatures has accounted for about 20% of all literatures [2]. The joint survey of the American National Science Foundation and Japan National Statistics Bureau shows that more than 90% of new technology patent document are made public through gray literature in their organizations, 95% of products information come from gray literatures [3].

Quotation are authors' responses on problems based on the inheritance of others' research results, and reflect respect for others as well as a number of important research questions of their own thinking and understanding on the learning process at work [4]. References characterize the links between literatures and author's academic record behavioral characteristics [5]. The phenomenon of treating gray literatures as references has always existed in many domestic and foreign journals. Throughout the literature, the perspective of researchers are more concentrated in terms of theoretical discussion of gray literature collection, development, construction and management, so far, they has not yet seen the empirical studies on gray literature citations based through the lens of gray literatures usage. This paper will take all the articles published on the library and information science journal "China Library Journal" between 2005 and 2014 as the sample, counts, summarizes, analyzes and explores Library information science field scholars' analysis of gray literature usage and its impact on academic research. We hope the paper can contribute to the development and use of gray literature and provide some reference recommendations.

For convenience, we will hereinafter collectively referred journal articles as papers; References in papers are collectively referred as citations; gray literature citation are referred to as gray Citations or GL citations; "Chinese Library Journal" is referred to as "CLJ".

Research Design

Research Methods

Firstly, we use literature research method to collect, identify and organize information resources related to gray literature to understand the connotation, classification, development and utilization status of gray literatures; secondly, we will take "Journal of FIG" as a research case, taking bibliometrics methods to perform a statistical analysis of the gray citation published for nearly a decade to describe and explain the data characteristics of the gray literature scholar research activities, variation principles and application status.

Sample Selection

In this study, we take all 60 volumes papers published on "CLJ Journal" between 2005 and 2014 as study sample, no sampling (excluding dynamic information, submission guidelines, author guidelines, Call for Papers, inspiration, congratulations, congratulatory, etc.). The following are the main considerations for the sample selection: ① "CLJ" is sponsored by China Library Association and National Library, charged by the Ministry of Culture People's Republic of China, many papers published on the journal reflect the highest level of Chinese Library Science [6]; ②CLJ has always been at the subject development forefront, and disseminated latest achievements in the field of Library and Information Science timely. CLJ has become the most influential library science professional journals in China [7],For over 50 years, "CLC" is regarded by scholars as the first Chinese library community field journal for its higher impact factor and the combined effects of factors[8];③Based on 18 kinds of CSSCI source journals among library, information and Documentation Science Researcher, Liu and other researchers took reputation, achievement and overall level as measure variables of journal academic status, studies found that China Library Journal located on the top of scores tower is in the first place of academic status among all Library and Information academia journals [9].

Data Collection

In the pursuit of accuracy of the data, we took papers published on CLJ with clear classification and scientific regulations as statistical basis, classified by gray citations of the papers. For citations and gray citations in a paper, when they come from the same sources, we count the citation numbers according to actual citation times.(For example:1,2,3 International Organization Standardization.ISO 14721-2003 spatial data and information transfer systems - Open archival information system - Reference model [S].2003.32-60, there are three formal references, when analyzing we still counted as three citations), and finally, we use bibliometrics methods and typed all data into Excel spreadsheet for statistical analysis.

Data Analysis

In the existing research, scholars often borrow the following three indicators to measure the impact and use of gray literature [10]:① use of frequency gray citations means the proportion of papers containing gray citations in all papers total, it only considers whether a paper's references contain gray citations, regardless of the number of gray citations it contains;②gray citation reference frequency, which means the proportion of gray citations in all cited papers, it need consider the number of citations in gray citations;③gray citation intensity of use that gray citation reference frequency divided by the frequency of use of the results, it is a comprehensive index to reflect the utilization of gray citations. This paper considers the above three indicators, while also analyzing the frequency distribution of gray citations, types and distribution source distribution.

Data Analysis Result

Use frequency of Gray Literatures

Authors reflect the vitality of a magazine, because they are implementer of a magazine's function and the provider of a magazine's dynamic sources. The consciousness level and abilities of authors directly determine the journal quality levels, positioning of periodicals and development trend [4]. Gray literatures use frequency analysis can directly reflect the authors' use awareness level of gray

literatures. Such as height, directly determines the level of the quality of publications, periodicals Position and Development Trend. As shown in Table 1, P represents the total journals published paper amount, G indicates the amount of the papers containing gray citations, G/P indicates gray citation use frequency. During 2005-2014, gray literatures use frequency in CLJ substantially follows the trend of increasing year by year, from 47.5% in 2005 to 92.8% in 2014, an increase of nearly 2 fold. Overall, among 1075 papers in the last decade, there are 740 published articles quoted gray literatures, which accounted for 68.8%. Thus, the authors of CLJ use awareness of gray information resources is relatively high, and the use behavior of gray literature also increased every year. Especially in 2012, the use of gray literature reached its peak, gray citation use frequency was as high as 93.5%, ie 100 Papers, nearly 94 papers cited gray literature.

Table1 the amount of paper and papers containing gray literatures annually in CLJ

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	sum
P	137	150	136	128	113	94	85	77	86	69	1075
G	65	68	83	92	81	77	71	72	67	64	740
G/P(%)	47.5	45.3	61.0	71.9	71.7	81.9	83.5	93.5	77.9	92.8	68.8

Gray Literature Citation Frequency

For statistical purposes, we use L to represent citations number of journal papers, GL representing gray citations and GL/ L indicating gray literature citation frequency. "CLC Journal" Citation amount of CLJ papers each year between 2005 and 2014, the amount of gray literatures and gray citation frequency are shown in Table 2. From 2005 to 2014, gray literature reference frequency is also substantially exhibit a rising trend. Among them, in 2013 the ratio reached the maximum of 42.2%, which is consistent with the gray citation frequency changes growth in Table 1. The underlying reasons may be while traditional conventional literature is still the main front of academic exchanges in network environment, but gray literature including online electronic documents has their strong advantages and get more and more popular by scholars, just as the sociologist of University of Chicago called Evans's study findings published in the journal Science: as more and more journals available online, papers cited more recent published articles, while fewer journals and articles are quoted [5]. This reflects from the side, with the rapid development of gray information resources at home and abroad, as well as the rapid development of computer and network technology, scholars 's attention to and use of gray literature is increasing year by year. It also shows that researchers are able to catch up with the pace of The Time and technological development, and can gradually use of cutting-edge technology to arm themselves, enriching their own scientific fields.

Table2 Citation amounts of citations and gray citations in CLJ annually

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	sum
L	1366	1516	1809	1832	1590	1958	2058	1917	2042	2299	18387
GL	326	261	473	508	498	590	573	775	862	713	5579
GL/L(%)	23.9	17.2	26.2	27.7	31.3	30.1	27.8	40.4	42.2	31.0	30.3

Analysis of Gray Literature Citation Frequency

Analysis of gray literature citation frequency=Papers amount with gray literatures in this journal/papers amount, its value may indicate the authors' utilization status of gray information resources to some extent. P represents Papers in journal, GL represents the amount of total gray citations, gray citation rate = GL/ P. Data in table 3 show that gray literature citation rate between 2005 and 2014 has increased significantly. From 2.4% in 2005 to 10.3% in 2014, the average gray literature citation rate is 5.2%, which means that each paper cited 5.2 gray literatures. One foreign study showed that the output number of a specific science project is formed "pyramid" shape according to document types, which means that there are two journal papers, 12 conference papers and 20 technical reports, although this data can't be regarded as final statistical results, but it can be seen gray literature have great impact on scientific Research of the United States[10]. On the other hand in FIG Journal" scientific practices in JCSL, authors' awareness about gray literature production and exploitation need improvement, gray

literature production, transmission, exchange and management system need to be further improved. Otherwise, as time goes by, this situation will not help sharing and exchange of academic knowledge, which isn't conducive to the transfer of scientific research achievements into productive forces and hinder the technological innovation and economic development of the country.

Table3 the amount of papers and gray literatures citations in CLJ

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	total
P	137	150	136	128	113	94	85	77	86	69	1075
GL	326	261	473	508	498	590	573	775	862	713	5579
GL/P(%)	2.4	1.7	3.5	4.0	4.4	6.3	6.7	10.1	10.0	10.3	5.2

Frequency Distribution of Gray Literature

Gray literature Citation frequency is the number of citations of each gray paper, which reflects the authors' preference level of gray literature use. For example, "the frequency is 1, the papers including gray literature Citation is 69" means the number of gray literature citation is 1 in 69 papers. The statistics found that in CLJ, the gray citation frequency is from 0 to 58. For statistical convenience, this paper takes frequency 5 as increments unit, counting the gray citations of papers according to the frequency range of statistics including the grey number of citations. The frequency statistics of gray literature citation of CLJ between 2005 and 2014 are shown in Table 4, among 1750 papers, there are 335 articles whose authors do not cite gray Literature Papers. Overall, the distribution of gray literature citation in the papers showing a clear trend of concentrated dispersion, that very few authors cited a lot of gray citations to support his research point of view in their writings, while a few gray citations are scattered in a lot of literatures. We found in the statistics, there are several papers whose citations are all gray literature, such as the paper "Network Information Resource Preservation Situation and Development Trends" in issue2 of 2009, which had 38 citations which were all gray literatures. The reason behind it may be that different scholars have different research areas, or they have different gray information resources utilization and recognition levels.

Table4 Frequency distribution of gray literature citation

Frequency range	Papers containing gray literature
0	335
1-5	420
6-10	156
11-15	71
16-20	39
21-25	20
26-30	13
31-35	5
36-40	4
41-45	6
46-50	3
51-55	2
56-60	1
total	1075

The Source of The Electronic Literature

As we all know, the internet has become the major scientific communication media of researchers, more and more scholars access information, discovery and interact through the internet, which greatly contributed to the emergence and development of electronic literatures. For traditional printed literatures, the citation sources refer to the journals or works containing the citations. For the source of the electronic literature, the source of the citation can be understood as the source address of the citation and the type of the source. By analyzing the sources of electronic literature, we can find various types of influence degree of online resource on scholars various types of organizations provided.

According to statistics derived from Table 5, the general online literatures from various organizations website are the most used frequently and the proportion was as high as 90.8%. From 2008, CLJ began to clearly classify electronic version of conventional literature in bibliographic citations, it can be seen from the comparison, these networked traditional literature began to play a more important role in researchers' research activities.

Table5 the distribution of electronic literature in CLJ

	General online literature			electronic edition of traditional document		
	Chinese	foreign language	sum	Chinese	foreign language	sum
2005	84	227	311			
2006	97	140	237			
2007	204	251	455			
2008	135	237	372	1	40	40
2009	123	183	306	3	32	38
2010	129	266	395	4	43	47
2011	72	229	301	12	40	58
2012	78	373	451	5	63	68
2013	81	429	510	14	56	70
2014	115	330	445	6	58	62
sum	1118	2665	3783 (90.8%)			383 (9.2%)

In Table6, for the electronic literature of papers in CLJ nearly ten years, authors divided them into seven categories according to domain name type they belong to. As can be seen from the statistics, in 4166 electronic literature citations, there are 1241 citations from the site of ORG domain name, with the highest proportion up to 29.8%; secondly, citations from the domain name of COM and the domain site name of EDU accounting for 20.1% and 13.2% respectively. While citation that Scholars referenced from the three domain name GOV, AC and NET three sites referenced is relatively less. This is similar to the results of other researchers [11]. From this analysis, the online information resources provided by business organizations, education departments and some organizations are more likely to be used by scholars in the field of Library and information science, and has a great impact on academic exchange activities. As the results of other researchers[12][13], with the birth of professional academic search engine and the third party integration tools online electronic resources are getting more and more are widely used by researchers.

Table6 Electronic documents belonging to a domain name in each of the type and quantity in CLJ

domain name	GL citations	the proportion(%)
ORG=organization	1241	29.8
COM=commercial	837	20.1
EDU=education	550	13.2
GOV=government	486	11.7
AC=academic	205	4.9
NET=network	169	4.1
other	678	16.3
sum	4166	100

Conclusions

Through the analysis of nearly ten years of the gray citation of papers in CLJ, it can be seen that scholars from Library and information science scholars' capacity of use gray information resources is improving year by year, especially for the use of electronic document is to achieve skilled level. But

use of conference papers and dissertations and other resources various colleges and universities and vigorously collected and constructed was not ideal, there is a phenomenon of attention to the construction, but contempt application, so that they should not act as an important role in academic exchange activities. So the author believed that the effective collection, selection and integration, dissemination and utilization of gray information resources should still be a research problem that we will continue to pay attention to in the future. In addition, this article only selected only one study sample from CSLJ, the conclusion is inevitable that there will be "Benevolence" is suspected, but given the number of statistical data, which we can still see a little clue; at the same time due to the impact of human factors, there may be some small deviations in statistics process.

Acknowledgements

This work was financially supported by the Scientific Research Development Fund Projects of Tianjin University of Technology and Education (sk13-02) and Young Foundation of Ministry of education, humanities and social science research projects (13YJC880029).

Reference

- [1] Lihua Xu. Research on the Intelligence Value and the Development and Use of Gray Literature[J]. Information Science, 2006(10):1486-1488. (in Chinese)
- [2] Xueli Ma. The Analysis and Definition for Connotation and Extension about Gray Literature-The Third Series about Studying of the Grade and Structure of Document [J].Library Development,2003(1):13-15. (in Chinese)
- [3] Lei Wang. Practice of the Systematic Construction of Gray Literatures [J].Library Development, 2012(12):25-28. (in Chinese)
- [4] Fengmei Zhao. Research on Web Citation Analysis of E-Education Research [J].E-Education Research, 2006(10):24-26+32. (in Chinese)
- [5] Siluo Yang. The Status and Trends of Web Citation Study Abroad[J].Journal of Library Science in China, 2010(4):72-82. (in Chinese)
- [6] Information on http://www.jlis.cn/jtlsc/ch/first_menu.aspx?parent_id=20080416210941001
- [7] Heping Zhou. The 50th anniversary Congratulatory Message of Journal of Library Science in China [J]. Journal of Library Science in China, 2007(5):5. (in Chinese)
- [8] Liqiang Liu, Fei Li. The 2007-2011 Statistical Review of China Library Journal [J].Journal of Modern Information, 2013(5):89-91. (in Chinese)
- [9] Yu Liu, Jiyuan Ye, Xilin Yuan. Hierarchical Structure of Library and Information Science Journals: An Empirical Study Based on Peer Review [J].Journal of Library Science in China, 2011(2):105-114. (in Chinese)
- [10] Min Chao, Jianjun Sun. Distribution of Chinese Electronic Gray Literature: A Sample Analysis Based on Google Scholar [J].Library and Information Service, 2014(24):73-79. (in Chinese)
- [11] Jie Xing, Dong Wei. The Analysis on the Domestic Research of E-government Based on the Keywords Statistics [J].Journal of Intelligence, 2012(1):115-118+122.(in Chinese)
- [12] Xinrong Li. Analysis of Difference between University Institutional Repositories and the Commercial Database on the Content of the Construction of the Gray Literature [J]. Library Theory and Practice, 2011(10):60-62. (in Chinese)
- [13]Yong Hu. The Study of Relationship Between Social Telepresence and Different Network Study Effect in the Process of Online Learning [J].E-Education Research, 2013(2):47-51.(in Chinese)