

基于功能视角的高校机构知识库建设策略探索

宗 良

(上海理工大学 图书馆, 上海 200093)

摘要: 基于 Open DOAR 中注册登录的各国高校机构知识库的数据, 运用网络调研法和文献分析法, 对国内外高校知识库的现状进行研究和分析。从机构知识库的功能角度出发, 提出可行的建设策略: 首先, 应搭建功能完备的开放获取 (OA) 平台, 提供资源搜索和下载、期刊查询、政策解读等基本功能; 其次, 充分发挥平台的 OA 功能, 实现学术资源共享; 再次, 增强平台的科研辅助功能, 如知识图谱展示和开放期刊推荐等; 最后, 加强开放获取理念和知识库平台的推广力度, 提高科研人员对开放科学的关注度和参与度。

关键词: 功能视角; 机构知识库; 开放获取

中图分类号: G 250.74

文献标志码: A

文章编号: 1009-895X(2023)04-0450-07

DOI: 10.13256/j.cnki.jusst.sse.210612199

Research on the Strategy of Functional Construction of Institutional Repository in Universities

ZONG Liang

(Library, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: Based on the data of the institutional repository of universities registered in Open DOAR, the current situation of the institutional repository of domestic and foreign universities is analyzed with the method of network document research. An effective and feasible strategy is proposed for the institutional repository from the functional point of view. First, a fully functional open access (OA) platform should be built up to provide literature search and download, journal query, and policy interpretation. Second, the OA platform should function as sharing academic resources. Thirdly, the platform should enhance its scientific research support functions, such as knowledge map display and open journal recommendation. Finally, the concept of open access and the institutional repository platform should be strengthened to draw more attention from scientific researchers.

Keywords: functional point of view; institutional repository; open access

“最珍贵的往往是免费的, 比如空气, 而知识是现代社会的空气, 知识获取成本越低, 越有利于创新型国家的建设。”2020 年两会期间, 全国政协委员、上海市教育委员会副主任倪闽景呼吁在政

府采购知识服务、合理确定成本的基础上, 让中国知网向国内民众免费开放^[1]。一时间, 关于开放获取 (Open Access, OA)、开放科学的探究成为科研学术圈内的热门话题, 并在“知乎”等知名网站

收稿日期: 2021-06-12

作者简介: 宗 良, 男, 馆员。研究方向: 图书管理、借阅及文献传递。E-mail: zl1226@163.com

上引发了热烈的讨论。在 1.2 万余人参与的投票中，有 82% 的网友认为知识聚合平台上论文的浏览与下载应向公众免费开放。由此，科研人员对于学术资源开放获取的渴求可见一斑。

机构知识库（Institutional Repository, IR）是高校和科研机构为保存教学和科研活动中产生的研究成果而建立的数字资源库，是开放获取运动的重要组成部分^[2]。高校机构知识库是实现高校学术资源和科研成果开放获取的重要平台，旨在消除学术交流的限制，推动开放科学的发展。很多国外高校已在知识库建设中取得了一定的成果^[2-4]，但国内的高校对于知识库的关注度和参与度仍稍显不足。如何利用本校的信息资源，为读者和科研人员提供便利，减少校内人员的资源收集和学术交流成本？本文基于 Open DOAR 中注册登录的国内外高校知识库相关数据，利用网络调研法调查我国部分高校知识库的现状，从知识库需要实现的开放获取、科研辅助、理念科普等功能上，对知识库的构建提出了可行性框架，以期为高校知识库的建设提供有效的借鉴。

一、国内外机构知识库实践概况

Open DOAR（Directory of Open Access Repositories，开放存取知识库目录）是英国的诺丁汉（Nottingham）大学和瑞典的隆德（Lund）大学图书馆在 OSI，JISC，SPARC Europe 和 CURL 等学术机构的资助下于 2005 年 2 月共同创建的开放存

取机构资源库、学科资源库目录检索系统，是全球收录资源最全面、最权威的开放存取知识库目录。2006 年 1 月登陆互联网提供服务。

Open DOAR 中知识库历年的数量统计结果如图 1 所示。截至 2020 年 11 月，在 Open DOAR 中注册登记的知识库已有 5 573 个。根据图 1 的统计结果可以发现，随着开放获取运动的兴起，机构知识库在世界范围内迅速发展，数量逐年递增，且增长速度迅猛。与 2015 年 2 月（2 670 个）相比，机构知识库的注册数量翻了 1 倍；与 2010 年 2 月（1 364 个）相比，增长了 3 倍。

（一）国内外高校注册登记知识库情况

美国作为最早提出机构知识库建设的国家，截至 2020 年 11 月，已有 903 个知识库在 Open DOAR 中注册，是该领域内影响力最高的国家。其次是日本（683 个）、英国（309 个）、德国（275 个），接下来是西班牙（168 个）、秘鲁（155 个）、巴西（151 个）和法国（151 个）。中国有 119 个注册机构知识库，与 2015 年的 96 个相比增幅不大，这表明我国的机构知识库建设增速缓慢，且远远落后于美日等国家。

通过对国外机构知识库相关论文数据的深度挖掘，发现国外高校院系是该领域研究和实践的主力^[4]，其系统性、深度化的研究为我国尤其是高校机构知识库的建设提供了至关重要的理论指导。对上述部分国家和地区的机构知识库进行统计，数据结果如表 1 所示。

由表 1 可以看出，全球机构知识库数量排名前

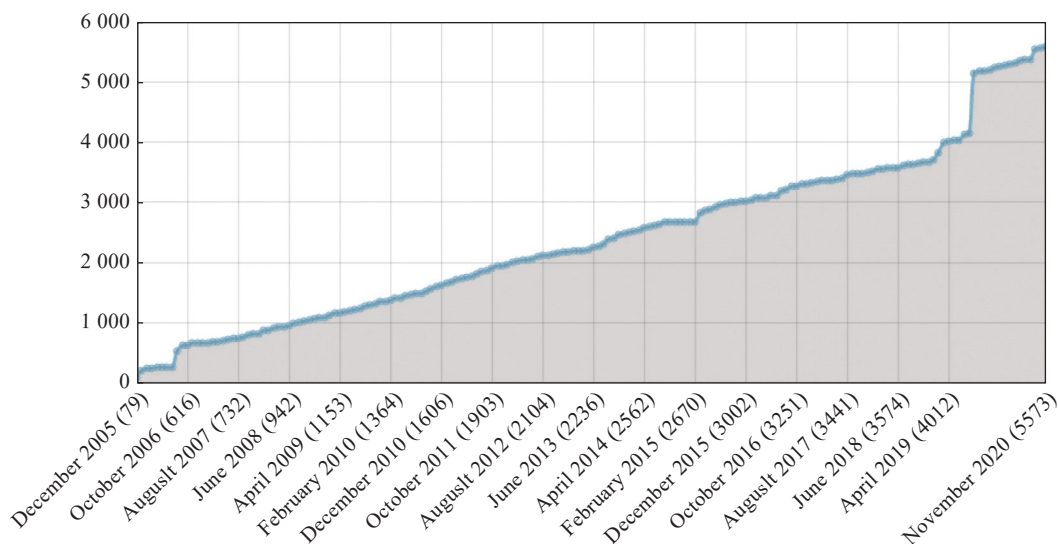


图 1 Open DOAR 中知识库历年数量统计结果

Fig. 1 Statistics of institutional repository in Open DOAR

20 位的国家和地区中,由高校或相关机构、组织注册登记的知识库比例绝大多数高于 60%,其中土耳其的比例最高为 97.1%。总数量排名前三位的美国、日本、英国,由高校或相关机构、组织注册的知识库数量排名依然保持在前三位,且占比均超过 75%,

这说明美国、日本和英国的高校在知识库建设方面遥遥领先。而中国共有 119 个机构知识库,其中有 73 个为高校、中国科学院及下属科研院所注册,占比为 61.3%,说明我国高校的知识库建设参与度仍处在较低的水平。

表 1 各国注册登记知识库数量统计结果
Tab. 1 Statistics of registered institutional repository

国 家	总数量/个	高校或相关机构 注册数量/个	占比/%	国 家	总数量/个	高校或相关机构 注册数量/个	占比/%
美国	903	747	82.8	土耳其	139	135	97.1
日本	683	631	92.4	克罗地亚	136	128	94.1
英国	309	237	76.7	中国	119	73	61.3
德国	275	192	69.8	波 兰	115	60	52.2
西班牙	168	99	58.9	乌克兰	103	99	96.1
秘 鲁	155	122	78.7	印 度	97	45	46.4
巴 西	151	103	68.2	加拿大	96	79	82.3
法 国	151	86	57.0	哥伦比亚	95	75	78.9
印度尼西亚	148	112	75.7	澳大利亚	90	70	77.8
意大利	141	109	77.3	荷 兰	73	51	69.9

(二) 国外高校知识库建设内容分析

对国外高校的机构知识库进行网络调研,包括平台功能、学术成果存储量、开放获取程度、OA 政策体系及其他功能,结果如表 2 所示。

在调研过程中发现,大多数国外高校的知识库网站功能简洁,一般包括作者登录、作品提交、作品检索、使用统计、常见问题解答、较完备的 OA 政策或隐私政策等。部分高校的 OA 政策中会提到,本校师生可以申请的基金项目对发表项目成果的要求,例如在 OA 期刊上发表或者规定在其他期刊发

表后的禁运期限。但是,一般不会在网站主页上显示现有的 OA 期刊或提示可选择的发布平台,用户不易关注和查找到相关信息。

大部分知识库的内容如论文、档案、著作、案例研究等,均可以自由访问和免费下载,也有极少的部分作品是链接到付费平台或联系平台后获取。对于知识库中学术成果的新增数量、下载次数、高频次下载作品的统计,有助于用户了解本校相关专业的热点方向和已有成果,同时知识库可以据此提升平台利用率并拓展更多功能。

表 2 部分国外高校知识库建设情况统计
Tab. 2 Statistics of institutional repository construction in some foreign universities

高校名称	知识库名称	存储量/件	是否可自由 访问、下载	是否有使用 情况统计	是否有完整的 OA政策体系
哈佛大学(文理学院)	DASH	23 163	是	是	是
康奈尔大学	eCommons	91 773	是	是	是
麻省理工学院	DSpace@MIT	115 348	是	是	是
牛津大学	ORA	257 754	是	是	是
剑桥大学	Apollo	175 915	是	否	是
诺丁汉大学	Nottingham ePrints	31 748	否	是	是

（三）我国高校知识库建设现状及分析

选取我国 15 个高校知识库的统计情况如表 3 所示。由表 3 可知，除了香港大学拥有 2 个知识库外，其他高校的注册数量均为 1 个；大多数知识库的数据内容为多学科的学位论文、期刊文章、会议和研讨会论文、专利、数据集，以及书或章节等。元数据存储量方面，15 个知识库的总存储量为 1 397 161 件，仅北京大学一所高校，其开放获取资源存储量就达到 510 704 件，占总数的 36.6%，而

福建师范大学的存储量仅为 326 件，说明我国高校间知识库的存储数据量差距悬殊。访问、下载方面：仅有 7 个知识库可以不受 IP 地址的限制自由访问和下载相关资源；北京大学和西安交通大学的知识库仅对校内 IP 或访客自由开放；南方科技大学和香港大学的 2 个知识库需要提供获取理由和电子邮箱，才可获得相关资源；有 4 个知识库的网页处于“无法访问”的状态。总体来说，我国高校知识库中科研成果的开放获取程度较低，未达到绿色开放获取的状态。

表 3 我国高校知识库统计结果

Tab. 3 Statistics of institutional repository in Chinese universities

高校名称	知识库数量/个	知识库名称	学科	语言	存储量/件	是否可自由访问、下载
北京大学	1	北京大学机构库	多学科	中文、英语	510 704	否
清华大学	1	清华大学机构知识库	多学科	中文	114 678	无法访问
北京科技大学	1	北京科技大学机构知识库	多学科	中文	95 207	无法访问
西安交通大学	1	西安交通大学机构知识库	多学科	中文	120 195	否
南方科技大学	1	南方科技大学知识苑	技术总论	中文、英语	24 485	电子邮件获取
兰州理工大学	1	兰州理工大学机构知识库	技术总论	中文、英语	65 304	是
福建师范大学	1	图书馆学系机构库	社科综合	中文	326	无法访问
厦门大学	1	厦门大学学术典藏库	多学科	中文、英语	156 359	是
广西民族大学	1	广西民族大学机构知识库	多学科	中文	3 937	是
澳门大学	1	澳门大学机构资料库	多学科	中文、英语	3 146	无法访问
香港城市大学	1	CityU Institutional Repository	多学科	英语	2 788	是
香港理工大学	1	PolyU Institutional Repository	多学科	英语	81 076	是
香港大学	2	香港大学学术库	多学科	中文、英语	195 817	电子邮件获取
		香港大学论文在线	多学科	中文、英语	11 850	电子邮件获取
香港科技大学	1	香港科技大学机构知识库	多学科	中文、英语	11 289	是

根据文献 [5] 中荷兰莱顿大学科学技术研究中心发布的《2019 年 CWTS 莱顿大学排名》数据，全球参与 OA 出版的高校中，我国的高校总数达 165 所，仅次于排名第一的美国（173 所），但 OA 论文的比例均处于中低区间，说明我国高校对于 OA 出版有一定的认知，但缺乏广泛的关注度和深入的参与度。这一结论与本文统计结果所反映的问题是—致的。

对表 3 中部分知识库，以及未在 Open DOAR 中注册可开放获取的高校知识库共 7 所高校的知识库功能模块进行统计和分析，以探讨如何在功能视角构建知识库平台，统计结果如表 4 所示。

由表 4 可以看出，对于知识库的基本功能即资源检索、作者服务、数据分析，大多数国内高校知识库都已具备。其中，大连理工大学知识库根据对各学院和机构的数据分析，提供有针对性的友好期刊清单，以提升机构师生的投稿利用率。香港大学的知识库平台还列出了该校学者中心的所有资助基金，包括进行中和已完结的基金项目和资助计划，以帮助本校师生了解和申请各类资助。

二、我国高校机构知识库的构建策略

在推进 OA 出版的过程中，为实现学术成果的

表 4 国内 7 家高校知识库的功能模块分析
Tab. 4 Functional modules of institutional repository in seven domestic universities

高校名称	知识库名称	资源搜索	知识图谱或 数据分析	作者个性化 服务	OA政策 及解读	国内外OA 平台接口	其他辅助工具
北京大学(医学部)	北京大学医学部机构知识库	√	√	√	—	—	—
清华大学	清华大学学者库	√	√	√	—	—	—
武汉大学	武汉大学机构知识库	√	√	√	√	—	—
南京大学	南京大学机构知识库	√	√	√	√	—	—
大连理工大学	大连理工大学机构知识库	√	√	√	√	—	友好期刊清单
香港大学	香港大学学术库	√	√	√	√	√	基金项目
香港科技大学	HKUST SPD	√	—	√	√	—	—

真正免费在线获取,需要相关利益方的共同努力。具体做法有:在国家战略层面将开放获取作为科学数字化转型综合战略的一部分,提出和制定相关的支持政策,建立开放科学推进机制;项目资助方、出版机构和科研人员各方在开放获取的资金支持、利益保障和资源共享需求等方面寻求平衡;高校可以为本校不同形式的科研成果提供开放获取的平台和接口,为科学出版体系的全面转型推波助澜,为相关项目和出版物拓宽资助渠道、提高社会效益,以及加快科研成果转化。高校知识库依靠高校优秀且丰富的学术资源优势,庞大且稳定的作者和读者用户群体,来自各级政府、企事业单位的项目资金支持,具有实现上述功能的有利条件。

(一) 搭建功能完备的开放获取平台

高校机构知识库应具备的基本功能如下:

- (1) 基于文献题目、关键词、研究专题、文献类型、学科分类、发布日期、出版者、资助项目及作者姓名等为搜索条目的资源搜索功能;
- (2) 对系统中作品进行统计汇总或分析处理,形成从不同维度或层面可视化揭示系统中知识作品分布情况的多种知识图谱;
- (3) 根据关键词、学科门类、期刊名称等提供开放获取期刊名录查询和投稿建议;
- (4) 作者登录系统后,提交作品、作品认领与全文提交、编辑个人信息、用户权限、收藏管理、个人作品统计、个人作品收录引证查询、交流与反馈等功能;
- (5) 针对不同用户提供个性化服务,如条目推荐、结果收藏与分享、Endnote 文件导出、关联搜索等;

- (6) 提供开放获取相关政策解读,以及优惠政策和基金项目的申请信息和跳转接口等;
 - (7) 提供其他辅助工具和国内外开放获取平台接口,以帮助科研人员挑选高质量的期刊或渠道进行学术出版和学术成果分享。
- 为实现上述知识库的相关功能,搭建如图 2 所示的开放获取平台框架。

(二) 充分发挥平台的开放获取功能

科研数据的共享使用和科研成果的开放获取,有助于国家在开放科学运动中占据相关标准制定的主导权,增强国家在参与国际科研环境治理的话语权。中国的开放科学已经“在路上”,但仍需要多层次的推进和多方面的努力。比如,航空航天领域是关系国家命脉的重要领域,其相关数据和成果的共享能够推动国际合作,是人类共同探索宇宙奥秘的体现。2020 年 11 月 24 日,嫦娥五号探测器发射升空,美国 NASA 在其官方推特账号上发文,希望中国能够向全球科学界分享数据。很多人在嘲讽回应的同时,并不知晓一个基本事实:中国早已在“中国探月工程数据发布与信息服务系统”中公布了嫦娥一号到四号的相关原始数据,并且为发布嫦娥五号带回的月壤数据预先开辟了“月球样品”的栏目。中国的科研工作者已经意识到,共享科研数据、降低科研成果的获取成本、提高知识成果转化效率,才是推动开放科学发展的有效途径。

高校机构知识库的构建旨在促进本校科研成果的广泛传播,实现学术资源共享,因此其开放程度决定了知识库的有效性和利用率。从表 2 的统计结果可以看出,我国高校知识库平台的开放程度、读者获取相关资源的便捷程度距离自由、免费地访问

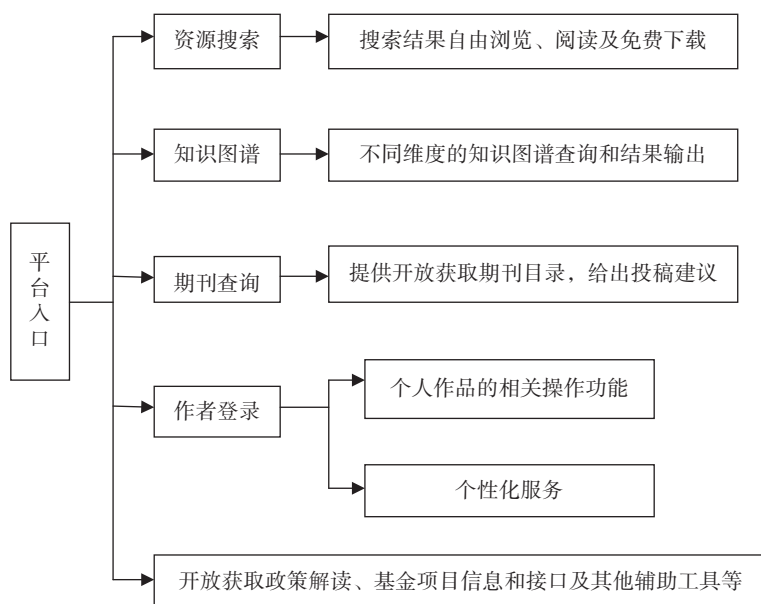


图2 高校知识库平台搭建框架

Fig. 2 Framework of university institutional repository platform

和下载这一目标仍相去甚远。而刘聃等对30所加拿大大学的40个知识库的统计结果显示，有90%的知识库可以不受身份限制地成功访问和下载^[6]。由此可见，为充分发挥高校知识库平台的开放获取基本功能，要尽可能地满足所有访客的访问需求，最大限度地消除访问障碍或限制，并在学校或图书馆门户网站设置导航和链接，提高平台资源的易获取性。

（三）增强平台的科研辅助功能

1. 知识图谱

在图书情报界，知识图谱又被称为知识域可视化或知识领域映射地图，用于揭示知识领域的动态发展规律，为学科研究提供切实的、有价值的参考^[7]。在机构知识库平台中增加知识图谱功能，可以可视化的形式，从不同维度或层面提供对系统中作品的统计汇总或分析处理结果。

例如，以系统中知识成果的研究机构和作者关系数据为基础，构建相关领域的研究合作网络，加强作者之间的学术关联；以系统中学术成果全文信息数据为基础，进行关键词共现分析和计算，展示相关领域内该关键词云图和共现图谱，有助于研究人员识别和了解学科研究的前沿和热点，更好地融入主流研究领域，把握学科发展趋势和演化轨迹等。

2. 开放期刊推荐

全球的开放期刊增长迅速，开放期刊数据仓储

越来越多。如：2003年瑞典隆德大学推出的开放获取期刊目录（Directory of Open Access Journals, DOAJ）包含300种开放获取期刊，如今已达到15607种经过同行评审的OA期刊，涵盖科学、技术、医学、社会科学、艺术和人文科学的所有领域；2010年中国科学院主管的中国科技期刊开放获取平台（China Open Access Journals, COAJ）发布上线，目前已收录的期刊数量达660种，可检索论文的期刊数量为340种。但如何对期刊数据进行整合利用，加强平台与访客之间的互动和联系，为作者和读者提供更高效率的个性化服务，仍是研究人员关注的重点之一。

很多科研人员对开放获取的理念和政策有一定的了解，但如何在发布科研成果时寻求或选择合适的平台，这方面仍存在很大的困扰。在作者投稿时，引导其优先选择开放获取期刊，在高校知识库平台上提供相关期刊目录和开放资源列表，作为投稿建议，体现出平台的科研辅助功能。当然，仅仅是目录和列表的展示远不能满足科研人员的发现、获取和利用需求，还要对期刊数据进行深度挖掘和聚合，如收录情况、影响力、评价、投稿指南等信息，并在此基础上根据关键词为作者挑选高影响因子的开放获取期刊，提高投稿命中率，促进知识的交流和传播。

（四）加强开放获取理念和知识库平台的推广力度

2003年，中国科学院路甬祥院长代表中国科学

家在《开放获取柏林宣言》上签字,成为国际开放获取运动的先驱缔约者。2005年7月,在武汉大学信息管理学院“数字时代图书馆合作与服务创新”国际研讨会上,来自国内50多所高校的图书馆馆长举办了“中国大学图书馆馆长论坛”,发表了《图书馆合作与信息资源共享武汉宣言》^[8]。2010年10月,“第八届开放获取柏林国际会议”在中国北京举行,这是该系列会议第一次在欧洲以外的国家召开,国内百余家研究教育机构、图书馆届的180名代表参加了会议。2015年“第十二届柏林开放获取会议”上通过“OA2020”倡议,截至2020年10月,已有147个学术机构和组织正式签署意向书,其中包括18所中国高校和研究机构。由此可见,开放科学作为全球范围内新兴的科学理念和政策理念已在我国引起了一定的关注,但是仍有待进一步提升,相关的基础设施建设水平、对开放获取的关注度和参与度与科研人员对开放科学的迫切需求仍存在较大差距。

为了推广开放获取的理念、建设高质量的知识库平台,可以从以下两个方面开展相关工作。一方面,高校可依托知识库平台,开展形式多样的开放获取主题活动。如:加拿大大学的“开放获取周”^[6]等,以讲座、专题研讨的方式,邀请相关方面的专家、学者解读政策,普及开放获取观念和知识,并鼓励师生向开放获取期刊投稿或将学位论文、科研论文及其他科研成果存储至本校知识库,在促进学术成果的共享和交流的同时,为知识库积累宝贵的学术资源;汇聚项目资助方、科研人员和出版机构,提供以开放出版学术成果为前提的项目资助平台,探讨解决利益诉求冲突的途径,促进更多创新成果的诞生,并加快科研成果的转化,形成良好的学术生态圈。

另一方面,在扩大开放程度、丰富内容资源、提高更新频率的基础上,高校知识库也应积极参与校外知识库联盟平台的建设,加强国际合作。借鉴日本的开放获取知识库联盟(Japan Consortium for Open Access Repository, JPCOAR)^[9],该平台推出的云服务使得成员机构的知识库可以在统一的数据标准下得到维护和管理,各知识库本身只需专注于资源建设。在国内外权威开放资源平台注册登记,也是拓展学术成果共享渠道的有效途径。目前国内较有影响力和权威性的平台有Open DOAR,

ROAR,中国高等教育文献保障系统(CALIS)等。

三、结语

知识在数字化时代应是全球共享的,建设机构知识库是实现这一目标的重要途径,其重点在于用户以更加便捷的方式自由、免费地访问和获取所需的学术资源。随着科学技术的发展进步和人们对开放科学理念的深入理解,作为各国机构知识库重要组成部分的高校机构知识库的建设会更加成熟。在提高开放程度、提供科研辅助、推广开放理念等功能的基础上,提高资源的整合利用,加快移动平台的开发,建立稳定的基金项目和科研人员的合作关系,扩大知识库的建设规模和利用效率,这些仍有待进一步深入研究。

参考文献:

- [1] 张惠娟. 倪闽景委员:中国知网,免费?免费![EB/OL]. (2020-05-21) [2020-11-06]. <http://www.rmzxb.com.cn/c/2020-05-21/2577013.shtml>.
- [2] 刘莉,刘文云,马伍翠,等. 英国高校机构知识库开放获取政策研究[J]. 情报理论与实践, 2019, 42(3): 171-176.
- [3] 杨杰,张新鹤. 美国高校图书馆开放获取政策研究[J]. 图书馆学研究, 2017(23): 86-91.
- [4] 王婧,田爱苹,李文兰. 国外机构知识库领域研究脉络及热点前沿研究[J]. 图书馆, 2019(10): 64-70.
- [5] 郁林羲. CWTS排名中全球高校开放获取概况及对我国的启示[J]. 中国科技期刊研究, 2020, 31(2): 204-214.
- [6] 刘聃,高波. 加拿大大学图书馆开放获取现状分析[J]. 图书情报工作, 2018, 62(19): 129-137.
- [7] 秦长江,侯汉清. 知识图谱——信息管理与知识管理的新领域[J]. 大学图书馆学报, 2009(1): 30-37.
- [8] 陆宝益. 以学习《宣言》为契机,进一步推动高校图书馆合作与资源共享——论《图书馆合作与信息资源共享武汉宣言》的内容特点与发表意义[J]. 图书馆建设, 2006(4): 85-88.
- [9] 孙海燕,解登峰,齐晓晨. 日本机构知识库联盟JPCOAR的开放获取策略及启示[J]. 数字图书馆论坛, 2019(6): 46-53.

(编辑:程爱婕)