

商业引文数据库与开放获取引文系统的比较研究*

蔡明月

(台湾政治大学图书资讯与档案学研究所, 台北 116)

摘要: 本文旨在以商业引文数据库和开放获取引文系统为研究对象, 深入比较两种学术传播机制的特色及涵盖资源的完整性、重复性与独特性。研究以经济学开放获取期刊论文系统Research Papers in Economics收录的2001—2012年共25位诺贝尔经济学奖得主的著作为样本, 采用文献计量法, 交叉比对两两数据库的重复性及独特性。所选的6个数据库及系统, 包括商业引文数据库Web of Science及Scopus, 开放获取引文系统Google Scholar、Microsoft Academic Search、OAster、OpenDOAR。其中, Google Scholar及Microsoft Academic Search为综合性学科搜索引擎系统, OAster、OpenDOAR为汇集式机构典藏系统。研究结果表明: 在数据完整性、重复性和独特性上, 搜索引擎最优, 其次为商业引文数据库, 再次为汇集式机构典藏系统。

关键词: 商业性引文数据库; 开放获取引文系统; 数据库完整性; 数据库重复率

中图分类号: G255.51; G252.7

DOI: 10.3772/j.issn.1673-2286.2016.10.004

1 引言

学术研究是国家发展的重要基础。随着传播媒体及学术传播方式的改变, 学者传递知识的途径也不断演进。2000年以来, 计算机及网络的快速发展, 促使信息生产、检索、传递和利用发生很大变化。商业机制与开放获取机制的引文数据库, 皆因学术传播而产生, 但其产生的目的和方法不同, 因此, 本研究旨在通过商业引文数据库和开放获取引文系统, 深入比较两种学术传播机制的特色及涵盖资源的完整性、重复性与独特性, 以期为研究者与图书馆选择引文索引资料库或引文系统提供参考。此外, 基于各系统的评估结果, 提出学术评鉴指标与工具的建议。

2 文献探讨

文献重复性及独特性相关研究由Bradford于1937年提出, 主要探讨期刊文献被索引摘要工具收录的重复性, 即引用文献在2个以上的数据库同时出现的情

况; Yerkey等提出的分散理论, 也指同主题论文分散在不同数据库中收录的情况^[1]。Martyn利用Bradford提出的重复性理论, 对其研究的摘要期刊收录范围进行分析, 发现超过一半的文献被收录在1个以上的索引与摘要数据库, 认为索引与摘要数据库的重复性会导致用户查询到重复数据, 造成人力浪费^[2]。

Wood等从Biological Abstracts、Chemical Abstracts及Engineering Index数据库选取14 592种期刊, 并对标题字段进行分析, 结果显示: 1%的期刊同时被3个数据库收录, 27%的期刊同时被2个数据库收录, 72%的期刊被1个数据库收录, 显示其唯一的独特性^[3]。

Read等针对图书馆与信息科学领域的数据库Library and Information Science Abstracts (LISA)、Library Literature and Information Science (LLIS) 及Information Science Abstracts (ISA), 包含CD-ROM和在线两种类别数据库进行重复性研究^[4]。结果显示, LLIS包含的数据最多(30 542篇), ISA最少(5 094篇); 在重复性比较上, LISA和ISA数据重复性为12.2%, LISA和LLIS的数据重复性为10.3%, ISA和LLIS的数据重复性为5.8%, 3个数

数据库同时存在的重复性数据低于3%，由此可知，3个数据库的数据独特性很高。

Walters等对7个单一学科数据库及5个跨学科数据库作内容比较，二者分别收集1990—2000年美国与加拿大出版的later-life migration主题相关的数据，经检验后选取155篇进行研究^[5]，发现5个跨学科数据库中有4个数据库收录的数据比任何单一学科数据库收录的数据都要完整，不同数据库间平均重复率约45%。

Rather等对Google、AltaVista、HotBot、Scopus、Bioweb 5个搜索引擎进行重复性研究^[6]。结果表明，HotBot与其他搜索引擎的平均重复性最高，其次是Google，另外，复合和复杂式的检索查询会产生更多重复结果。

Esmacil等研究比较searchenginewatch.com网站介绍的6个常用搜索引擎系统及数据库的重复性^[7]。研究发现，在不同的搜索引擎系统中，Yahoo与其他搜索引擎系统的平均重复率最高（约40%）；Curry Guide检索物理学类文献的召回率为77%，与其他5个系统的平均重复率约44%。

Wang等以乳腺癌为检索词对Google、Yahoo、Bing和Ask.com进行分析比较^[8]。研究结果显示，4个搜索引擎对乳腺癌的6个标准检索结果均排列前30名。从有效性看，Google最好，其次为Bing与Ask.com，最差为Yahoo；从重复性看，各搜索引擎两两对比，其收录数据重复性均超过50%；从内容看，4个搜索引擎皆强调不同类型的内容；从使用者满意度看，Bing的满意度最好，其次为Yahoo与Google，最差的是Ask.com。研究认为依据不同的信息类型需求，选取适合的搜索引擎才能事半功倍。

重复性与独特性的研究范围十分广泛，既包括出版社、数据库与搜索引擎等数据源提供者，也包括各种数据源类型（如期刊、专利、网络资源等）。重复性和独特性的优劣，是图书馆选择索引与摘要数据库的重要依据。

3 研究方法

本研究采用文献探讨法及文献计量法比较Web of Science、Scopus、Google Scholar、Microsoft Academic Search、OAIster、OpenDOAR数据库及系统，以RePEc的IDEAS系统提供的诺贝尔经济学奖得主2001—2012年的文献清单为基准^[9]，针对上述各数据库及系统进行文献核对；利用Excel工具并辅以人工方式加以查证，评鉴各数据库及系统的完整性；再用两两交叉比对的方式比较各数据库及系统间的重复性及独特性，评鉴

其优劣。一般而言，数据库评鉴的假设：收录数据越完整，与其他数据库重复率越低，说明其独特性越高，涵盖数据范围越优异。

RePEc用户使用的界面为IDEAS，IDEAS特别设置诺贝尔经济学奖得主所著的图书及各类文献，并在每月的第1天更新诺贝尔得主的文献清单^[10]。本研究考虑期刊文献为学术传播的主体，因此，图书、网站资源、数字资源等皆不列入本研究范围。

4 研究结果

本文就商业引文数据库及开放获取引文系统现况与特性、完整性、重复性与独特性进行研究分析。

4.1 商业引文数据库及开放获取引文系统现况与特性

商业数据库受营利性出版商的建设和定期维护，具有数据完整、学科范围多元且选择数据政策严谨等特色，因此其数据具有权威性。

4.1.1 商业引文数据库

（1）Web of Science为美国汤森路透（Thomson Reuters）于1997年建立的网络版引文索引数据库，提供理、工、医、农、人文及社会科学等各学科领域的文献书目、摘要及引用文献等资料，收录期刊超过1.3万种，更新频率为1次/周^[11]。该数据库特有的文献间相互引用关系检索，可提供研究人员更完整的研究参考信息，使其可依循前人研究轨迹，进一步了解同行及竞争对手的研究。

（2）Scopus是由Elsevier公司于2002年建立，2004年11月正式推出的索引摘要数据库。截至2016年1月，Scopus收录超过2.1万种同行评审期刊（包含4 200种开放获取期刊），且有超过5 000条国际出版社提供的摘要与引文索引（更新频率1次/天），包括文献纪录6 000万篇，回溯建档文献2 200万篇；内容包括学术期刊、会议论文及书评等。通过Scopus读者可检索到广泛的人文与艺术领域、社会科学及生命科学领域等学术文献，并通过单一接口提供整合检索的功能。Scopus收录论文除依据使用者意见外还遵循①收录文献必须为学术性刊物；②收录期刊需有同行评审制度，以确保论文质量；③论文必需提供英文摘要，且期刊必须定期出

版,以保持资料库收录文献的稳定性^[12]。

4.1.2 综合性学科搜索引擎

(1) Google Scholar于2004年开始运营,是免费的学术文献检索系统。其利用爬虫(crawlers)在各合作者的资料库和网络中漫游获取文献资料,包含①许多大型且知名的出版者和大学机构(如美国电子电机工程师学会等)的资料库;②数位主机(如HighWire Press、MetaPress、Ingenta等);③政府及其他学术组织和政府基金会(如美国物理学会、国家卫生研究院等);④预印、重印服务者(如arXiv.org、Astrophysics Data System、RePEc及CiteBase等)^[13]。此外,Google Scholar也欢迎个人用户将其作品及个人资料库加入其爬虫清单。Google Scholar收录资料类型包含同行评审论文、学位论文、图书、预印本、摘要、技术报告、电子刊物等;按照出版物来源划分,包含学术出版方、专业学会、预印本寄存出版商、大学机构及网络资源等;按照提供管道划分,包含出版社收费取用、开放取阅信息及图书馆登入取用信息等^[14]。

(2) Microsoft Academic Search是微软亚洲研究院于2009年开发的开放获取学术搜索引擎系统。自2012年不更新系统收录的内容^[15],其提供超过1.5亿篇文献,及数十亿学术图、表与文献的链接,以知识概念和语义网作为其检索系统的基础,改进系统检索方式,并强化与外部全文的连结。因此,Microsoft Academic Search的学术检索除提供相关学术文献外,还在页面显示相关研究者、研究机构、会议、期刊等学术信息,以构建学术网络社群,推动学术传播^[16]。另外,与传统搜索引擎系统相比,其开发时参考H指数排名,因此系统能查找到不同领域内H指数较高的科学家或期刊文献。

4.1.3 汇集式机构典藏系统

(1) 开放存取知识库名录(the Directory of Open Access Repositories, OpenDOAR)旨在利用全球性开放获取系统进行数据搜集、组织及传递,提高开放获取资源的取用效益,推动开放获取的发展。OpenDOAR受开放社会研究所、英国联合信息系统委员会、英国大学学术图书馆联盟、欧洲学术出版与学术资源联盟资助,由英国诺丁汉大学和瑞典兰德大学于2005年共同创办。此外,OpenDOAR和开放获取期刊指南

(Directory of Open Access Journals, DOAJ)并列为当前开放获取学术资源检索的主要平台,其数据类型包含期刊论文、会议论文、学位论文、技术报告、专利数据、学习资源、多媒体资源、数据集、工作手稿及数据、预印本等,涵盖29个学科领域^[17]。截至2016年6月,有超过3 000家研究机构参与该系统建设,包含3 099个数据库,其中有2 621个数据库为机构典藏数据库,并以美国、英国、日本及德国的机构典藏数据库为主,四者共占总数35.6%^[18]。

(2) OA Ister为美国密歇根大学图书馆于2002年建立的汇集式机构典藏资源联合目录系统。主要提供数字资源搜寻服务,其典藏虽非自身产出的数字内容,但目标是建立开放获取的数字资源,使散布在网络的学术性数字资源被方便地检索使用。为扩大系统使用率及曝光度,2009年10月,密歇根大学将OA Ister资料移至Online Computer Library Center,使全球使用者可通过WorldCat.org检索OA Ister的资料。OA Ister收录资料类型包括图书、学位论文、期刊论文、研究报告、报纸、手稿、声音文件、图片、图像文件、统计资料等^[19]。OA Ister包含全球1 500多所机构提供的数字资源,数据超过3 000万篇(部分电子全文)。

综上所述,商业引文数据库的成立宗旨是为特定用户建立,而开放获取引文系统是为全球用户建立;除Google Scholar无明确收录范围,其他数据库及系统皆有明确收录范围;学科主题皆为综合性学科;从时间看,Web of Science网络版建立最早(1997年),Microsoft Academic Search建立最晚(2009年);各商业引文数据库及开放获取引文系统皆为全球性,在语言上均为多种语言,并以英文为主。

4.2 商业引文数据库与开放获取引文系统完整性分析

2001—2012年诺贝尔经济学奖得主在期刊发表文献1 888篇,在商业引文数据库Web of Science中检索到1 034篇文献,占总量的54.77%;Scopus可检索到1 117篇文献,占总量的59.16%;可见相较Web of Science,Scopus资源拥有量更多。综合性学科搜索引擎系统Google Scholar、Microsoft Academic Search及开放获取引文系统OA Ister、OpenDOAR所检索到的文献数占诺贝尔经济学奖得主发表文献总数的比例分别为94.86%、93.59%、45.13%和28.81%(见表1)。可见,搜

索引的期刊文献数据最丰富, 其次是商业引文数据库, 再次是汇集式机构典藏系统。

表 1 商业引文数据库及开放获取引文系统
期刊文献完整性比较

商业引文数据库及 开放获取引文系统名称	含诺贝尔经济学奖者 发表文献数量/篇	比 例
Web of Science	1 034	54.77%
Scopus	1 117	59.16%
Google Scholar	1 791	94.86%
Microsoft Academic Search	1 767	93.59%
OAIster	852	45.13%
OpenDOAR	544	28.81%

4.3 商业引文数据库与开放获取引文系统重复性分析

4.3.1 Web of Science

Web of Science和Scopus检索到重复期刊文献714篇, 重复文献量占Web of Science含诺贝尔经济学奖者发表文献总量的69%, 占Scopus的64%。Web of Science与Google Scholar、Microsoft Academic Search、OAIster、OpenDOAR的交叉比较及重复情况, 具体见表2。其中, Web of Science与Google Scholar检索到重复的期刊文献数量最多, 与OpenDOAR重复的期刊文献数量最少; Web of Science在收录期刊数据重复性的表现不及Scopus、Google Scholar、Microsoft Academic Search, 但优于OAIster和OpenDOAR。

表 2 Web of Science与其他数据库及系统重复性比较

	Web of Science		
	重复文献 数量/篇	重复文献占各数 据库及系统比例	重复文献占Web of Science比例
Scopus	714	64%	69%
Google Scholar	1 014	57%	98%
Microsoft Academic Search	995	56%	96%
OAIster	521	61%	50%
OpenDOAR	325	60%	31%

4.3.2 Scopus

Scopus和Google Scholar检索到重复期刊文献1 085篇, 重复文献量占Scopus含诺贝尔经济学奖者发表文献总量的97%, 占Google Scholar的61%。Scopus与Web of Science、Microsoft Academic Search、OAIster、OpenDOAR的交叉比较及重复情况, 具体见表3。其中, Scopus与Google Scholar检索到重复的期刊文献数量最多, 与OpenDOAR重复的期刊文献数量最少; Scopus在收录期刊文献数据重复性的表现远不及Google Scholar、Microsoft Academic Search, 但优于Web of Science、OAIster和OpenDOAR。Web of Science、OAIster和OpenDOAR与Scopus的期刊文献重复率均不足70%, 说明其表现虽较差但仍有其重要性。

4.3.3 Google Scholar

Google Scholar和Microsoft Academic Search检索到重复期刊文献数为1 694篇, 重复文献量占Google Scholar含诺贝尔经济学奖者发表文献总量的95%, 占Microsoft Academic Search的96%。Google Scholar与Scopus、Web of Science、OAIster、OpenDOAR的交叉比较及重复情况, 具体见表4。其中, Google Scholar与Microsoft Academic Search检索到重复的期刊文献数量最多, 二者收录期刊文献数据重复性相近; 与OpenDOAR重复的期刊文献数量最少, 整体而言, Google Scholar的重复性表现优于其他检索数据库及系统。因此, 在检索文献数据时, 几乎可以使用Google Scholar系统替代其他数据库及系统。

表 3 Scopus与其他数据库及系统重复性比较

	Scopus		
	重复文献 数量/篇	重复文献占各数 据库及系统比例	重复文献占 Scopus比例
Web of Science	714	69%	64%
Google Scholar	1 085	61%	97%
Microsoft Academic Search	1 034	59%	93%
OAIster	512	60%	46%
OpenDOAR	353	65%	32%

表 4 Google Scholar与其他数据库及系统重复性比较

	Google Scholar		
	重复文献数量/篇	重复文献占各数据库及系统比例	重复文献占Google Scholar比例
Web of Science	1 014	98%	57%
Scopus	1 085	97%	61%
Microsoft Academic Search	1 694	96%	95%
OAster	823	97%	46%
OpenDOAR	527	97%	29%

4.3.4 Microsoft Academic Search

从表5可见, Microsoft Academic Search与Web of Science、Scopus、Google Scholar的重复文献量占Microsoft Academic Search含诺贝尔经济学奖者发表文献总量的比例与重复文献量占各数据库及系统的比例相比, Microsoft Academic Search的重复性远低于Web of Science与Scopus, 仅超出Google Scholar 1%。

Microsoft Academic Search和OAster重复文献量占Microsoft Academic Search含诺贝尔经济学奖者发表文献总量的46%, 占OAster的96%; Microsoft Academic Search和OpenDOAR重复文献量占Microsoft Academic Search含诺贝尔经济学奖者发表文献总量的30%, 占OpenDOAR的96%。

虽然Microsoft Academic Search与Google Scholar重复性相近, 但优于Scopus、Web of Science、OAster及OpenDOAR数据库及系统。

表 5 Microsoft Academic Search与其他数据库及系统重复性比较

	Microsoft Academic Search		
	重复文献数量/篇	重复文献占各数据库及系统比例	重复文献占Microsoft Academic Search比例
Web of Science	995	96%	56%
Scopus	1 034	93%	59%
Google Scholar	1 694	95%	96%
OAster	814	96%	46%
OpenDOAR	522	96%	30%

4.3.5 OAster

OAster和OpenDOAR检索到重复期刊文献331篇, 重复文献量占OAster含诺贝尔经济学奖者发表文献总量的39%, 占OpenDOAR的61%。OAster与Web of Science、Scopus、Microsoft Academic Search、Google Scholar的交叉比较及重复情况, 具体见表6。

表 6 OAster与其他数据库及系统重复性比较

	OAster		
	重复文献数量/篇	重复文献占各数据库及系统比例	重复文献占OAster比例
Web of Science	521	50%	61%
Scopus	512	46%	60%
Google Scholar	823	46%	97%
Microsoft Academic Search	814	46%	96%
OpenDOAR	331	61%	39%

其中, OAster收录期刊数据重复性不佳, 只优于OpenDOAR系统, 而OpenDOAR整体表现均不及其他数据库及系统。

通过期刊文献数据重复性比较可见, 搜索引擎优于商业引文数据库, 商业引文数据库优于汇集式机构典藏系统。搜索引擎Google Scholar优于Microsoft Academic Search; 商业引文数据库Scopus优于Web of Science; 汇集式机构典藏系统OAster优于OpenDOAR。Google Scholar对Microsoft Academic Search的重复率最高, 近100%, 推测原因为Google Scholar及Microsoft Academic Search皆为搜索引擎, 互相抓取的数据范围基本相同。

4.4 商业引文数据库与开放获取引文系统独特性分析

从表7可见, 与Scopus相比, Web of Science独有的期刊文献量占Web of Science含诺贝尔经济学奖者发表文献总量的31%, Scopus独有的文献占Scopus含诺贝尔经济学奖者发表文献总量的36%, 二者独特性相差不大。各数据库及系统两两独特性交叉比较, Web of Science期刊文献独特性与Google Scholar相比, Web of Science独特性为2%, Google Scholar独

特性为43%，可见Web of Science独有的期刊文献资源不及Google Scholar。此外，Web of Science期刊文献独特性也远不及Microsoft Academic Search，但优于OA Ister和OpenDOAR。

Scopus与Google Scholar相比，Scopus独有的期刊文献量为32篇，占Scopus含诺贝尔经济学奖者发表文献总量的3%；Google Scholar独有的期刊文献量为706篇，占Google Scholar含诺贝尔经济学奖者发表文献总量的39%；Scopus独有的期刊文献资源不及Google Scholar。此外，Scopus期刊文献独特性不及Microsoft Academic Search，但优于OA Ister和OpenDOAR。

Google Scholar与Microsoft Academic Search期刊文献独特性相近，均优于Web of Science、Scopus、OA Ister、OpenDOAR。汇集式机构典藏系统OA Ister和OpenDOAR独特性不及其他4个数据库及系统；二者相比，OA Ister期刊文献独特性略优于OpenDOAR，推测原因为OpenDOAR所拥有数据量最少，故对应其独特性最低。

综上所述，本研究中的商业引文数据库及开放获取引文系统期刊文献独特性优劣顺序为Google Scholar、Microsoft Academic Search、Scopus、Web of Science、OA Ister、OpenDOAR。

表7 商业引文数据库及开放获取引文系统期刊文献独特性比较

	Web of Science	Scopus	Google Scholar	Microsoft Academic Search	OA Ister	OpenDOAR
Web of Science	-	31%	2%	4%	50%	69%
Scopus	36%	-	3%	7%	54%	68%
Google Scholar	43%	39%	-	5%	54%	71%
Microsoft Academic Search	43%	41%	4%	-	54%	70%
OA Ister	39%	40%	3%	4%	-	61%
OpenDOAR	40%	35%	3%	4%	39%	-

5 结论与建议

本研究主要探讨各数据库及系统特性与现况，归纳其优、缺点，进而分析各数据库及系统的完整性。此外，还进一步交叉对比两两数据库及系统的重复性及独特性。

(1) 搜索引擎数据完整性优于商业引文数据库，商业引文数据库优于汇集式机构典藏系统。搜索引擎Google Scholar与Microsoft Academic Search的数据相对完整；Google Scholar因包含各种子系统，所以可检索到多元数据，在研究中，Google Scholar最常链接到的子系统为Google Book及Google Search；Microsoft Academic Search收录数据也非常完整，二者完整性约95%与94%。商业引文数据库完整性次之，两个数据库与搜索引擎数据量相差均超过34%。汇集式机构典藏系统完整性表现最差。各数据库及系统完整性由高到低的顺序为Google Scholar、Microsoft Academic Search、Scopus、Web of Science、OA Ister、OpenDOAR。

(2) 各数据库及系统两两对比重复率后发现，

Google Scholar与其他各数据库及系统重复文献量占Google Scholar含诺贝尔经济学奖者发表文献总量的比例均低于重复文献量占各数据库及系统的比例，其重复性最低。其次是Microsoft Academic Search；Scopus的重复性低于Web of Science、OA Ister与OpenDOAR，却高于Google Scholar、Microsoft Academic Search；Web of Science的重复性低于OA Ister与OpenDOAR，却高于Scopus、Google Scholar与Microsoft Academic Search；汇集式机构典藏系统OA Ister重复性只低于OpenDOAR，却远高于其他系统与数据库，即OpenDOAR的重复性是各数据库及系统中最高的。

(3) 各数据库及系统两两对比独特性后发现，综合性学科搜索引擎Google Scholar与Microsoft Academic Search独特性优于商业引文数据库Web of Science与Scopus；汇集式机构典藏系统OA Ister与OpenDOAR独特性较差，其中OA Ister独特性略优于OpenDOAR；商业引文数据库的独特性则介于搜索引擎与汇集式机构典藏系统之间，独特性低于搜索引擎36%—41%，却超过汇集式机构典藏系统11%—33%；

其中Scopus独特性又优于Web of Science。换言之,商业引文数据库质量不及搜索引擎却优于汇集式机构典藏系统。

参考文献

- [1] YERKEY N,GLOGOWSKI M.Scatter of library and information science topics among bibliographic data bases[J].Journal of American Society for Information Science,1990,41(4):245-253.
- [2] MARTYN J.Tests on abstracts journals:coverage overlap and indexing[J].Journal of Documentation,2013,23(1):45-70.
- [3] WOOD J L,FLANAGAN C,KENNEDY H E.Overlap in the list of journals monitored by BIOSIS,CAS,EI[J].Journal of the American Society for Information Science,1972,23(1):36-38.
- [4] READ E,SMITH C.Searching for library and information science literature:a comparison of coverage in three database[J].Library Computing, 2000,19(1):118-126.
- [5] WALTERS W H,WILDER E I.Bibliographic index coverage of a multidisciplinary field[J].Journal of the American Society for Information Science and Technology, 2003,54(14):1305-1312.
- [6] RATHER R A,LONE F A,SHAH G J.Overlap in web search results:a study of five search engines[J].Library Philosophy and Practice,2008(12):1-6.
- [7] ESMAEIL S M,KIAIE R M,KETAB F.A comparison between search engines and meta-search engines in retrieving information related to physics and the extent of their overlap[J].Library and Information Studies, 2011,22(3):130-140.
- [8] WANG L,WANG J,WANG M,et al.Using Internet search engines to obtain medical information:a comparative study[J].Journal of Medical Internet Research,2012,14(3):74.
- [9] Nobel Prize.All Nobel Prizes in economic sciences[EB/OL]. http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates.
- [10] RePEc. RePEcservices[EB/OL].[2016-08-01].<http://repec.org>.
- [11] Web of Science.Introduction to web of science[EB/OL].[2016-08-01].<http://concert.stpi.narl.org.tw/fdb/Web of Science /index.html>.
- [12] Scopus.Scopus in Detail: What does it cover?[EB/OL].[2016-08-01].<http://info.scopus.com/detail/what>.
- [13] O'LEARY M.Google Scholar:what's in it for you?[J].Information Today,2005,22(7):35-39.
- [14] Google Scholar[EB/OL].[2016-08-01].<https://scholar.google.com.tw>.
- [15] Microsoft.AcademicSearch[EB/OL].[2016-08-01].<http://academic.research.microsoft.com/default.aspx>.
- [16] Microsoft Academic.Research more, search less[EB/OL].[2016-08-01].<http://academic.research.microsoft.com>.
- [17] 何琳.OpenDOAR和机构知识库发展现状[J].图书馆工作与研究, 2009(2):30-33.
- [18] OpenDOAR.The directory of open access repositories[EB/OL].[2016-08-01].<http://www.opendoar.org>
- [19] OCLC.The OAster® database[EB/OL].[2016-08-01].<http://www.oclc.org/oaister.en.html>.

作者简介

蔡明月,女,1956年生,台湾政治大学图书资讯与档案学研究所特聘教授,兼图书馆馆长暨社会科学资料中心主任,研究方向:信息计量学、网络计量学、图书情报学、用户研究,E-mail: mytsay@nccu.edu.tw。

A Comparison Study on Citation Index: Commercial Databases and Open Access Systems

CAI MingYue

(Graduate Institute of Library, Information and Archival Studies National ChengChi University, Taipei 116, China)

Abstract: In this study, scholarly communication system of commercial services and open access will be examined through user interface and comprehensiveness of coverage for commercial citation index databases (Web of Science and Scopus) and open access citation system (Google Scholar, Microsoft Academic Search, OAster and OpenDOAR). Retrievals will be conducted in the six citation index databases and systems stated above to analyze and compare their retrieval interfaces, and evaluations of each system will be made as well according to the output of retrieval results. Noble laureates in economic sciences from 2001 to 2012 are selected as samples in this study. Bibliographic records of their publications will be retrieved and downloaded from each system, and a computer program will be developed to perform the analytical tasks of sorting, comparison, elimination, aggregation and statistics. Bibliographic records retrieved from the six databases and systems will undertake quantitative analyses and cross references to determine the comprehensiveness of their system coverage. The results of the study may provide better references for libraries to acquire citation index databases, to build institutional repositories, or to create citation index systems on their own in the future. Suggestions on indicators and tools for academic assessment will be presented based on the comprehensiveness assessment of each system as well.

Keywords: Citation Index Database; Open Access System; Database Coverage; Database Overlap

(收稿日期: 2016-08-19)