

# 医科院校图书馆联盟机构知识库建设模型

范晓磊,陈 静

(南京医科大学图书馆,江苏 南京 211166)

**摘要:**医科院校与附属医院的图书馆关系紧密,文章以南京医科大学图书馆及其附属医院分馆联盟的学术交流与资源共享为基础,提出共同建设区域型医药学科特色的联盟机构知识库,设计具有较好实践性的、以点带面且独具特色的建设模型,采取分散建设最终集中呈现的模式来逐步完成全国性的大型机构知识库的建设。

**关键词:**医科院校;图书馆联盟;机构知识库;模型设计

中图分类号:G205.76

文献标志码:A

文章编号:1671-0479(2017)01-063-004

doi:10.7655/NYDXBSS20170115

机构知识库(institutional repository, IR)是学术科研机构在网络环境下建立的一种学术资源开放共享平台,平台对该机构产生的各类型学术成果进行直接或数字化后的存储、管理和发布,以实现可开放资源的共享,提高机构学术影响力。机构知识库始于开放存取(open access)运动,伴随这种全新学术交流与资源共享方式而兴起。紧随国际发展步伐,我国机构知识库建设也在蓬勃发展,本文基于南京医科大学图书馆及其附属医院分馆联盟的学术交流与资源共享,提出构建一种小型但具有医药学科特色的联盟机构知识库的模型。

## 一、我国现有机构知识库模型及建设方式

### (一)国际形势

2002年11月,Dspace系统的问世为全球大规模建设机构知识库提供了技术保障,截至2016年5月25日,开放存取知识库名录(The Directory of Open Access Repositories, OpenDOAR)全球共登记有3090个<sup>[1]</sup>机构知识库成立和投入使用,尤其欧美等国家的机构知识库发展已趋于成熟,很多国家及机构都有开放获取期刊和各种级别的机构知识库供科研人员免费获取和利用。在机构知识库方案设计和建设、不同类型的文献搜集、配套的计算机软件技术等多方面均有理论基础和实践经验,同时欧美

发达国家大量机构知识库建设的成熟技术和成功案例模型都为我国机构库建设指明了目标和方向。

### (二)国内现状

相对于全球在OpenDOAR注册的3090个机构知识库,中国仅有39个机构进行了注册,这一数字在2014年1月1日是36个<sup>[2]</sup>,我国机构知识库数量少、规模小,发展进程比较缓慢,尤其是在全球资源共享和推送上,与国外发达国家相比还存在一定差距。

实际上我国主要研究机构 and 高等院校都是中央集权行政体制下的组织机构,其科研经费、图书馆经费、运行费用等都由政府统一拨发,这种管理模式显然非常便于机构知识库的建设<sup>[3]</sup>。然而我国现有的机构知识库的实际数字却没有体现出这种优势,已建成机构知识库的资源条目数也远远低于国际水平。

### (三)成功案例

截至2016年5月25日,OpenDOAR中的统计数据涉及医学与健康(health and medicine)学科的机构知识库有310个,中国仅有3个,分别为中国科学院自然科学史研究所机构知识库(Institutional Repository of Institute for the History of Natural Sciences, Chinese Academy of Sciences, IHNS OpenIR),中国科学院心理研究所机构知识库(Institutional Repository of Institute of Psychology,

基金项目:南京医科大学科技发展基金面上项目“高校图书馆分馆联盟机构知识库研究”(2014NJMU016)

收稿日期:2016-07-11

作者简介:范晓磊(1983—),女,江苏南京人,硕士研究生,馆员,研究方向为图书馆资源建设与共享,通信作者。

CAS, PSYCH OpenIR), 和北京大学机构知识库(Institutional Repository of Peking University, PKU OpenIR)。这些数字显示,在全球生物医学文献占据重要地位的形势下,我国在建设机构知识库过程中对于生物医学的学科并没有有效关注。

中国科学院心理研究所的 PSYCH OpenIR 阐述了机构知识成果的收集、长期保存、合理传播以至开展综合知识管理的目标,截至 2016 年 5 月 25 日现有条目数量 7 920 个,全文数量 6 692 篇,浏览总量 1 075 387 次,下载总量 25 583 次。最早的记录为 1951 年,最新记录为 2016 年 5 月 18 日。该机构知识库的提交内容包括该机构及其员工、合作者创造的各种形式的研究和工作成果,具体包括正式出版的研究论文、机构出版物、课题作品、专著,和其他学术作品,如未经同行评议的论文手稿、会议 PPT、实验记录、多媒体文档等,专著等内容需取得出版商的授权,同时也声明了传播授权许可<sup>[4]</sup>。

## 二、医科院校图书馆及分馆联盟机构 知识库建设模型

### (一)优势

无论是国家级别还是区域性质的现有机构知识库,在由科研机构的科研工作人员或对文献内容做出贡献的高校师生主动提交资源的内容征集方法下,由于没有意识到机构知识库的重要性,或者保障机构知识库长期运行的人力与经费不足,使得提交的内容与机构实际产生的知识成果严重不符,机构知识库的架构下没有丰富的知识内容。有多篇文献指出大部分机构知识库包含的文献内容都少于 1 000 条,而较小规模的学术机构大多没有足够的资源支持建设机构知识库系统或提供服务。在现有情形下我们反思,基层学术机构的机构知识库建设并不能照搬固有理论和模式,而应采取灵活的应对措施来实现小型机构知识库的建设,以期以小逐大,在资源丰富的基础上再商区域、国家、国际的机构知识库建设与共享大计<sup>[5]</sup>。

与其他科研机构或高校不同,医科院校与附属医院间具有非常紧密的实质联系,包括科研产出人员的高校教师与医院医师的双重角色,学习、教学与实习将医科院校与附属医院的关系联系得密不可分。这使得医科院校在建设机构知识库的过程中,无论是宣传还是内容征集都能够得到附属医院从上至下的支持和主动关注。南京医科大学图书馆附属医院分馆联盟联合了 18 家附属医院,同时也为南京市卫生局下属的其他医疗机构提供深层次信息服务,

其中就包括了各家机构的学术成果搜集工作,这些已搜集完成的资源为该机构知识库的建设提供了基础保障。而在联盟资源共享服务中,因学校医院间紧密关系产生的联盟机构知识库的建设需求也将得到科研人员的有力支持。

南京医科大学同时也参与中国高等教育文献保障系统(CALIS)和江苏省高等教育数字图书馆(JALIS)的机构知识库建设项目,为联盟机构知识库的建设提供了政策和决策上的保证。联盟资源共享服务人员和现有机构知识库建设人员也提供了人力资源和长期建设与维护、利用的要素。

### (二)模型主体

#### 1.人员

决策者制定相关政策,决定项目是否启动,制订机构知识库许可协议以获得合法身份,解决项目资金来源和技术保证。其中的项目负责人制定发展方向和目标,组织具体实施人员开展工作,并最终评估机构知识库建设的阶段进展和成果。

馆员是项目的实际执行人,联盟机构知识库建设初期最为重要的角色。负责搜集资源,建立元数据模式和标准,组织和发布内容;搭建平台,系统维护,确保功能完善与数据安全;沟通科研工作人员与资源,负责机构库的推广,帮助研究人员提交条目,建立学术社区或群,提交资源与利用资源培训。

科研人员也是注册用户,既负责主动提交自己的研究成果也接受系统推送的资源和服务。

#### 2.要素

硬件条件包括,必须遵守的知识产权相关政策法律,开启项目和维持长期发展的资金来源,适用的具有前瞻性的技术保障。

核心内容有资源和知识,正式的学术出版物有学术论文、专著、会议论文、学位论文等内容;基金项目和实验的相关数据、资料、成果等,这一部分的收集工作可能会非常困难,但其价值也非常高;多媒体资源有 MOOC(慕课)资料、微课视频、手术录像等,这部分资料存储空间需求大,为现今信息技术发展的方向所在,必须预见到此类资源的重要性。

#### 3.服务

提供检索平台。每一个建成或在建的机构知识库都有或将有一个统一检索平台,联盟机构知识库也不例外,这是库内成果的展现平台,也为吸引用户利用和提高机构学术影响力最终目标而存在。同时,也是分散建设后集中呈现的接口。

进行数据挖掘。数据与元数据的加工格式决定了数据挖掘的深度。

#### 4.整合

一是集中呈现。以可扩展形式加工的元数据和模块化建设的检索平台使得本联盟机构知识库能够简便迅速地集中在更大范围的机构知识库中呈现并融合,如江宁地区机构知识库、南京高校机构知识库、JALIS 机构知识库等。

二是深层次服务。除用户自行检索与利用库内知识外,馆员可利用本联盟机构知识库生成机构学术竞争力评估报告、机构成果展示平台等。

##### (三)建设目标

根据 CALIS 的相关说明,建设机构库联盟的目标一般可以归结为:实现机构内部知识产出的集中揭示和统一管理;实现机构知识产出的长期保存和持续利用;实现机构联盟之间的知识产出开放获取和共建共享,促进学术交流;通过知识产出的开放服务和利用,扩大和提升机构及科学家个人的学术影响和声望<sup>[6]</sup>。

而根据不同机构的实际情况,机构库联盟建设目标又可以分成不同层次,包括开放存取与自我存档,知识管理和数字科技产出的长期存档<sup>[6]</sup>。本文案例拟实现的目标以第一和第二层次为主,在联盟所产出的期刊、论文、学位论文、会议论文等正式发表成果基础上,加大对其他类型的科研产出如实验室报告、相关医疗记录进行搜集和保存,重点在于资源管理后的知识共享,机构库建设完成后根据上级机构的机构知识库建设情况考虑其资源的长期存档问题。

相较于其他大型机构知识库吸引诸多下级机构的科研人员主动参与知识成果的提交不同,本文的联盟机构知识库因其医药学科特色,除南京医科大学的基础研究外,重点在于附属医院的临床学术成果的利用。由于机构正式学术成果的有限性,数据上传工作主要由专门的图书馆人员完成,重要的是对于已经产生的学术成果的共享与利用上,尤其是实力较弱的机构会以较强实力机构为风向标,对于应用性和实践性较强的临床医学,能够转化为实际工作的知识储备是最受欢迎的,机构知识库内的资源反射出机构的研究方向和学科特点,在此基础上,深度挖掘和分析相关数据是此机构知识库管理和服务的要务。

### 三、联盟机构知识库建设的问题和解决方法

#### (一)软件的选择

OpenDOAR 中显示的软件和系统有 160 余种,其中业内广泛应用的构建机构知识库的传统软件有 DSpace、EPrints、Fedora 等<sup>[7]</sup>,近年来的新型开源 IR

软件的典型代表有 IR+、Zentity、Islandora 和 Hydra 等<sup>[8]</sup>。机构知识库建设过程中,开源知识库软件选择时系统的兼容性、易用性、可扩展性、资源管理功能,尤其是现代的新型多媒体资源(图片、MOOC 和微课)的保存与管理功能,都应考虑或做预留。而另外一个在实际工作中同样至关重要的因素是构建成本,包括机构知识库建设的人员技术成本和购置或开发知识库的经费。

除国字头的机构在自上而下组织建设机构知识库外,国内现有的中国知网、万方数据和重庆维普等资源数据库商,和如雨后春笋般出现的各种小型学术数据加工公司,面对机构知识库建设的大蛋糕也想分一杯羹,这些小公司通常会依托于知名大学的图书情报部门的知识管理背景。

南京医科大学图书馆及附属医院分馆联盟机构知识库建设中需要收集和提交的数据量,与国家级机构知识库或专业学术数据库的数据量相比较少,软件的选择应简洁并考虑经费成本,更为重要的是关注可持续发展,将能够建设并共享的资源融合到整体大环境或者更大系统中去<sup>[9]</sup>,故应选取国际通过的、质量有保证的软件,或者直接选择对应的小型公司,完成构建时则优先考虑系统的可扩展性。

#### (二)资源建设与收集

虽然本文案例指出了前期建设的大量资源上传工作由专门的馆员负责,但实际上越来越深度的学科服务和其他资源建设会分散人力资源,因此在可能的情况下,更合理更便捷的资源上传方式也在考虑范围内。

我国现有的科研机构包括高校、医院等,都会进行一定周期的科研产出统计与考核工作,这个周期通常为年度,而事业单位的职称评定工作也涉及到学术成果的提交,南京医科大学每年都有这样的学术产出统计工作,目前该工作由相关的科技处和人事处负责,这种机构内不同部门与图书馆的合作有极大的可能会被促成。其最直接的结果就是可以用机构内的行政规定和流程使得科研人员主动提交自己的学术成果,图书馆需要做的只是提供一个成熟的搭建好的平台,就可以直接享受到“上门”的资源建设,同时图书馆也能为科研人员的成果作出相关鉴定,达到双赢<sup>[10]</sup>。

充分利用网络 OA 资源也是非常重要的途径,现有的 OA 资源非常丰富且不乏精品,这些 OA 资源既可以在联盟机构知识库作相应的推广,搭建资源获取的桥梁。同时也可以批量收割,这里的收割并非简单的数据下载和存储,而是通过联盟机构知识

库的自有分类或主题系统深度加工,再推荐给科研人员。

### (三)使用和服务

医科院校及附属医院图书馆是师生及医护人员科研资源的最重要来源,但实际工作中,附属医院的图书馆因其地位、资金和人员配备的限制,发挥的作用有限,科研资源的查找和信息素养的培训工作在很大程度上都依赖于院校图书馆的支持,这一特点使院校图书馆在附院图书馆中自然而然地处于领导地位,对于联盟机构知识库的建设非常有利,但同时也对机构知识库资源利用和知识服务提出了更高要求。

首先是范围的扩大,联盟机构知识库的人员不仅要服务于院校师生,还要服务于多家附属医院的医护人员,后者的数量甚至已经超过前者。同时,附属医院的资源与知识建设、服务特点实际上与院校的基础研究也有较大区别,医护人员的资源与知识需求更为实际和深入,多伴有科室特点。以上都要求联盟机构知识建设过程中不仅要重视正式出版的学术论文,对于一些有特色的精品专著、手术视频、药物指南等多种类型的资源也要保证,使得联盟机构知识库真正发挥“智囊”的作用,而不是又一座“电子图书馆”而已<sup>[11]</sup>。

其次是时间的持久,联盟机构知识库的建设并非一个阶段性的项目,而是要伴随信息管理机构(现阶段的图书馆)长期存在的,它需要的并非某一任领导或某一个项目,而是需要机构生长过程中产生的资源和知识的不断“灌溉”。这种持续建设的要求要在项目启动时即向相关负责单位和负责人说明,并由专门的联盟机构组织相关事宜。

第三是数据挖掘功能,现有观点都将关注的重点放在机构知识成果产品的长期保存上,对于利用机构知识库内的资源所含有的“知识”没有足够的重视,笔者认为虽然这与现阶段的机构知识库建设初期的现实相符。但从长远来看,最终目标“全球机构知识库”目的绝非仅为了保存,其最终目标是为了更好的利用,这从各开源软件的高扩展性也能窥见一二。基础数据加工过程中的可扩展性,软件的可定制性,和数据表现形式的多样性,都是后期数据挖掘的重要保证。

最后是知识产权问题,这是不容商榷的法律问题,也是每个机构知识库建设中资源获取与提供的核心问题。此问题应一分为二来解决,对于已正式出版的知识成果,现有法律和其出版机构都能够保障其权益;而对于用户提交的非正式出版物的知识产

权保护,较为困难,机构知识库虽然为某一或联盟的机构所有并开放,但价值较高的非正式出版物可能会被窃取或抄袭,是每个科研人员在提交资源时的最大顾虑。故笔者认为现阶段机构知识库建设过程中的知识产权问题,是任何下属机构都无法通过个体努力解决的国家层面立法和政策问题,因而最好的途径就是依循 CALIS 等高级别的项目先例。

本文拟建的联盟机构知识库建设模式采用集中与分散相结合的建设模式。一方面,CALIS 和 JALIS 的相关机构知识库项目已经达到院校及图书馆层面,得到政策、资金和人力的支持,跟随大潮集中时间建设;另一方面,图书馆在院校与附属医院紧密联系和合作的基础上共同建设区域型医药学科特色的机构知识库,笔者认为可以作为样本在小型机构知识库建设中推广,这种以点带面且独具特色的建设模式具有较好的实践性,分散建设最终集中呈现的模式也适应我国国情。

### 参考文献

- [1] OpenDOAR [EB/OL]. [2016-05-25]. <http://www.opendoar.org>
- [2] 张晓丹,张志平. 科研机构的机构知识库构建研究[J]. 现代情报,2014,34(7):45-49
- [3] 何琳,吴芬. 我国机构库发展对策探讨[J]. 图书馆论坛,2008,28(1):77-79
- [4] 中国科学院心理研究所机构知识库 (PSYCH OpenIR) [EB/OL]. [2016-05-25]. <http://ir.psych.ac.cn:8080/>
- [5] 周素芬. 国内机构库联盟建设的初步规划[J]. 图书馆学研究,2013(3):72-74
- [6] 马建霞,祝忠明,洪梅,等. 机构仓储系统规划与政策框架研究[J]. 图书馆理论与实践,2008(3):64-66
- [7] 王颖洁. 我国机构知识库模式构建与流程设计[J]. 图书情报工作,2008,52(4):104-107
- [8] 张旺强,祝忠明,卢利农. 几种典型新型开源机构知识库软件比较分析[J]. 现代图书情报技术,2014(2):17-24
- [9] 田丽君,张静鹏,芬兰 Doria 和 Theseus 联盟机构知识库建设模式及其启示[J]. 图书馆学研究,2014(5):37-41
- [10] 渠芳. 高校教学联合体机构知识库联盟建设研究——以徐州高校教学联合体为例[J]. 情报理论与实践,2010,33(11):83-85
- [11] 付伟棠,廖璠. 我国机构知识库资源建设调查[J]. 国家图书馆学刊,2014,23(2):64-72