DOI:10.3969/j.issn.1671-3982.2017.06.004

·信息组织与信息服务·

中美高校图书馆机构知识库资源建设比较及启示

李　强

[摘要]以我国 C9 联盟高校和美国常青藤联盟高校为例,从资源的丰富性、易用性、规范性和可持续性 4 个维度对比分析了中美两国高校图书馆机构知识库资源建设现状,结果显示,如我国高校图书馆机构知识库的建设仍处于初级阶段,当前机构知识库资源建设现状不容乐观,资源组织的易用性和规范性不足,资源建设的可持续发展前景堪忧等。

[关键词]高校图书馆;机构知识库;资源建设;比较研究

[中图分类号]G253　　[文献标志码]A　　[文章编号]1671-3982(2017)06-0017-05

Comparison of resources development for institutional repository between college and university libraries in China and USA and its enlightenments

LI Qiang

(Library of Nantong University, Nantong 226019, Jiangsu Province, China)

[**Abstract**] The resources development for institutional repository between college and university libraries in China and USA was comparatively analyzed in aspects of its richness, usability, normality and sustainability with C9 League Colleges and Universities in China and Ivy League Colleges and Universities in USA as examples, which showed that the construction of institutional repository in domestic college and university library is still at its initial stage, its usability and normality are limited, and its sustainability is not so promising.

[**Key words**] College and university library; Institutional repository; Resources development; Comparative research

开放存取随着其资源的日益增长,正成为学术信息资源管理的新兴课题之一[1]。而机构知识库(Institutional Repository,IR)正是伴随着学术信息资源开放存取的兴起而兴起的,其具有协助教育科研机构知识资产保存和知识管理的重要作用[2]。相比于 2010 年前后全球机构知识库数量保持年均 30% 的增长率[3],相关统计显示,近两年全球范围内机构知识库的建设已经进入一个增长相对平缓的阶段[4]。在此背景下,对知识库资源的“精耕细作”则更加受到重视,相关研究也日益增多。有学者指出,资源建设作为机构知识库建设的核心问题,目前也是机构知识库建设的瓶颈所在[5]。而高校图书馆作为推动全球机构知识库建设的骨干力量,其资源建设的好坏直接关系到学术资源开放存取服务的推进。

1　国内外高校图书馆机构知识库资源建设概况

2002 年 11 月,美国麻省理工学院(MIT)和惠普公司(HP)共同开发的 DSpace 系统机构知识库面世,并作为开源软件向广大机构用户开放[6]。DSpace 具有成本低廉且易于配置的特点,它的出现极大地促进了机构知识库在全球高校图书馆中的普及与应用,也拉开了全球范围内大规模建设机构知识库的序幕。麻省理工学院图书馆也是全球范围内较早推进机构知识库建设的高等院校。高校图书馆机构知识库承担着保存机构知识资产和支持知识开放共享的重要职责,是新型学术信息交流体系和教育科研知识基础设施的有机组成部分。学术出版与学术资源联盟(Scholarly Publishing and Academic Resources Coalition, SPARC)的 Raym Crow[7] 认为,

[作者单位]南通大学图书馆文献建设部,江苏 南通　226019

[作者简介]李　强(1979-),男,江苏南京人,硕士,馆员,研究方向为资源建设。

高等院校的机构知识库是收集并保存单个或多个大学内的数字知识资产的知识库,是大学提供给其内部社群的一系列服务,用来管理和传播由此机构及其内部社群创立的数字资源。而美国网络信息联盟的常务董事 Clifford A. Lynch[8]认为,大学机构知识库是学校为师生员工所提供的一系列服务,其目的是用于管理和发布由本机构及其成员所创造的数字资源。

2005 年以来,我国台湾地区[9]、香港特区[10]和大陆[11]的高校在高校图书馆机构知识库建设方面进行了一些实践探索。为保证数据的统一性,笔者以国际通行的机构知识库注册平台 Open DOAR[12](The Directory of Open Access Repositories)、ROAR[13](Registry of Open Access Repositories)和 Ranking Web of Repositories[14](The Spanish National Research Council)中的各项数据为基础,对中美两国高校机构知识库建设现状进行调研和分析。截至 2017 年 2 月 20 日,我国在以上国际注册平台中注册的机构知识库共 41 个(包括港澳台地区),来自高校图书馆的机构知识库有 11 个,其中大陆地区仅有北京大学、清华大学、厦门大学、西安交通大学、北京理工大学、浙江大学和广西民族大学 7 所高校明确推出了独立的机构知识库服务。美国共登记注册机构知识库 477 个,其中来自高校的机构知识库有 255 个,所占比例超过 50%。相比之下,我国高校图书馆的机构知识库无论在绝对数量上,还是在相对数量上,都与美国有着较大差距。

截至 2015 年,我国高等学校共 2 845 所,美国共有大专院校 4 180 家左右[15]。需要注意的是,我国高校机构知识库建设尚处于初级阶段,相当数量的机构知识库虽然已经投入使用,但由于语言等各种原因尚未在国际注册网站中注册,如中国人民大学图书馆机构知识库、大连理工大学机构仓储和上海交通大学机构学术信息库等。2016 年 9 月,由北京大学图书馆倡导,北京理工大学、北京师范大学、北京邮电大学、重庆大学等 17 所高校图书馆共同发起了“中国高校机构知识库联盟”(CHAIR)。但截至目前,大部分联盟高校尚未推出独立的机构知识库建设计划。

2 我国 C9 联盟高校与美国常青藤高校的机构知识库资源建设比较

为了深入比较中美两国高校图书馆机构知识库资源现状的现状,本文选择我国 C9 联盟和美国常青藤联盟(Ivy League)两大顶尖高校联盟进行深度比较。2016 年 10 月至 2017 年 2 月,通过网络访问的方式对以上两大联盟高校的机构知识库情况进行了在线访问和跟踪(表 1)。我国 C9 高校中,有 5 所开通了机构知识库;常青藤联盟高校中,有 6 所开通了机构知识库。笔者从资源的丰富性、易用性、规范性和可持续性 4 个维度对比分析其资源建设现状。

表 1 C9 高校与常青藤高校机构知识库情况

国别	高校	机构知识库名称
中国	北京大学	北京大学机构知识库
	清华大学	清华大学机构知识库
	浙江大学	浙江大学机构知识库
	复旦大学	--
	上海交通大学	上海交通大学机构学术信息库
	南京大学	--
	中国科学技术大学	--
	哈尔滨工业大学	--
	西安交通大学	西安交通大学机构知识门户
美国	哈佛大学	Digital Access to Scholarship at HarvardHarvard Dataverse Harvard Smithsonian Digital Video Library Video Library
	耶鲁大学	Cross Collection DiscoveryManuscripts and Archives Digital Image DatabaseYale Medicine Thesis Digital Library
	普林斯顿大学	Dataspace
	哥伦比亚大学	Columbia University Academic Commons
	宾夕法尼亚大学	Celebration of Women Writers Scholarly Commons@ Penn
	达特茅斯学院	--
	布朗大学	--
	康奈尔大学	arXiv. org e-Print Archive DigitalCommons@ ILR eCommons@ Cornell Kinematic MOdels for Design Digital Library

2.1 丰富性

资源建设的丰富性有 3 层含义,一是机构知识库的收录资源的个体数量,二是机构知识库收录资源的类型数量,三是机构知识库资源覆盖的学科或主题范围。

首先,从资源收录的绝对数量上来看,C9 高校

和常青藤高校机构知识库的资源收录数量差距很大(表 2), C9 高校机构知识库的平均资源收录量仅为常青藤高校平均资源收录量的 14.45%。其中我国收录资源数量最多的浙江大学的资源规模仅为美国耶鲁大学的机构知识库资源规模的 14.26%。

表 2　C9 高校与常青藤高校机构知识库资源收录情况比较

C9 高校	所属 IR 资源收录量/条	常青藤高校	所属 IR 资源收录量/条
浙江大学	254009	耶鲁大学	1781499
西安交通大学	120195	哈佛大学	1265527
清华大学	114678	康奈尔大学	1176827
北京大学	27756	普林斯顿大学	68463
上海交通大学	2857	哥伦比亚大学	17822
--	--	宾夕法尼亚大学	5298

注:资料来源于相应机构知识库网站

其次,从资源的类型上看,C9 高校和常青藤高校图书馆的机构知识库基本都收录了多种类型的数据资源,既包括正式出版的研究论文和学位论文,也包括工作报告、实验数据、课件等灰色文献。中美机构知识库在资源收录的类型方面差别不大,均以期刊论文、学位论文和会议论文为主。不同的是,包括哥伦比亚大学、耶鲁大学在内的美国高校图书馆机构知识库除了收录以文字为载体的资源数据外,还收录了如录音材料、视频资料和工具软件等多媒体资源。

最后,从资源覆盖的学科或主题范围看,高校图书馆机构知识库的学科和主题范围基本与所属高校的学科和研究范围重合。C9 高校和常青藤高校图书馆的机构知识库资源都属于多学科(主题)范畴。

2.2　易用性

评价资源的易用性可以从资源组织方式的易用程度和资源的全文获取容易度两个方面展开。从资源组织方式的易用程度来看。C9 高校图书馆机构知识库都是使用开放源码软件 DSpace 建设的,数据对象组织模型采取层级制。整个机构知识库由若干个 Community 构成,每个 Community 下包含若干个 Collection,每个 Collection 下包含若干个具体的条目。因而,机构知识库的资源组织和浏览方式相对于成熟的商业文献数据库要单一许多,一般仅提供作者、题名、时间、资源类型等几种方式进行浏览。如浙江大学机构知识库仅提供了按照期刊论文、会议论文和学位论文 3 种浏览方式。在这一点上,虽然常青藤高校机构知识库的建设软件要多元化一些,但其最终提供的资源组织和浏览方式与国内也基本相同。由于机构知识库只对资源的进行简单的 DC 元数据描述,缺乏对资源的更深层次的描述和揭示,从而对 IR 中资源的进一步加工整理和控制造成不利影响,也降低了 IR 资源的在浏览和检索上的易用性。

从获取资源全文的容易度上看。IR 的本意是要促进学术信息资源的开放获取。然而调研中却发现,C9 高校机构知识库极少提供可开放获取的全文链接,部分提供全文链接的也存在“链接无效”和“IP 限制”的情况。如西安交通大学的机构知识门户,仅提供了论文资源的题录浏览功能,并不提供全文链接[16]。常青藤高校机构知识库均提供了资源的有效全文链接,约 60% 的资源提供了可开放获取的全文链接。其中康奈尔大学的机构知识库[17](arXiv. org e-Print Archiv)抽样全文获取率为 100%,麻省理工学院图书馆的机构知识库(DSpace at MIT)已基本实现全部论文资源的开放获取[18]。

2.3　规范性

完善制度是确保机构知识库走向规范的前提。Open DOAR 将相关的标准和规范设定为 5 类[12],即元数据再利用制度、内容再利用制度、资源内容制度、提交制度和保存制度。C9 高校和常青藤高校机构知识库资源建设标准和规范见表 3。

表 3　C9 高校与常青藤高校机构知识库标准和规范情况比较

标准与规范		C9 高校机构知识库数量/个	常青藤高校机构知识库数量/个
元数据再利用制度	允许非商业目的元数据复用	0	3
	允许商业目的的元数据复用	0	
	无	5	3
内容再利用制度	允许机器人收割于完整数据项	0	1
	允许非商业目的使用数据项	1	2
	允许商业目的使用数据项	0	0
	对不同数据项种类限定不同	0	0
	无	4	3
资源内容制度	无	5	6
提交制度	无	5	6
保存制度	无	5	6

注:数据来源于 OpenDoar 及各机构知识库网站

从调研数据看,高校图书馆鲜有进行资源建设标准和规范设定,国内方面仅有清华大学机构知识库对资源的内容再利用制度进行设定;即使在美国,进行了资源建设标准和规范的设定的机构知识库数量也很少[5]。

2.4 可持续性

机构知识库资源建设的可持续性是指当前的资源建设政策是否能够推动学术信息资源开放共享的持续发展。

Kathleen Shearer[19]曾表示,"(资源提供者的)输入行为和机构知识库的利用率是决定机构知识库成功的两个重要因素。"高校图书馆机构知识库的资源采集政策正是影响其资源输入的关键。当前,机构知识库的资源采集政策可以分为 3 类,一是强制类政策,即主管单位强制要求学术研究人员上传论文、会议文献等资源;二是鼓励性政策,即将资源的上传采集与某种激励机制挂钩,如中科院部分院所的科研人员向机构知识库提交论文,将作为年终考核绩效和晋升的依据;三是建议性政策,即仅向科研人员发出资源采集的倡议。

从本次调研的情况来看,C9 高校和常青藤高校的机构知识库都采用的是建议性政策,呼吁科研人员上传资源。而从全球范围来看,这种模式获得的资源输入比例非常低,仅有 15% 左右[20]。机构知识库的资源采集的障碍成为阻碍学术资源开放利用可持续发展的重要因素。

为了解决单个机构知识库在资源建设上面临的诸多难题,很多发达国家和地区开始构建机构知识库联盟借以实现更广泛的学术资源共享,扩大学术影响。如常青藤高校之一的康奈尔大学牵头建立的机构知识库联盟(arXiv. org e-Print Archive),由特定的知识库建设中心建立一个统一、集中的服务系统与平台,并长期管理所收集的元数据与内容,运行及维护整个知识库系统,各成员机构则需要积极地将本机构的学术资源和对应的元数据上传到知识库中心系统中,而不需要承担维护及管理系统的任务[21]。包括 C9 高校在内我国大陆高校图书馆机构知识库普遍加入了 CALIS 机构知识库联盟,成员机构根据共同的数据交互标准和协议分别建立各自独立的知识库,并在数据与资源本身都是分布式和实现独立搜索的基础上,将元数据采集并存储到一个集中的搜索数据库中,通过建立统一检索平台的方式为用户提供其所有成员机构知识库资源的检索[22]。但这种分布式的 IR 联盟成员库之间缺乏必要的交流共享,加之全文获取率偏低的原因,通过 CALIS 机构知识库统一检索的使用体验并不好。

3 C9 联盟高校与常青藤高校的机构知识库资源建设的启示

通过对中美两国顶尖高校机构知识库资源建设现状进行深度比较,得到以下几点启示。

第一,我国高校图书馆机构知识库的建设仍处于初级阶段。从中美两国高校机构知识库的数量来看,我国高校机构知识库的保有量明显低于美国,收录资源的总量上也存在巨大差距。从机构知识库的质量上来看,根据 Ranking Web of Repositories 2016 年 1 月最新发布的排名结果,我国大陆地区高校机构知识库排名第一的是厦门大学机构知识库,全球排名仅为 278 名,北京大学机构知识库全球排名 1 830 名。从机构知识库联盟的角度看,国内成熟的仅有中科院机构知识库服务网格和 CALIS 机构知识库,其中后者仅有 37 所加盟高校。就全中国的科研系统而言,开放学术资源的辐射范围相当有限。

第二,当前机构知识库资源建设现状不容乐观。相比机构知识库近年来在数量上的高歌猛进,其资源建设质量还有待提升,无论中国或美国高校都存在提升的空间。从调研的情况来看,中美普遍存在资源整体质量参差不齐,全文资源获取困难等问题。尤其是各机构知识库之间的资源收录量两极分化现象非常严重。作为中美著名高校,上海交通大学和宾夕法尼亚大学机构知识库的资源收录量甚至不足 10 000 条。统计结果也显示,数据在 10 000 条以上的机构知识库数量仅占知识库总数的 30%[23],少数发展较好的机构知识库占据了较大的资源量,绝大多数机构知识库的资源稀少,有的甚至近乎空壳。

第三,资源组织的易用性和规范性不足。虽然 IR 在学术界处于研究热潮,但 IR 的建设和应用达到预期效果的寥寥无几。目前我国很多机构知识库沦为了"鸡肋"[24],大多数研究人员更希望从专业数据库中查资料[25]。这一点与机构知识库资源组织的易用性和规范性不足有直接关系。从本次调研

的结果来看,中美两国的高校图书馆都缺乏对资源建设确定必要的标准和规范,更缺少对资源的深度组织和揭示,信息服务流于表面。

第四,资源建设的可持续发展前景堪忧。高校机构知识库资源采集政策一方面受制于版权因素,另一方面也受制于领导对 IR 建设的不重视。调查显示[2],虽然超过总量60%的用户知晓 IR,但40%的用户对 IR 用途不清楚或完全不明白,其中 30%用户从未听说过;而实际上仅有 43%的用户使用过 IR 。在此背景下,机构知识库资源采集难以通畅、全面,机构成员产出的新成果不能及时存缴到 IR 中,学术信息资源的深度开放获取更难以持续推进。

4　结语

C9 高校和常青藤高校作为中美两国顶尖高校的代表,因其在各自国家都具有极高的社会地位和学术影响力,其获取机构知识库资源建设支持的能力的也应当更强,其在机构知识库资源建设工作中的表现具有典型性,且对其他学校能起到榜样作用的。然而调研结果表明,当前高校图书馆机构知识库在数量急速增长的同时,质量并未同步提升。中美两国的高校机构知识库在资源建设的丰富性、易用性、规范性和可持续性方面都存在不足,我国高校在一些方面的欠缺更为明显。对于这些共性问题,中美高校应当竭力避免重蹈覆辙,这也将是未来深入研究和解决的方向。

【参考文献】

[1] 张晓林,李　麟,刘细文,等. 开放获取学术信息资源:逼近“主流化”转折点[J]. 图书情报工作,2012,56(9):42-47.

[2] 姚晓霞,聂　华,顾立平,等. 我国教育科研机构知识库建设现状调查与分析[J]. 现代图书情报技术,2014(5):1-9.

[3] 张宏胜. 中美高校机构知识库比较研究[J]. 科技情报开发与经济,2011,21(34):111-113.

[4] 陈美华,刘文云,毕　煜,等. 美国机构知识库联盟的建设研究及对我国的启示[J]. 图书馆,2015(11):59-64.

[5] 邱小红,姜　颖. 中美机构知识库资源建设现状与特点剖析[J]. 图书情报知识,2015(3):16-21.

[6] ansley R, Bass M, Stuve D, *et al*. The DSpace institutional digital repository system: current functionality[C]// Proceedings of the 3rd ACM/IEEE-CS joint conference on digital iibraries. IEEE Computer Society, 2003:87-97.

[7] Raym Crow. The case for institutional repositories: a SPARC position paper [EB/OL]. [2016-04-14]. ARL Bimonthly Report, 2002 (223). http://www. sparc. arl. org/sites/default/files/media_files/instrepo. pdf.

[8] Lynch CA. Institutional repositories: essential infrastructure for scholarship in the digital age[J]. Portal Libraries & the Academy, 2003, 3(2):327-336.

[9] 包心萍,刘兹恒. 台湾高校图书馆机构知识库发展研究[J]. 大学图书馆学报,2012,30(4):61-65.

[10] 顾立平. 提升机构知识库作用的实践:多元学科教育资数字图书馆源的集成与再利用[J]. 图书情报工作网刊,2011(2):55-82.

[11] 聂　华,韦成府,崔海媛. 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.

[12] The directory of open access repositories [EB/OL]. [2016-02-14]. http://www. opendoar. org.

[13] The registry of open access repositories [EB/OL]. [2016-04-14]. http://roar. eprints. org.

[14] Ranking web of repositories[EB/OL]. [2016-04-14]. http://repositories. webometrics. info/en.

[15] 教育部公布全国高等学校名单[EB/OL]. [2015-05-21]. http://www. chinanews. com/gn/2015/05-21/7293105. shtml.

[16] 西安交通大学机构知识门户[EB/OL]. [2016-02-20]. http://www. ir. xjtu. edu. cn/.

[17] arXiv. org e-Print Archive[EB/OL]. [2016-02-20]. http://arxiv. org/.

[18] DSpace at MIT[EB/OL]. [2016-02-20]. http://dspace. Mit. edu/.

[19] ShearerK. Institutional repositories: TowardstheIdentifica-tionof-CriticalSuccessFactors[EB/OL]. [2015-12-05]. https://dspace. ucalgary. ca/bitstream/1880/43357/6/CAIS-IR. pdf.

[20] Les Carr, Alma Swan. Repositories for institutional open access: mandated deposit policies[EB/OL]. [2016-02-21]. http://eprints. soton. ac. uk/263099/2/abs77. pdf

[21] 陈美华,刘文云,毕　煜,等. 美国机构知识库联盟的建设研究及对我国的启示[J]. 图书馆,2015(11):59-64.

[22] Digitalcommons@ carli. [EB/OL]. [2015-03-30]. http://digitalcommons. carli. illinois. edu/.

[23] 付希金,田丽君,张艳婷,等. 国内外区域 IR 联盟建设比较及我国区域 IR 联盟构建路径研究[J]. 情报科学,2015(3):35-40.

[24] 郑　萍,薛德军,蔡　郧,等. 基于全方位服务机制建设机构知识库研究[J]. 图书馆研究,2015,45(5):16-21.

[25] 钱建立,李　鹏,李若溪. 机构知识库可持续发展策略研究[J]. 情报杂志,2012(11):176-180.

[收稿日期:2017-04-14]

[本文编辑:王天津]