

国内开放存取的研究热点：基于共词分析的文献计量研究

李 武 董 伟

摘 要 以 CNKI 期刊全文数据库作为数据来源,在对关键词进行一定程度人工干预的基础上,采用共词分析法探索国内开放存取的研究热点。研究认为,2003—2009 年间国内学者对开放存取的研究热点集中在七个方面:OA 期刊出版、自存储和 OA 知识库、OA 期刊的质量评价、OA 资源的长期保存、开放存取的发展对策、开放存取对图书馆的影响、开放存取与学术传播的关系。目前的研究仍停留在形态研究层面,未来应在受众分析和效果研究两个层面加强力度,针对国内的科研人员和学术传播体系开展扎实的本土化研究工作。图 1。表 4。参考文献 67。

关键词 开放存取 研究热点 共词分析

分类号 G230

ABSTRACT Based on co-words analysis of key words of papers on Open Access published from 2003 to 2009, the authors conclude that there are mainly seven hotspots of OA research in China: OA journals publishing, self-archiving, OA journals evaluation, OA resources long-term preservation, OA developing strategies, the influence of OA on libraries, relationship between OA and scholarly communication system. The authors point out that user analysis and effect analysis should be enhanced and localization should also be conducted. 1 fig. 4 tabs. 67refs.

KEY WORDS Open access. Research hotspots. Co-words analysis.

CLASS NUMBER G230

1 引言

开放存取(Open Access)作为一种新的出版模式和学术传播模式,引起了国际社会越来越广泛的关注。开放存取不仅是一种机构的战略,更应该是国家的战略^[1]。自从该概念引入国内学术界后,许多研究者对这种全新的学术传播模式投入了极大的研究热情。如王云才、陈红勤和刘锦宏等人分别对开放存取在不同时段的研究状况进行梳理^[2-4]。这些文章对于我们了解开放存取的研究现状具有较大的参考价值,但这些研究均采用简单的频次统计法来考察作者、作者机构、来源期刊和研究主题的分布情况,在考察开放存取的研究现状,尤其是对于

研究热点的揭示方面存在一定的局限性。

本研究采用共词分析法,试图克服以往研究中存在的不足,更加完整、客观地反映开放存取在我国的研究热点情况。共词分析法属于内容分析方法的一种,其原理是对一组词两两统计它们在同一篇文献中出现的次数,以此为基础对这些词进行聚类分析,从而反映出这些词之间的亲属关系,进而分析这些词所代表的学科和主题的结构变化^[5]。就目前文献来看,共词分析法主要是通过共关键词或主题词的方法,分析鉴别某一学科或主题的主要知识结构和研究热点^[6]。与共被引分析法相比,共词分析法是对当前发表文献的直接统计,所寻找的是目前已有论文集中关注的主题,反映的是在趋势形成之后的焦点,适合寻找新兴学科的范

式或共同体;而共被引分析法则是通过分析以往发表的论文的被引用情况来探究人们目前关注的焦点,更适合于寻找成熟学科的范式^[7]。这也正是本研究为什么没有采用共被引分析而采用共词分析的原因。

2 研究材料和方法

本研究选定 CNKI 期刊全文数据库作为获取数据来源。具体来说,在 CNKI 期刊全文数据库中选取 2010 年 1 月 2 日以前的检索主题为“开放存取”或“开放获取”的相关文献(精确匹配),将时段限定为 2003 年到 2009 年,最后获得 856 条记录。通过删除与主题不符的文章和未提供作者关键词的文章,最后保留 728 条记录。然后采用共词分析法,分三个阶段,进行分析。

第一阶段:统计高频词和制作共词矩阵。主要步骤如下:①关键词的预处理。由于关键词未受词汇控制,本研究在进行词频统计之前进行了一定程度的人工干预。如将 open access 的中文译名统一为“开放存取”,将 self-archive 的中文译名统一为“自存储”等;②计算关键词的出现频次,并选择频次不小于 5 的 46 个关键词作为代表学科主题研究方向的高频词;③两两统计这些高频词在同一篇文章中同时出现的次数,形成共词矩阵。

第二阶段:基于共词矩阵进行聚类分析,同时考虑到聚类分析在面对聚类谱系图难以确定最佳分类数的局限性,本研究参考因子分析所确定的因子个数寻找聚类分析的分类点来解决这个问题^[7]。具体步骤包括:①利用 Lillifors 检验法和卡方检验法对共词矩阵进行数据检验,结果判定矩阵数据不符合正态分布规律,也不符合均匀分布规律;②鉴于矩阵数据不满足正态分布,本研究采用 Spearman 分析法分析 46 个关键词之间的相关关系,并抽取相关分析中的相关系数形成相关矩阵进行因子分析。具体方法是利用主成分法、协方差矩阵与平均正交旋

转方法进行因子分析,萃取 8 个公共因子,因子个数碎石图显示最佳的公共因子个数区间为 [6,8];③考虑到矩阵数据为离散数据,为了更好地满足进行聚类分析的条件,本研究用“1”与上述获得的相关矩阵相减,得到表示两两词间相异程度的相异矩阵;然后基于该相异矩阵进行分层聚类分析。具体而言,本研究将聚类个数的范围值选定为 6-8 个,在聚类方法方面选择离差平方和法,在距离测度方法上选择离散型数据类型中的德尔斐方法。

第三阶段:研究结果的分析和讨论。通过上述两个阶段,获取了对国内开放存取研究论文的关键词进行共词分析的聚类结果,然后根据论文本身对这些研究结果进行进一步的阐释说明。由于论文数量众多,质量参差不齐,本研究的重点为引用发表在核心期刊上的高被引论文。核心期刊的确定参考 CSSCI(中文社会科学引文索引)2010-2011 年来源期刊目录(不包括扩展版来源期刊),论文的被引次数则参考本研究于 2010 年 2 月 1 日在 CNKI 的引文数据库中的检索结果(检索式:被引文献关键词 = “开放存取” + “开放获取”):至少被引 1 次以上的论文共有 392 篇,其中被引频次不小于 6 的论文共有 145 篇。

3 研究主要发现

3.1 确定的高频关键词

通过关键词词频统计,本研究发现共有 615 个关键词。其中,频次为 1 的关键词共有 471 个,占 76.6%;频次为 2-4 的关键词共有 98 个,占 15.9%;而频次在 5 次或 5 次以上的关键词有 46 个,占 7.5%。表 1 列出了前 45 个关键词,最后一个为“影响因子”,由于表格空间受限未列出。可以看出“开放存取”是出现最为频繁的关键词,其频次远远高于排在第二和第三的“图书馆”和“学术传播”。

表 1 国内开放存取研究论文的高频关键词统计(频次≥5)

序号	关键词	词频	序号	关键词	词频	序号	关键词	词频
1	开放存取	575	16	学术出版	27	31	发展现状	11
2	图书馆	103	17	出版模式	26	32	学术文献	11
3	学术传播	103	18	影响	23	33	电子期刊	10
4	机构知识库	98	19	预印本	23	34	数字资源	8
5	OA 期刊	96	20	质量控制	19	35	图书情报学	8
6	信息资源	95	21	OA 出版	17	36	DOAJ	7
7	学术期刊	73	22	自存储	17	37	科学信息	7
8	大学图书馆	70	23	公共存取	15	38	文献计量	7
9	知识产权	59	24	数字图书馆	15	39	定量分析	6
10	OA 知识库	40	25	信息服务	15	40	开源软件	6
11	信息共享	39	26	政策	14	41	数字出版	6
12	OA 资源	38	27	经济机制	12	42	非正式交流	5
13	对策	38	28	期刊评价	12	43	同行评审	5
14	互联网	36	29	长期保存	11	44	学科知识库	5
15	资源建设	33	30	发展	11	45	引文分析	5

3.2 构建的共词矩阵及经由转换的相异矩阵

表 2 是基于上述选定的 46 个高频词构建的共词矩阵(部分),代表两两关键词在同一篇文献中出现的次数。从中看出由于“开放存取”本身出现的词频远远高于其他关键词,因此“开放存取”与其它关键词同时出现在同一篇文献中

的概率也是最高的。表 3 则是由表 2 转换而来的相异矩阵,其中经过相关矩阵转换这一环节。转换为相异矩阵的目的是为了消除由共词频次差异所带来的影响,从而更好地满足进行聚类分析的前提条件。下面的聚类分析就是以该相异矩阵作为源数据而开展的。

表 2 国内开放研究研究论文关键词的共词矩阵(部分)

	开放存取	图书馆	学术传播	机构知识库	OA 期刊	信息资源	学术期刊	大学图书馆	知识产权	OA 知识库	信息共享	OA 资源	对策	互联网
开放存取	572	90	91	77	61	80	53	60	42	37	28	21	35	28
图书馆	90	103	13	12	6	21	3	0	2	5	5	3	9	2
学术传播	91	13	101	19	7	14	5	6	1	4	3	0	4	5
机构知识库	77	12	19	92	0	5	0	6	7	3	4	0	3	0
OA 期刊	61	6	7	0	97	5	8	4	7	25	2	2	4	0
信息资源	80	21	14	5	5	96	1	7	5	3	2	2	6	11
学术期刊	53	3	5	0	8	1	73	3	4	2	3	0	0	5

续表

	开放存取	图书馆	学术传播	机构知识库	OA 期刊	信息资源	学术期刊	大学图书馆	知识产权	OA 知识库	信息共享	OA 资源	对策	互联网
大学图书馆	60	0	6	6	4	7	3	70	1	0	6	7	8	3
知识产权	42	2	1	7	7	5	4	1	58	5	2	3	0	1
OA 知识库	37	5	4	3	25	3	2	0	5	42	2	0	1	0
信息共享	28	5	3	4	2	2	3	6	2	2	38	0	0	0
OA 资源	21	3	0	0	2	2	0	7	3	0	0	37	2	4
对策	35	9	4	3	4	6	0	8	0	1	0	2	38	1
互联网	28	2	5	0	0	11	5	3	1	0	0	4	1	36

表 3 国内开放存取研究论文关键词的相异矩阵(部分)

	开放存取	图书馆	学术传播	机构知识库	OA 期刊	信息资源	学术期刊	大学图书馆	知识产权	OA 知识库
开放存取	.000	.225	.314	.427	.529	.171	.604	.304	.336	.361
图书馆	.225	.000	.510	.488	.672	.339	.756	.377	.525	.512
学术传播	.314	.510	.000	.667	.754	.333	.539	.529	.513	.354
机构知识库	.427	.488	.667	.000	.790	.547	.999	.779	.536	.501
OA 期刊	.529	.672	.754	.790	.000	.544	.375	.746	.581	.429
信息资源	.171	.339	.333	.547	.544	.000	.525	.360	.433	.427
学术期刊	.604	.756	.539	.999	.375	.525	.000	.754	.630	.580
大学图书馆	.304	.377	.529	.779	.745	.360	.754	.000	.733	.767
知识产权	.336	.525	.513	.536	.581	.433	.630	.733	.000	.385
OA 知识库	.361	.512	.354	.501	.429	.427	.580	.767	.385	.000

3.3 聚类分析形成的类别分属聚类表与树状图

通过聚类分析,形成了如表 4 所示的分属聚类表,包括区分三种类别情况下各个关键词的归属状态。通过比较,发现在 6-8 类的取值期间,有 4 类的结果一直是稳定的,即由“开放存取”、“图书馆”、“信息资源”、“信息共享”、“对策”和“政策”6 个关键词组成的第一研究主题;由“学术传播”、“机构知识库”、“预印本”、“图书情报学”、“科学信息”、“发展现状”、“学术文献”、“非正式交流”和“学科知识库”9 个关

键词组成的第二研究主题;由“OA 期刊”、“学术期刊”、“学术出版”、“质量控制”、“OA 出版”、“经济机制”和“数字出版”7 个关键词组成的第三研究主题;由“大学图书馆”、“OA 资源”、“互联网”、“资源建设”、“影响”、“公共存取”、“信息服务”和“发展”8 个关键词组成的第四研究主题。

剩下的 16 个关键词归属则随着分类个数的变化有所区别。通过分析和对比,本研究认为分为 7 类最为合适。以区分为 7 类所形成的 7 个研究主题作为参考系的话,分为 8 类是将“文

献计量”与“定量分析”从原先的第七研究主题中独立出来,但这两个关键词组成的新的研究主题在内容方面并没有太大的指示性作用。分为6类则是将原先的第三研究主题和第五研究主题进行了合并,从内容的角度来看,是将关于OA 期刊出版的研究和关于自存储和 OA 知识的研究进行了合并,合并的结果可看成是关于开

放存取两种主要实现途径的研究。但本研究认为开放存取的两种实现途径各成体系,对之进行区别有助于更加精确地揭示国内学者对这两种实现途径的研究现状。因此,本研究最终选择对相异矩阵区分为七大类进行聚类的分析结果。

表 4 类别分属聚类表(6-8 类)

关键词	8 类	7 类	6 类	关键词	8 类	7 类	6 类
开放存取	1	1	1	公共存取	4	4	4
图书馆	1	1	1	数字图书馆	6	6	5
学术传播	2	2	2	信息服务	4	4	4
机构知识库	2	2	2	政策	1	1	1
OA 期刊	3	3	3	经济机制	3	3	3
信息资源	1	1	1	期刊评价	7	7	6
学术期刊	3	3	3	长期保存	6	6	5
大学图书馆	4	4	4	发展	4	4	4
知识产权	5	5	3	发展现状	2	2	2
OA 知识库	5	5	3	学术文献	2	2	2
信息共享	1	1	1	电子期刊	7	7	6
OA 资源	4	4	4	数字资源	6	6	5
对策	1	1	1	图书情报学	2	2	2
互联网	4	4	4	DOAJ	7	7	6
资源建设	4	4	4	科学信息	2	2	2
学术出版	3	3	3	文献计量	8	7	6
出版模式	5	5	3	定量分析	8	7	6
影响	4	4	4	开源软件	6	6	5
预印本	2	2	2	数字出版	3	3	3
质量控制	3	3	3	非正式交流	2	2	2
OA 出版	3	3	3	同行评审	7	7	6
自存储	5	5	3	学科知识库	2	2	2
				引文分析			

图 1 为本研究将类别确定为 7 类的情况下进行分层聚类形成的聚类结果树形图。本研究将这 7 类或这 7 个研究主题分别命名为(按树

状图从上往下的顺序):OA 期刊质量评价;自存储和 OA 知识库;OA 期刊出版;OA 资源长期保存;开放存取与学术传播;开放存取的发展对

策;开放存取对图书馆的影响。在讨论部分本

研究主题进行重新排序并援引文献本身对它们

研究将根据研究主题的内容逻辑关系对这 7 个

进行详细说明。

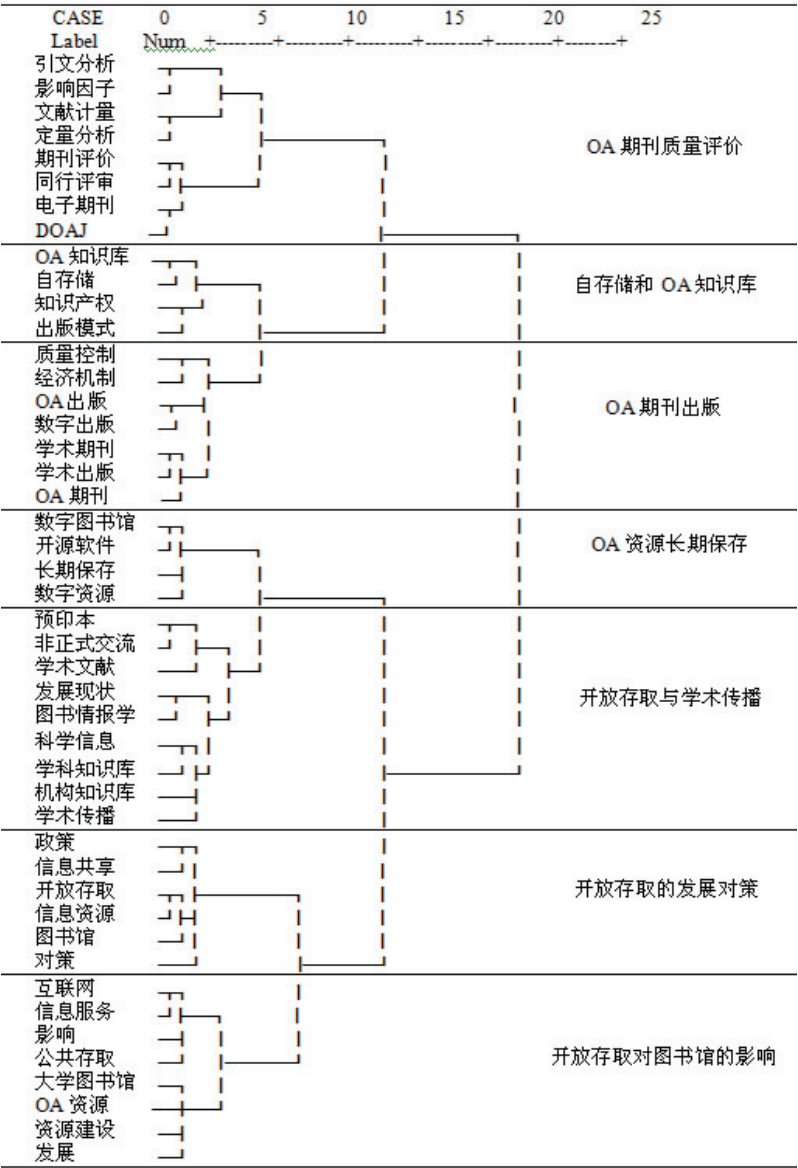


图 1 聚类结果树状图

4 讨论

内开放存取研究领域的七大主题,进而探讨目

目前国内在开放存取研究领域所呈现出来的

特征。

本部分将根据分层聚类的结果详细分析国

4.1 研究的七大主题

4.1.1 OA 期刊出版

布达佩斯开放存取先导计划提出了开放存取的两种主要实现途径:自存储和 OA 期刊^[8-9]。国内学者最初对开放存取的关注集中在 OA 期刊。李武和刘兹恒在 2004 年发表了《一种全新的学术出版模式:开放存取出版模式探析》一文,这是国内第一篇系统介绍 OA 期刊出版的学术论文,该论文基于对两大 OA 期刊出版商 BMC 和 PLoS 的案例分析了 OA 期刊出版的三大核心问题:经济机制、质量控制和知识产权^[10]。任胜利、李麟和初景利等人也较早地从出版角度对开放存取开展了探索性研究^[11-12]。在此基础上,王应宽以个案分析的形式详细介绍了两大 OA 期刊出版商的创办历史和运作模式^[13-14],傅蓉对上述提到的三大核心问题分别开展了深入研究^[15-17]。而刘海霞等人也对 OA 期刊的经济运作模式给予了较大的关注^[18-19]。在结合我国学术期刊开展的本土实证调查方面,主要有李麟对我国科研人员对学术信息开放存取态度的考察^[20],李武等人调查了中华医学会系列期刊对开放存取的认知和态度^[21],而王应宽则对中国学术界对 OA 期刊认知度和认可度进行了大样本调查分析^[22]。

4.1.2 自存储和 OA 知识库

紧随 OA 期刊之后,自存储也很快引起了国内研究者的关注。乔冬梅对 e 印本文库及其创建工作给予了较早的关注^[23-25]。秦珂一直关注于 OA 资源及作者自存储的版权问题的研究,提出在构建自存储制度中于满足期刊社的合理要求的同时应尊重作者的自存储权利^[26-28]。在 OA 知识库的两种主要形式之一的机构知识库的概念引进方面,常唯做了较好的推广工作^[29-30]。随后,研究者从不同的角度对机构知识库进行了多维广泛研究。比如说,陈和、唐兆琦和祝忠明分别结合自己所在单位的实践工作从技术角度论述了机构知识库的开发工作^[31-33]。王学勤、邵晶、李大玲等人对机构知识库建设的关注集中在激励机制和建设政策方面^[34-37]。张晓林则结合国家科学图书馆的实际需要,对机构知识库的政策、功能和支撑机制

进行颇为深入的分析^[38]。在结合我国科研机构 and 科研人员开展的本土实证调查方面,主要有韩珂和祝忠明合作开展的科研机构对机构知识仓储认知和服务需求调查分析^[39];何燕等人则以中国科学院人员为例开展了自存储需求分析和态度调查^[40-41]。

4.1.3 OA 期刊的质量评价

对 OA 期刊的出版模式研究通常包括经济机制、质量控制和知识产权,尽管 OA 期刊的评价问题与质量控制有着密切的联系,但对其研究往往是相对独立的。来自中南大学湘雅医学院医药信息系的研究团队对 OA 期刊的质量评价有着深入系统的研究。他们调查了 DOAJ 数据库和 ISI 期刊引文报告的相关数据^[42],建议采用综合评价法对 OA 期刊质量进行评价,建立一套科学、合理、实用的质量评价体系^[43],同时对 OA 期刊论文和非 OA 期刊论文的影响力进行了对比分析^[44]。张红芹和黄水清两人也于 2008 年初步探索了 OA 期刊质量评价指标体系^[45]。随后,他们以 DOAJ 收录的 44 种化学化工期刊为研究对象进行了多元指标的相关性分析及核心期刊评价研究,从实证角度验证了所构建的指标体系的合理性与实用性^[46]。另外,刘辉也曾将 DOAJ、HWP、PMC 和 BMC 这四个具有代表性的 OA 数据库与 Elsevier Science Direct 系统进行比较,分别从学科分布、检索特性、影响因子和相关性角度对其进行了评价^[47-48]。

4.1.4 OA 资源的长期保存

数字信息的长期保存问题是数字图书馆建设的核心问题之一,OA 资源的传播主要依赖于互联网媒介,因此 OA 资源的长期保存问题也得到国内许多学者的重视。孙红娣指出数字资源的长期保存是开放存取发展成败的关键,并分析了开放存取中不同类型的数字资源长期保存的策略问题,包括技术策略、系统框架以及法律问题,并就我国长期保存数字资源提出了一些建议^[49]。李春明从长期保存的角度分析了图书馆与开放存取之间的关系,提出了图书馆进行开放存取资源长期保存需要的相关因素,包括技术框架、经费预算和合作机制等^[50]。随后,陈月婷和李春明两人结合我国实际,探讨了我国

开放存取资源长期保存模式选择的相关问题,具体包括长期保存整体框架、责任体系及建设原则等^[51]。黄如花等人则调查了澳大利亚开放存取项目 ARROW 中各机构知识库的长期保存政策制定情况,并选取国外具有代表性的机构知识库 ALADIN Research Commons、Nottingham Eprints 和 Griffith Research Online 的长期保存政策进行了对比分析,提出了 OA 资源长期保存政策的发展方向^[52]。

4.1.5 开放存取对图书馆的影响

开放存取的兴起在很大程度上就是为了解决期刊价格不断攀升和图书馆相对有限的经费支出之间的矛盾,OA 资源的快速增长已经给图书馆的资源建设和读者服务两大核心业务都带来了较大影响。因此这一研究主题不可避免地成为国内学者关注开放存取的切入点之一。黄凯文在 2005 年就较为系统地分析了开放存取运动对图书馆的影响,包括对图书馆经费预算和图书馆在学术传播体系中的角色和职能变化的影响等^[53]。任真也较早地分析了开放存取环境下的图书馆,重点剖析开放存取对资源建设、信息服务和图书馆的角色变化所带来的影响,认为在开放存取环境下图书馆将在信息流而不是在资金流中发挥更好的作用^[54]。随后,众多学者纷纷撰文分析开放存取的发展给图书馆带来的深远影响。

4.1.6 开放存取的发展对策

开放存取的发展有赖于多方的通力合作。初景利在政策层面系统地分析了开放存取的发展的推动因素,并指出开放存取还是一项新的事业,需要科研人员和科研机构、学术团体、科研资助机构、大学师生、图书馆员、政府机构以及每个公民的积极参与和鼎力支持,并在其中发挥应有的作用^[55]。黄凯文明确提出了开放存取的可持续发展战略^[56]。陈振英等人在分析开放存取在国内发展困境的基础上提出了发展开放存取的中国特色之路^[57]。张晓林等人则站在战略高度,提出通过落实责任机制、坚持法律有效性、提高可持续性、建立系统化的支撑机制等促进开放存取发展^[58]。值得注意的是,国内很多学者对开放存取的对策研究通常立足于

图书馆,往往在论述开放存取对图书馆影响的同时提出图书馆的应对措施。马景娣较早地提出图书馆应积极支持并参与开放存取运动^[59],并进一步明确了大学图书馆在开放存取运动中的角色:领导者、实施者、宣传者和开发利用者^[60]。

4.1.7 开放存取与学术传播

为了解决传统的学术传播危机,开放存取的兴起已对传统的学术出版和学术传播产生了深远的影响。乔冬梅在 2004 年就基于学术交流的视角对国外的开放存取运动作了较为全面的梳理和介绍,并总结了开放存取运动对我国学术交流发展的几点启示^[61]。吴建中在 2005 年开始思考开放存取环境下的信息共享空间的构建问题^[62]。李武等人在揭示 OA 期刊发展现状的基础上深入分析了 OA 期刊出版对当前学术交流系统的各利益关系方(包括科研人员、传统出版机构、图书馆和科研资助机构等)所产生的积极影响^[63]。随后,黄如花等人探讨了由开放存取带来的新的科学信息交流和利用模式,阐述了它对科学信息交流和利用活动参与者的影响^[64]。在传播效果方面,尹达从 CNKI 收录的 1094 种医学期刊中遴选出国内医学领域的 OA 期刊,采用定性与定量结合的方式,从认知效果、态度效果和行为效果三个层面对我国医学领域的 OA 期刊的传播效果进行了实证研究^[65]。

4.2 研究特点分析

通过研究发现,国内学者对开放存取的研究已经形成了相对集中的主题领域。首先,关于开放存取本身形态的研究是目前国内学者的重点研究内容,包括 OA 期刊的出版模式和机构知识库的建设问题,前者的研究内容细化为经济机制、质量控制、知识产权和期刊评价等,后者的研究内容则涉及技术开发、版权问题、建设政策和 OA 资源的长期保存等。在传播学者丹尼斯看来,学者们对新媒介的研究最初通常会集中在形态研究上^[66],这也就较好地解释了目前该领域以开放存取本身形态研究为主导的局面。其次,开放存取与学术传播系统(尤其是图

书馆)的关系也是目前国内学者所关心的研究主题,这与开放存取运动的动因和目标直接相关。开放存取运动旨在解决“学术传播危机”,建立真正服务于科学研究的学术传播体系。而图书馆是学术传播体系中的重要环节,同时关注和研究开放存取的学者大多数来自图书馆理论和实务界,这就很好地解释了为什么“开放存取对图书馆的影响”也是学者们研究的焦点,同时也解释了为什么国内很多学者对开放存取对策的研究通常是基于图书馆立场的。

通过对上述七大研究主题内容和研究深度的比较,我们也发现尽管目前对开放存取的研究同时涵盖了丹尼斯所谓的新媒介在不同阶段的三个层面研究,即形态研究、受众分析和效果研究,但从研究的广度和深度来看,基本还是停留在形态研究阶段。换言之,对开放存取本身形态的研究是目前研究的主导内容。在受众分析层面,虽然有少数学者从用户接受的角度实证调研了科研人员的接受行为,但这些研究不仅数量稀少,而且都隶属于用户对 OA 期刊或自存储行为的认知和态度的描述性研究,并未通过挖掘变量之间的关系开展解释性研究,比如通过构建理论模型探索和确定阻碍或促进用户接受开放存取的相关影响因素。而在效果研究层面,尽管不少文献也阐述了开放存取对图书馆和学术传播的影响,但这些研究基本属于陈学飞等人所谓的思辨性研究、文献研究或规范性研究^[67],存在研究方法不规范、研究力度不深入等问题。结合目前的研究现状,本研究认为今后国内学者对开放存取的研究应该在受众分析和效果研究这两个层面加强力度,针对国内的科研人员和学术传播体系开展扎实的本土化研究工作。

5 结论

本研究通过对 728 篇文章关键词的共词分析,直观形象地揭示了国内开放存取研究的七大热点,具有一定的现实意义。这一研究发现不仅对于我们了解国内学者对开放存取的关注热点具有较大的参考价值,而且也可以说在某

种程度上为我们今后继续开展研究指明了方向,包括加强某一现有研究领域和开拓某一新的研究领域。但总体而言,本研究尚属探索性研究,还存在许多不足之处。比如说,在高频关键词的阈值确定方面,没有非常充足的理论支撑,很大程度上只是根据以往经验选定 5 作为高频词的阈值,具有一定的主观性。另外,所考察的文献类型特指期刊论文,并没有涉及其它类型的学术文献。以学位论文为例,自从 2005 年以来,国内已有多篇开放存取研究领域的硕博论文问世。如果将这些学位论文也统计在内的话,也许对本研究的结论会起到一定的修正作用。

参考文献:

- [1] 初景利,李麟.国内外开放获取的新发展[J].图书馆论坛,2009(6):83-88.
- [2] 王云才.国内外“开放存取”研究综述[J].图书情报知识,2005(6):40-45.
- [3] 陈红勤.2003-2007 年我国开放存取研究文献计量学分析[J].情报科学,2008(9):1317-1322.
- [4] 刘锦宏.我国开放获取研究论文计量分析[J].图书情报知识,2009(6):108-114.
- [5] 崔雷,郑华川.关于从 MEDLINE 数据中进行知识抽取和挖掘的研究进展[J].情报学报,2003(4):425-433.
- [6] 张勤,马费成.国外知识管理研究范式——以共词分析为方法[J].管理科学学报,2007(12):65-74.
- [7] 马费成,宋恩梅,张勤.IRM-KM 范式与情报学发展研究[M].武汉:武汉大学出版社,2008:24.
- [8] BOAI. Open Access as a Means to Promote Academic Publishing [J/OL]. [2010-01-25]. <http://www.soros.org/openaccess/meeting.shtml>.
- [9] 李武.开放存取出版的两种主要实现途径[J].大学图书馆学报,2005(4):58-63.
- [10] 李武,刘兹恒.一种全新的学术出版模式:开放存取出版模式探析[J].中国图书馆学报,2004(6):66-69.
- [11] 任胜利.开放存取(Open Access):现状与展望[J].中国科技期刊研究,2005(2):151-154.

- [12] 李麟,初景利. 开放获取出版模式研究[J]. 图书馆论坛, 2005(6):88-93.
- [13] 王应宽,王锦贵. 基于赢利模式的开放存取期刊出版:BioMed Central 案例研究[J]. 中国科技期刊研究, 2006(3):354-359.
- [14] 王应宽. 开放存取期刊出版:PLoS 案例研究[J]. 出版发行研究, 2006(5):59-64.
- [15] 傅蓉. 开放存取期刊的经济机制[J]. 中国图书馆学报, 2006(5):32-35.
- [16] 傅蓉. 开放存取的质量控制[J]. 情报理论与实践, 2006(6):694-696.
- [17] 傅蓉. 开放存取的版权问题[J]. 图书馆理论与实践, 2006(6):40-42.
- [18] 刘海霞,孙振球,胡德华,等. 开放存取期刊的经济学分析[J]. 情报理论与实践, 2007(1):48-51.
- [19] 刘海霞,孙振球,胡德华,等. 开放存取期刊质量的双边市场模型研究[J]. 情报杂志, 2007(6):118-121.
- [20] 李麟. 我国科研人员对科技信息开放获取的态度——以中国科学院科研人员为例[J]. 图书情报工作, 2006(7):34-38,50.
- [21] 李武,沈锡宾,季燕江. 医学期刊对开放存取的认知和态度研究——基于中华医学会系列期刊的问卷调查[J]. 中国科技期刊研究, 2008(1):44-47.
- [22] 王应宽. 中国科技界对开放存取期刊认知度与认可度调查分析[J]. 中国科技期刊研究, 2008(5):753-762.
- [23] 乔冬梅. e 印本文库(e-print archive)研究——介绍一种非正式科学交流系统[J]. 图书情报知识, 2006(1):66-69.
- [24] 乔冬梅. 我国 e 印本文库的发展策略[J]. 图书情报工作, 2006(1):59-63.
- [25] 乔冬梅. e 印本文库创建过程研究[J]. 情报理论与实践, 2006(2):232-235.
- [26] 秦珂. 开放存取版权管理的特点分析[J]. 情报理论与实践, 2006(4):409-412,502.
- [27] 秦珂. 试论开放存取资源的版权保护[J]. 大学图书馆学报, 2006(5):59-62,99.
- [28] 秦珂. 开放存取自存档(self-archiving)的版权问题分析[J]. 图书与情报, 2008(1):103-105,114.
- [29] 常唯. 机构知识库: 数字科研时代一种新的学术交流与知识共享方式[J]. 图书馆杂志, 2005(3):16-19.
- [30] 常唯. 数字学术环境中的机构知识库探析[J]. 图书情报工作, 2006(7):46-50.
- [31] 陈和,萧德洪,林丽敏. 基于 DSpace 构建机构仓储的本地化实践[J]. 现代图书情报技术, 2007(3):13-17.
- [32] 唐兆琦,郑巧英,潘卫. 机构仓储系统(IR)功能扩展初探[J]. 图书馆杂志, 2007(12):53-57.
- [33] 祝忠明,马建霞,张智雄,等. 中国科学院联合机构仓储系统的开发与建设[J]. 图书情报工作, 2008(9):90-93,144.
- [34] 王学勤. 机构知识库建设相关政策研究[J]. 中国图书馆学报, 2007(3):44-47.
- [35] 都平平. 机构仓储的自存储和强制存储策略研究[J]. 图书馆杂志, 2008(9):15-18.
- [36] 邵晶. 绿色 OA 仓储的“存档”与“开放”策略研究[J]. 图书情报工作, 2008(11):78-80,142.
- [37] 李大玲,柯平. 基于知识管理的学术机构知识库激励模式研究[J]. 图书情报工作, 2009(10):98-101.
- [38] 张晓林. 机构知识库的政策、功能和职称机制分析[J]. 图书情报工作, 2008(1):23-27,19.
- [39] 韩珂,祝忠明. 科研机构对机构知识仓储认知和服务需求调查分析[J]. 现代图书情报技术, 2008(3):12-17.
- [40] 何燕,初景利,张冬荣. 自存储的需求分析[J]. 情报理论与实践, 2008(3):428-431.
- [41] 何燕,初景利,张冬荣. 我国科研人员自存储态度调查——以中国科学院人员为例[J]. 图书情报工作, 2008(5):121-124.
- [42] 刘海霞,方平,胡德华. 开放存取期刊的质量评价研究[J]. 图书馆杂志, 2006(6):23-27.
- [43] 陶雯,胡德华,曲艳吉,等. 开放存取期刊质量评价方法研究[J]. 图书情报工作, 2006(10):72-75.
- [44] 胡德华,常小婉. 开放存取期刊论文质量和影响力的评价研究[J]. 图书情报工作, 2008(2):61-64.
- [45] 张红芹,黄水清. OA 期刊质量评价指标体系初探[J]. 情报杂志, 2007(3):124-126.
- [46] 张红芹,黄水清. 开放获取期刊质量评价的指标体系构建与评价实现——以化学类期刊为例[J]. 情报理论与实践, 2008(3):386-390.

- [47] 刘辉. 开放获取期刊的现状与评价分析[J]. 图书馆建设, 2006(4):44-47.
- [48] 刘辉. 开放存取期刊数据库的评价[J]. 大学图书馆学报, 2007(1):59-63.
- [49] 孙红娣. 论开放存取中的数字资源长期保存问题[J]. 图书馆学研究, 2005(11):15-18.
- [50] 李春明. 图书馆与开放存取资源之长期保存[J]. 图书馆建设, 2007(4):52-55.
- [51] 陈月婷, 李春明. 开放存取资源长期保存模式探析[J]. 图书情报工作, 2008(5):33-36, 120.
- [52] 黄如花, 胡文琴. 开放存取资源长期保存政策的调查与分析[J]. 图书与情报, 2009(5):70-74.
- [53] 黄凯文. 试析网络科学信息的 OA 运动对图书馆的影响[J]. 图书馆论坛, 2005(2):112-114.
- [54] 任真. 开放获取环境下的图书馆[J]. 大学图书馆学报, 2005(5):44-47.
- [55] 初景利. 开放获取的发展与推动因素[J]. 图书馆论坛, 2006(6):238-242.
- [56] 黄凯文. 科学信息开放获取的可持续发展探究[J]. 情报理论与实践, 2007(6):760-763.
- [57] 陈振英, 何小军, 陈益君. 开放存取在中国的困境及对策分析[J]. 大学图书馆学报, 2008(2):35-40, 67.
- [58] 张晓林, 刘兰, 李麟, 等. 科技信息开放获取的内涵演变、责任意义和实施战略[J]. 图书情报工作, 2009(5):28-33, 114.
- [59] 马景娣. 学术文献开放访问和图书馆的应对策略[J]. 中国图书馆学报, 2005(4):38-41.
- [60] 马景娣. 开放访问运动中大学图书馆的定位和作用[J]. 图书馆建设, 2005(4):91-93.
- [61] 乔冬梅. 国外学术交流开放存取发展综述[J]. 图书情报工作, 2004(11):74-78.
- [62] 吴建中. 开放存取环境下的信息共享空间[J]. 国家图书馆学刊, 2005(3):7-10.
- [63] 李武, 杨屹东. 开放存取期刊出版的发展现状及其影响分析[J]. 图书情报工作, 2006(2):25-30.
- [64] 黄如花, 冯晴. 论开放存取出版对科学信息交流和利用的影响[J]. 出版科学, 2008(3):75-78.
- [65] 尹达. 我国医学领域开放存取型学术期刊传播效果实证研究[J]. 图书情报工作, 2009(5):145-149.
- [66] [英] 丹尼斯·麦奎尔. 受众分析(中文版)[M]. 刘燕南, 等, 译. 北京: 中国人民大学出版社, 2006.
- [67] 陈学飞, 茶师俊. 理论导向的教育政策经验研究探析[J]. 北京大学教育评论, 2007(10):31-42.

李 武 上海交通大学媒体与设计学院助理研究员。通讯地址: 上海市东川路 800 号。邮编: 200240。

董 伟 河北大学管理学院硕士研究生。通讯地址: 河北省保定市五四东路 180 号。邮编: 071002。

(收稿日期: 2010-02-11; 修回日期: 2010-03-06)