

英美机构知识库的比较

周永红 吴振寰

(湘潭大学公共管理学院, 湖南湘潭 411105)

摘要: 文章针对OpenDOAR网站所登记注册的英美机构知识库进行调查, 分析英美机构知识库发展概况, 并主要从机构知识库的主要类型及创建主体、创建软件应用情况、知识库容量大小、资源收录类型、学科分布、语种、开放存取情况等方面进行比较, 总结两国的主要相同点和差异, 为中国机构知识库建设提供相应的启示。

关键词: 英国; 美国; 知识库; 机构知识库; 比较

中图分类号: G250

文献标识码: A

DOI: 10.3772/j.issn.1674-1544.2014.03.015

Comparison on the Institutional Repositories between United Kingdom and United States

Zhou Yonghong, Wu Zhenhuan

(Institute of Public Administration, Xiangtan University, Xiangtan 411105)

Abstract: Based on the investigation of the institutional repositories development status between United Kingdom and United States from OpenDOAR, this paper firstly analyses their outlines, then it compares the institutional repositories from the types, the softwares, the sizes, the digital resources kinds, the subject, the languages, the open access, and so on. Their main sameness and difference are summarized to provide some enlightenments to construct the institutional repositories in China.

Key Words: United Kingdom, United States, repositories, institutional repositories, comparation

英国和美国拥有发达的经济、先进的科学技术和频繁的学术交流, 两国都是目前世界上机构知识库建设得最好的国家, 无论是机构知识库的数量还是机构知识库的质量方面, 都领先世界, 具有一定的可比性。著名的开放存取知识库目录网站OpenDOAR (Directory of Open Access Repositories) 提供的数据显示, 截止到2013年9月6日, OpenDOAR收录的机构知识库总数为

2401个, 英国的机构知识库数量为218个, 约占总数的9.1%, 居世界第二; 美国机构知识库数量为408个, 约占总数的17%, 居世界第一^[1]。再根据2013年9月6日西班牙赛博计量实验室 (Cybermetrics Lab) 对全球机构知识库排名显示, 排名前100机构知识库中, 英国有4个; 美国有23个, 居世界第一^[2]。

本文根据2013年9月6日OpenDOAR网站登

作者简介: 周永红* (1974-), 女, 博士, 湘潭大学公共管理学院副教授, 研究方向: 信息资源共享与数字化信息服务; 吴振寰 (1988-), 女, 博士, 湘潭大学公共管理学院硕士研究生, 研究方向: 数字化信息服务。

基金项目: 国家社科基金项目“基于企业联盟的知识共享研究”(08CTQ007); 湖南省图书馆学会课题“免费开放时代图书馆的服务体系及其演化”。

收稿日期: 2013年12月15日。

记注册的机构知识库的相关数据（注：文章中的相关数据，没有特别说明的，都来源于本网站），对英美机构知识库的发展进行比较分析。

1 英美机构知识库发展概况

英国较早进行机构知识库的建设。据 OpenDOAR 网站的数据显示，截止到 2005 年 1 月 12 日，英国已有 25 个机构知识库在该网站上注册登记。从 2005 年中期到 2011 年，英国机构知识库数量增长非常迅速。如截止到 2007 年 5 月 1 日时，网站上已有 100 个注册登记的机构知识库；到 2011 年 6 月 1 日，网站上已有 200 个注册登记的机构知识库。英国机构知识库数量的急剧增长主要由于英国相关政府部门和高校对机构知识库建设的重视。如英国高等教育团体曾提出 2006 年 7 月底建立开放存取型大学机构知识库的计划目标。英国的信息系统联合委员会（JISC）为建立一个国家知识库服务的基础设施或框架，曾委托进行了一项英国知识库技术和组织模型方面的研究，鼓励所有的高等教育机构建立知识库，并且提出要制订各种政策来保证内容的收藏，这与英国高等教育团体的开放获取知识库计划目标是一致的，且该研究同时也提出作为国家整个知识库服务计划的一部分，还要给那些私人企业发展的机会^[3]。针对高校机构知识库的建设目标，Hull 大学在 JISC 知识库项目的资助下，曾设立

知识库、元数据和管理项目，主要目标就是要建立一个标准的、灵活的工作流程管理工具，使用户能够通过该工具实现与 Fedora 知识库平台进行交互，并能够自动生成元数据^[4]。2011 年至 2013 年 9 月，英国机构知识库数量增长的速度趋于平缓。2011 年 6 月后的 2 年多时间里，OpenDOAR 网站上登记注册的机构知识库只增加了 18 个，增长速度也有一定起伏。因为这一时期英国除了注重机构知识库数量的增长外，更注重机构知识库质量和服务的提升。英国机构知识库数量增长的总体情况如图 1 所示。

2001 年美国就开始建立相应机构知识库——俄亥俄州立大学知识库。相对于英国，美国机构知识库的发展更加迅速。如据 OpenDOAR 网站的数据显示，截止到 2005 年 1 月 12 日，在该网站上注册登记的美国机构知识库是 38 个。2006 年是美国机构知识库数量增长最快的一年，如 2006 年 4 月 1 日注册登记的美国机构知识库是 100 个，短短 4 个多月的时间里，到 2006 年 8 月 11 日时就已增加到 211 个。这一时期美国机构知识库数量的急剧增长与相关计划的推动是密切相关的。如 2006 年美国博物馆及图书馆服务事业局曾启动了“构建机构知识库的协作学习环境”项目（MIRACLE），以发放问卷方式对美国学术机构进行了调查，形成了《MIRACLE 项目的研究发现》报告。报告指出共发放问卷 2147 份，回

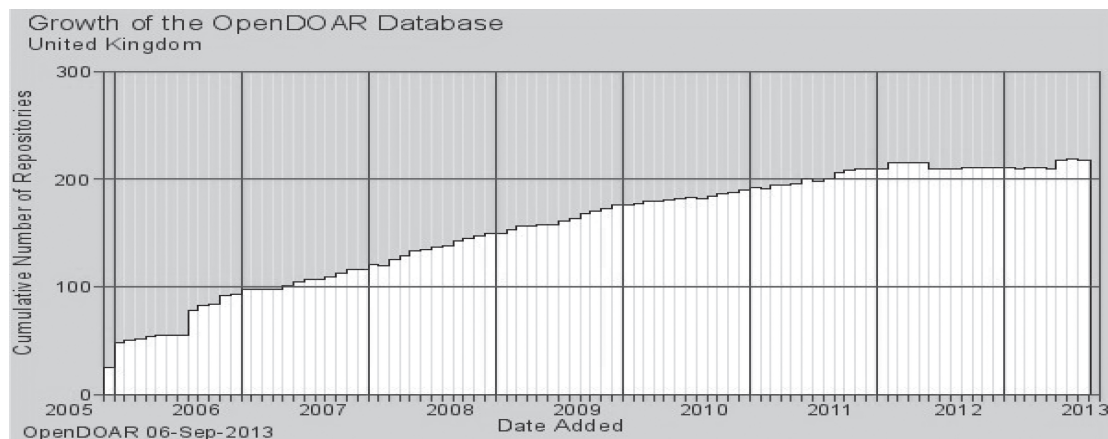


图 1 英国机构知识库数量增长情况

收问卷446份。在回收的446份问卷中,显示有48家已经运行和使用了机构知识库,有70家在积极地计划并尝试机构知识库建设,有90家才刚开始计划机构知识库建设,有236家约占60%的机构还没有相关计划;报告也总结指出根据高等教育机构卡内基分类表分析,大多数研究性大学已经计划、试验或者运行机构知识库了,而回答“没计划”或者“计划中”的单位多数是硕士或本科院校^[5]。2006年8月到2008年3月期间,美国机构知识库增长速度有所减慢,在一年半多的时间里,增加的数量还不到100个,如到2008年3月1日时,注册登记的美机构知识库是301个。2008年3月到2011年7月,美国机构知识库的总体数量有所增加,但增速不明显,这3年多的时间里增加才近100个,达到了400个。2011年7月至2012年4月期间,登记注册的机构知识库有所增长;但令人不解的是2012年4月后,登记注册的机构知识库数量还有所减少,具体原因不清楚。从2012年4月到2013年9月的这段时间里,登记注册的机构知识库数量基本维持在408个左右。美国机构知识库数量增长的总体情况如图2所示。

2 英美机构知识库的比较

2.1 知识库的主要类型及创建主体

OpenDOAR网站将知识库一共分为机构型、学科型、聚合型和政府型4种类型。其中,机构知识库的创建主体多为高校和科研机构;学科型知识库的创建主体包括高校、图书馆、政府组织等,提供专业的学科知识;聚合型知识库的创建主体包括高校、图书馆、政府和科研机构等,提供某一或者某些方面的综合性知识;政府型知识库创建主体为政府,主要提供政务方面的相关知识。

目前,在OpenDOAR上登记注册的知识库绝大多数是机构知识库,其中注册的2401个知识库中,机构知识库就有1982个,约占总数的82.5%。英国和美国的知识库类型(表1)也是以机构型为主,学科型和聚合型次之,政府型最少。英国和美国拥有世界一流的高校及科研机构,学术成果也较多,这些都为机构知识库的创建提供了有利条件。除高校及科研机构等学术性机构外,英美的公共图书馆和政府等已开始重视和推进机构知识库的建设,有些还成立了机构知

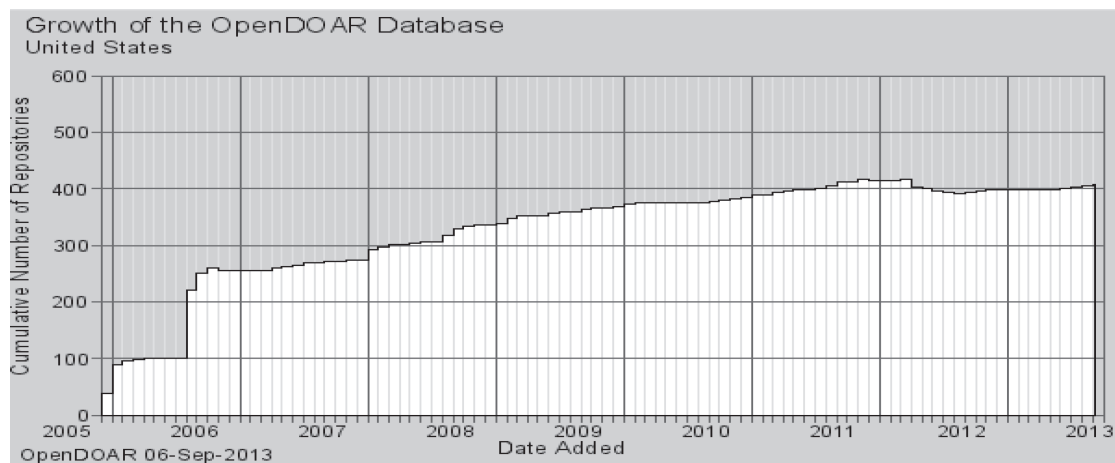


图2 美国机构知识库数量增长情况

表1 英美机构知识库类型

单位:个

国家	机构型	学科型	聚合型	政府型	总数
英国	170	41	4	3	218
美国	291	81	24	12	408

识库联盟，如 JISC RepositoryNet 联盟等^[6]。

2.2 应用的软件

根据 OpenDOAR 网站的统计，目前世界上机构知识库所应用的软件种类已达 100 多种，但应用得较多的主要是 DSpace、E-prints、Digital Commons 等软件。英国和美国在机构知识库的创建过程中虽然都应用了多种软件，但软件的选择和应用方面还是有明显的差异，英美机构知识库创建软件应用排名的前 5 名如表 2 所示。

英国机构知识库应用的创建软件主要有 E-prints、DSpace、Open Repository 等，其中英国有近一半的机构知识库应用了 E-prints 软件，使用率达到了 48.6%。E-prints 是由英国南安普顿大学开发的一款适用于机构知识库的免费开源软件。DSpace 软件在英国的应用排名居第二，如英国剑桥大学机构知识库（DSpace@Cambridge）就应用了 DSpace 软件。

美国机构知识库创建软件的应用种类比英国多，主要有 DSpace、Digital Commons、CONTENTdm、E-prints 等。其中，由美国麻省理工学院图书馆（MIT Libraries）和美国惠普公司实验室（Hewlett-Packard Labs）合作开发的 DSpace 软件应用率在美国排名第一。但美国 DSpace 软件的

应用没有英国 E-prints 软件集中，如从 DSpace 和 Digital Commons 软件的应用比率来看相差不大，分别为 24.8% 和 21.3%。

2.3 知识库的容量

机构知识库容量大小是反映机构知识库建设质量的一个重要衡量指标。在 OpenDOAR 网站上，主要以资源收录条目数来反映机构知识库容量。为了便于比较，将英美两国机构知识库分别按其收录资源的内容条目数进行排名，都选择排名为前 10 的机构知识库（表 3）。

从所选择的机构知识库资源条目数来看，虽然美国的机构知识库总体容量要略胜于英国，但总体上相差不大，两国排名前 10 名的机构知识库收录资源的总条目数都超过了 3000 多万条，其中每个机构知识库的资源条目数都超过了 10 多万条，而且两国排名为第一的机构知识库的资源条目数都超过了 1000 多万条。

2.4 资源收录类型

目前，OpenDOAR 将机构知识库收录的资源分为期刊论文、多媒体和视听资料、学位论文、未出版报告和工作草稿、会议和讨论资料、其他的特殊文档、图书与部分章节、学习资料、书目参考资料、数据集、软件、专利等 12 类。

表 2 英美机构知识库的软件应用情况

国家	1	2	其他
英国	E-prints	DSpace	Open Repository、Unknow
美国	DSpace	Digital Commons	CONTENTdm、Eprints、Unknow

表 3 英美机构知识库按容量大小排名的前 10 名

排名	英国		美国	
	知识库名称	资源条数	知识库名称	资源条数
1	Europe PubMed Central	17415153	Research Papers in Economics	14000000
2	English Heritage ViewFinder	8000000	Internet Archive	8585416
3	Geograph British Isles	3682907	PubMed Central	2800000
4	PhilPapers	563834	University of Michigan Library Repository	2321577
5	UCL Discovery	297139	Cross Collection Discovery	1759467
6	DSpace @ Cambridge	193480	Education Resources Information Center	1397164
7	ESRC Research Catalogue	181592	Marshall Technical Reports Server	986780
8	Manchester eScholar Services	169916	NASA Technical Reports Server	986780
9	Oxford University Research Archive	116593	Social Science Research Network	930816
10	Visual Arts Data Service	111997	arXiv.org e-Print Archive	885327
总和	30732611		34653327	

对于OpenDOAR所列的12类资源,英国和美国的机构知识库都有所收录,且都以期刊论文为主,如英国的218个机构知识库中就有150多个以期刊论文为主,美国的408个机构知识库中就有200余个以期刊论文为主。除了期刊论文以外,会议和讨论资料、学位论文、未出版报告和工作草稿等灰色文献都是两国机构知识库收录的主要资源类型(表4)。

2.5 学科分布

OpenDOAR网站将机构知识库分为29个学科类型,包括多学科、自然科学、农业和食品、生物学、化学和化工、地球科学、生态和环境、数学和统计、物理和几何、健康和医疗、一般技术、建筑、民生工程、计算机和IT、电子工程、机械工程、艺术和人文学、实用技艺、地理和地区研究、历史和考古、语言和文学、哲学和伦理、社会科学、商业与经济、教育、法律和政治、图书馆和信息科学、管理和运筹学、心理学等类型。

目前,在OpenDOAR网站登记注册60%以上的机构知识库都是多学科型的。从英国和美国机构知识库的学科分布来看,OpenDOAR所列的29个学科范围都有所分布,且都有一半以上是多学科型的。如英国的多学科型知识库有120多个,美国则超过了200个。除了主要的多学科型机构知识库外,英美两国的机构知识库都以人文

科学和社会科学型的居多,如两国都主要涉及了历史和考古、健康和医疗等方面的学科(表5)。

2.6 语种

机构知识库所收录资源涉及的语种多少是衡量机构知识库资源收录完备程度的一个重要指标,也是影响机构知识库利用情况的一个重要因素。根据OpenDOAR的统计,目前世界机构知识库一共收录了64种语言类型的资源,其中英语资源超过了一半,排名前三种的英语、德语和西班牙语资源,所占比例分别约为总数的71%、12%和8%。

英国和美国几乎所有的机构知识库都以收集英语资源为主,这是因为英语为英国和美国的母语,资源本身要多一些。除了英语之外,法语、西班牙语、德语资源都是两国机构知识库收录的重点(表6)。

2.7 开放存取情况

开放存取是机构知识库的一个重要特征。开放存取带动了机构知识库的兴起和发展,使其成为了开放存取的有效实现方式^[7]。OpenDOAR根据机构知识库的开放存取情况,将已注册登记的机构知识库分为4种,即具有全部开放存取功能的可操作型知识库(Operational repository)、试验型知识库(Trial repository)、由于技术故障无法开放存取的中断型知识库(Broken repository)、不提供开放存取服务的封闭型知识库(Closed

表4 英美机构知识库排名前5名的资源类型

国家	1	2	3	4	5
英国	期刊论文	会议/讨论资料	未出版报告/工作草稿	图书/部分章节	学位论文
美国	期刊论文	多媒体/视听资料	学位论文	未出版报告/工作草稿	会议/讨论资料

表5 英美机构知识库按学科排名的前5名

国家	1	2	3	4	5
英国	多学科	计算机/IT	历史/考古	健康/医疗	教育
美国	多学科	历史/考古	法律/政治	健康/医疗	地理/地区研究

表6 英美机构知识库排名前5名的资源语种

国家	1	2	3	4	5
英国	英语	法语	西班牙语	威尔士语	德语
美国	英语	西班牙语	法语	汉语	德语

表7 英美机构知识库的开放存取情况

单位:个

国家	可操作型	试验型	中断型	封闭型	总数
英国	198	9	4	7	218
美国	396	6	3	3	408

repository)。

目前,英美在OpenDOAR网站上登记注册的机构知识库90%以上都具有全部的开放存取功能,是可操作型知识库;但两国目前都也有一些试验型、中断型、封闭型知识库,说明机构知识库的开放存取功能还在进一步完善(表7)。

3 结语

英国和美国的机构知识库,除了绝对数量和创建软件等方面有一些区别外,主要的创建主体都是高校和科研机构,容量大小总体上相差也不大。两国机构知识库都收录了OpenDOAR所列的12类资源,且都以期刊论文为主,会议和讨论资料、学位论文、未出版报告和工作草稿等灰色文献也是收录的主要资源类型;两国的机构知识库有一半以上是多学科类型的,且以人文科学和社会科学的机构知识库居多;两国几乎所有的机构知识库都以收集英语资源为主,其次是法语、西班牙语、德语资源;两国机构知识库90%以上都能顺利地进行链接,能提供较好的开放存取服务,并通过政府强制开放存取方式和开放源代码软件方式等来解决其中的知识产权问题^[8]。

目前,我国学术机构正在加强机构知识库建设,如中国科学院目前已有100多家研究所建立了所级机构知识库,并以此为基础建立全院机构知识库集成服务网络平台。其中,可开放获取全文的成果已达到70%以上^[9]。但从目前的情况来看,我国机构知识库建设还存在建设模式选择、资源收集、数据加工、相关标准、政策支持、宣传与推广、开放获取与知识产权、实施平台和工具、联盟与合作、大数据与云环境下机构知识库的进一步发展和服务应用等方面的问题^[10-11]。为了能更好地解决这些问题,我国机构知识库应充分借鉴英美两国机构知识库建设与服务中的主要经验,加强对机构知识库的宣传,增加资源的种

类,扩展资源收集的语种范围,合理解决相应的知识产权问题,使更多的机构重视机构知识库的建设与服务,并促进机构知识库联盟的发展等。

参考文献

- [1] University of Nottingham. Directory of Open Access Repositories[EB/OL]. [2013-09-06]. <http://www.open-doar.org/index.html>.
- [2] Cybermetrics Lab. Ranking Web of Repositories [EB/OL]. [2013-09-06]. http://repositories.webometrics.info/en/top_Inst.
- [3] Alma Swan, Chris Awre. Linking UK Repositories: Technical and Organizational Models to Support User—Oriented Services Across Institutional and Other Digital Repositories[EB/OL]. [2013-10-12]. <http://www.jisc.ac.uk/media/documents/programmes/digitalrepositories/linkingukrepositoriesreport.pdf>.
- [4] Green Richard, polphin Ian, Awre Chris, et al. The RepoMMan Project: Automating Metadata and Workflow for Fedora[J]. OCLC Systems & Services, 2007(2): 210-215.
- [5] Karen Markey, Soo Young Rieh, Beth St Jean, et al. Census of Institutional Repositories in the United States[EB/OL]. [2013-10-12]. <http://www.clir.org/pubs/reports/pub140/pub140.pdf>.
- [6] 曾苏,马建霞,祝忠明,等.机构知识库联盟发展现状及关键问题分析[J].图书情报工作,2009,53(24):106-109.
- [7] 李枫林,赵雪芹,胡吉明,等.机构知识库;开放获取的有效实现形式[J].情报杂志,2007(6):115-117.
- [8] 柳菁.美国机构知识库版权问题的解决方式及启示[J].情报科学,2013(5):157-160.
- [9] 宋喜群,刘晓倩.中科院建成国内最大规模机构知识库群[N].光明日报,2013-09-25(6).
- [10] 马建霞.机构知识库内容建设与服务设计的趋势[J].情报理论与实践,2010(9):23-27.
- [11] 聂华,韦成府,崔海媛,等.CALIS 机构知识库:建设与推广、反思与展望[J].中国图书馆学报, 2013, 39(2): 46-52.