

关于机构知识库构建的若干思考

李霞

(南京工业大学图书馆, 江苏南京 210009)

摘要: 概述机构知识库产生的背景和含义, 分析学术机构构建机构知识库的作用, 探讨目前机构知识库构建中存在的问题, 并提出构建机构知识库中可以采用的策略。认为机构知识库是由学术机构建立的, 用于收录并长期保存机构内部所产生的各种数字学术资源和某些相关的外部资源, 按照开放标准与相应的互操作协议, 供机构内外用户开放获取, 进行交流、工作的数据库服务平台。机构知识库是数字学术环境中基于开放理念、以用户为中心的一种新的学术资源组织形式和学术交流方式。

关键词: 机构知识库; 学术资源; 学术机构; 开放获取

中图分类号: G205.7 **文献标识码:** A **DOI:** 10.3772/j.issn.1674-1544.2009.06.009

1 引言

近年来, 科研人员竞相利用新技术进行科学研究, 使得科学发现的速度大大加快, 科研成果的数量急剧增加, 学术机构产生了大量形式多样的原生数字信息资源 (即已发表或未发表的期刊论文、会议论文、学位论文、科技报告、文件、教学课件、数据库、声像资料、动画、图片等, 以文本、图像、音频、视频及其他不同格式出现的数字化内容)^[1]。但由于传统学术出版体系的局限, 这些重要的学术资源因出版周期及版面限制等原因, 往往滞后发表甚至无法发表, 加之印刷型书刊大幅度涨价, 使得学术机构图书馆采购印刷型书刊的成本上升, 不少国家的学术机构没有经费购买足够的学术期刊。科研人员受学术机构资助的研究成果发表要向出版商付费, 科研人员利用学术参考期刊时要受种种限制, 传统的以纸质出版物 (如期刊) 为载体的学术交流方式, 明显无法满足科研人员在数字化环境下全面、准确、及时的科研需求^[2], 由此引起了全世界范围内的学术交流危机。正是基于原生数字信息资源的急剧膨胀、印刷型书刊价格的暴涨、传统学术出版体系的不合理性以及科研

人员对学术资源的迫切需求等多方面的原因, 伴随着网络信息技术的发展, 出现了一种基于开放获取理念的新的学术资源组织形式和学术交流方式——机构知识库。

机构知识库, 简称 IR, 又称为机构仓储、机构资料库、机构信息库、成果数据库、机构典藏库等。不同学者对机构知识库含义的认识与表述不尽相同, 迄今为止, 还没有统一的定义。笔者倾向于: 机构知识库是由一个或多个学术机构 (大学、大学共同体、研究机构等) 建立的, 用于收录并长期保存机构内部所产生的各种数字学术资源和某些相关的外部资源, 按照开放标准与相应的互操作协议, 供机构内外用户开放获取, 并进行交流、工作的数据库服务平台^[1]。据此, 可以这样来理解 IR: 第一, 构建的主体是学术机构, 建立和运行均以特定机构为轴心; 第二, 是一个基于网络的数据库; 第三, 存放的是学术资源; 第四, 数据记录具有累积性和长期性; 第五, 具有开放和互操作的性质, 是学术资源开放利用的工具; 第六, 用来进行学术资源的收集、存储和发布, 是一个大学或研究机构学术传播过程的组成部分; 第七, 是数字资源长期保存的工具, 收藏的是以任何形态出现的机

作者简介: 李霞 (1964-), 女, 南京工业大学图书馆馆员, 主要研究方向是图书情报学。

收稿日期: 2009年4月10日。

构内部产生的学术资源^[3]。

2 构建机构知识库的作用

2.1 全面收集学术信息 长期保存数字资源

在大学和研究机构里,教师与研究人员普遍使用计算机来处理各种信息,产生的原生数字信息资源通常分散保存于个人或院系、研究所(室)的计算机上。这种分散保存的方式,往往因时间的推移或外部条件的变化等原因丢失数字资源^[3]。构建机构知识库可以从机构整体的角度出发,将机构中这些分散存储的学术信息全面收集起来,经过适当的处理,永久完整地保存,克服各部门以及个人分散保存信息的弊端和风险,为大学或研究机构提供一个收集和长期保存学术资源的安全场所。

2.2 传播学术资源 拓宽学术交流渠道

出版滞后、价格飞涨、版权许可等传统学术出版体系的局限,使得学术成果无法快速、准确地反映出来,影响了学术传播,限制了学术交流。构建机构知识库,可以带来学术传播方式上的变革,拓宽学术交流渠道。通过 IR,借助于互联网这一便捷的运行平台,原先无法进入正式出版途径,无法及时、有效地融入学术交流过程中的大量学术资源,得以出版、发布、揭示、检索等。科研人员能够跨越传统学术出版体系的局限,低成本、高效益、全面准确、灵活方便地实时获取信息资源,加速学术研究成果的传播,及时交流学术领域内科技信息与研究成果。

2.3 知识有效管理 资源共享利用

机构知识库通过相应的技术手段对机构内不同格式的学术资源进行统一分类、标引、加工、存储、建库和保存,避免了学术资源的散乱存在,实现了知识的有效管理,方便机构内用户快捷地检索和利用。在此基础上,在权利义务平衡的前提下,通过开放的获取机制和相关的互操作协议,不同的大学和研究机构可以相互开放学术资源^[3],交流彼此的学术成果,实现学术机构内、学术机构之间、各地区乃至国家层面数字学术资源的共享利用。

2.4 发挥教育功能 提高教学质量

IR 存储的是机构内部的各类学术资源,对于

教育机构还包括丰富的教学信息资源。IR 构建后可充分发挥教育功能,发展新形势下各种在线教育。通过 IR,学生可以利用各类课件进行自主学习、作业和考试,教师可以交流、完善教学课件和教学方法,有助于师生提高教学质量。

2.5 展示机构学术成果 提升机构学术地位

在传统的学术交流体系中,一个机构成员创作的各种学术成果往往分散在众多期刊、数据库、网站中,不能从机构整体的角度为人所知,削弱了机构在科研中实力的显示度^[4]。构建机构知识库,可以使原先分散存储的机构成员的产品得以汇聚,学术机构的学术成果得到全面、系统、完整展示。IR 对学术机构科研人员全部科研活动的集中反映,还可以具体体现一个学术机构对科学研究、社会发展等多方面的贡献,彰显其科研实力与水平^[5],有助于扩大学术机构在学术界的知名度和影响力,提升机构的学术地位。

2.6 分析机构科研进展 评价机构科研活动

机构知识库是机构成果和能力的展示窗口,按时间序列保存了机构的研究历史,展现了机构的发展脉络,可以作为分析机构科研发展历史、预测机构科研发展趋势和评价机构学术能力的工具。同时,由于机构知识库全面保存了机构中科研人员的科研成果,为定量评价科研人员的绩效提供了客观的原始数据,方便学术机构了解科研人员的研究成果,可作为评价机构科研活动的有效工具^[6],实现对科研人员科研业绩的管理。

3 构建机构知识库面临的问题

3.1 政策问题

国家和机构所采取的政策对构建机构知识库具有极其重要的影响。但目前我国信息资源公共政策总体来说水平较低,效率较差^[7]。机构知识库作为一种由学术界发起的新生事物,国内还没有正式的国家政策或机构内的政策,未能对机构成员研究成果收集、保存、维护以及提供使用等方面做出规定,无法指导或支持机构知识库的部署和发展^[8]。机构知识库在构建过程中如不能做到有章可循、有法可依,就难免失去其规范性和系统性。缺乏相关的支持政策,成为机构知识库构建的主要问题,势必影响机构知识库的未来可持续发展。如果国家能够制

定政策,明确研究成果集中保存,并在一定条件下开放获取,那么必将大力推动机构知识库的发展。

3.2 法律问题

(1) 版权许可问题^[9]。机构知识库收录资源类型的多样性表明机构知识库的版权主体和版权客体法律关系具有复杂性。版权归作者所有的文献信息,作者可授予收录其文献信息的机构知识库拥有“非专属授权”传播权利。但作者提交到 IR 中的在正式期刊上经过同行评议发表的学术论文,如果没有得到出版商版权许可,IR 收录后就会构成侵权。

(2) 知识产权保护问题。目前,我国还没有自主研发应用于 IR 的开放源码软件,IR 构建多是利用国外相对已经发展成熟的开放源码软件。这些开放源码软件虽然允许修改代码,但我们在利用时应注意到软件的知识产权^[9]。同时,资源提供者提交的内容是否违反或侵犯其他人的版权、专利权和商标权等;机构是否有权发布、存储、移植、复制、删除或重新排列资源提供者提交的资源;资源获取者是否可以免费复制、分发、显示从机构库获取的资源,并由此生成派生资源,获取的免费资源是否用于商业目的等,都是和知识产权保护有关的法律问题。

由于立法滞后于技术的迅猛发展,随着新技术在机构知识库的使用,机构知识库的构建与服务将面临一些新的法律问题^[9]。

3.3 质量控制与评价问题

机构知识库不像传统出版具有全面、系统的质量控制评价体系,缺乏现有学术出版过程中进行质量控制的同行评议过程或^[9]专家评审制度。本着文责自负的原则,机构知识库只对作者提交论文做一定的形式审核(如特定格式、学术规范等),不做实质性的内容审核,其开放性又使得用户可以不受限制地上传论文,因而论文质量得不到有效控制^[7]。加之 IR 收录的学术资源既有已在传统学术期刊上发表的学术性论文,又有大量未发表的预印本论文、学习资料、工作文件等^[1],而不同类型的资源又有着不同的质量评定方法,如何针对不同类型的资源建立相应的质量控制和评价体制,以保证机构知识库的质量^[10],是机构知识库构建中必须面对的一个问题。

3.4 经费问题

运营经费也是 IR 构建发展需要解决的主要问题之一。机构知识库要长期存在并发展壮大,在起草机构知识库的政策、制定许可协议、改进指导方针、选择软硬件、迁移数据、开发和升级系统、管理和维护系统、选用网络服务器和数据库、培训用户、支付员工薪水等方面必须有足够的经费支持^[9]。由于机构知识库是以开放获取的方式提供服务的,每一个用户可以得到机构知识库里的记录,而无须支付额外的费用,其运营所需要的费用自然转嫁给学术机构^[11]。高昂的开发、运行费用可能降低决策者对 IR 的信任值和支持度,部分工作无法按计划进行,这将给 IR 的平稳发展及在世界各国尤其是发展中国家的普及带来负面影响^[1]。

3.5 认识问题

科研成果是整个社会的财富,科研论文的完全免费公开共享已经成为世界性的潮流,开放获取的理念得到了科研工作者的广泛支持。然而,开放获取运动在我国刚刚兴起,开放获取的理念还未得到广泛传播。在国内,人们对于“开放获取”和“机构知识库”的了解还不是很深,对于把学术成果存放在 IR 中的好处也不很了解。即使有所了解的人,有时也因为机构内部缺乏信任和合作的环境,而产生各种疑虑,甚至有抵触情绪。

4 构建机构知识库的策略

4.1 发挥学术机构图书馆的作用

机构知识库是学术机构保存本机构学术研究成果和资源,实现机构内部交流的重要方式。作为机构的知识中心、学术资源的集散地和传播者,图书馆承担着信息资源的搜集、管理、传播、开发和服务的重任。在构建机构知识库,明确组织管理责任单位时,应充分发挥学术机构图书馆的作用。

第一,国内外已有的机构知识库建设实践表明,图书馆是机构知识库主要创建力量。绝大多数的机构知识库是由图书馆负责建设与维护的,许多成功的机构知识库也都是由图书馆来领导并运行管理的。机构知识库的发展离不开图书馆,尤其是大学图书馆的引导、参与及支持^[12]。

第二,图书馆在机构知识库构建中有着自身优势。一方面,图书馆员对图书馆的财政压力与信

息获取障碍比其他人有着更深刻与直接的认识,因此具有建设机构知识库的意识和需求;另一方面,图书馆有承担和指挥机构知识库建设的能力,拥有其他部门所缺乏的机构知识库建设中所需要的专业技术,如元数据的创造与维护等,在机构知识库的方案设计、技术实现和内容维护等方面可以发挥重要作用^[2]。此外,参与数字图书馆建设,积累了丰富的经验,因而能够更深刻地从机构知识库的建设中发现与解决问题^[13]。

第三,图书馆本身具有保存学术资源、传播学术成果的职能。如果运用图书馆现有的人力资源与物力资源,采用适当的规划与管理策略,充分开发掌握在广大教学和科研人员手中的原生数字信息资源,实现对本机构学术资源的收集、发布、加工、组织、管理、保存、交流和提供利用,同时依靠已有的相关专业基础进行相应的用户培训工作,就可以为机构知识库的开发、运行、管理和维护等节约经费,降低机构知识库的运营成本。

4.2 配置高水准的工作团队

作为生产力的第一要素,人是机构知识库构建最重要的资源。为保证各项工作的顺利进行,在建设机构知识库前,需要配置决策领导者、系统技术专业人员、用户服务专业人员和法律工作者共同组成的对学术交流、技术应用或对推动机构知识资产使用具有热情的高水准的工作团队。

在团队中,决策领导者对外负责与学术机构管理者和政策制定者沟通,协调与学术机构内其他部门的关系,取得机构领导和相关职能部门的理解和支持;对内负责合理分配工作任务,以保证项目的推进。系统技术专业人员主要负责机构知识库系统的技术方面,包括数据库设计、网络的管理与维护、系统延伸功能的开发等。用户服务专业人员主要负责与用户的沟通和技术支持,特别是对为机构知识库提供内容的用户给予技术支持。法律工作者负责解决机构知识库建设中相关的法律问题。

所有成员既要各司其职,充分发挥专业技能,又要团结协作,共同致力于机构知识库的建设,以保证机构知识库的顺利建成和正常运行。

4.3 制订合适的机构知识库建设计划

制订合适的计划,对于按时、有效地完成机构知识库建设具有极其重要的作用。在计划制订前,工作团队要有目的地进行详细而深入的调查,摸清学术机构当前所拥有的知识资源的数量、类型和保存方式,计算机网络及其IT设备资源,机构内成员对数字资源存储、利用及其存在问题的认知程度,所需要的知识资源等情况。获得合理设计机构知识库的第一手资料。

在此基础上,通过需求分析,制订合适的机构知识库建设计划。计划通常由开始阶段、发展阶段和成熟阶段组成,对不同的阶段应设立不同的建设目标和服务要求^[3]。初期应将目标限制在较小和可控的资源内容、用户对象和服务功能上,在取得成果、取得共识后逐步提升建设目标,增加建设内容,扩大用户范围,扩充服务功能。

在IR构建计划中,要有明确的工作进程表和时间安排。建立内容提交、发布和维护管理的制度,确定对用户是全免费还是需要部分收费的政策。制定IR的短期目标和长期目标,确定优先服务功能和服务对象。确定图书馆与大学或研究机构内各单位之间对于IR构建的责任^[3]。

4.4 完善机构知识库的相关政策

IR构建是一项长期、艰巨的系统工程。需要通过国家或机构制订相关政策来加以保障,为IR的建设创造一个良好的环境。

政府应该利用我国科研资助机制和学术出版机制的特殊性,给机构知识库的构建以政策支持。要求凡是由国家科研经费资助完成的科研成果在发表后必须在规定时间内存入机构知识库^[3];要求学术机构建立或者合作建立机构知识库,并将机构知识库建设和使用情况纳入机构尤其是大学科研评价体系;以立法的方式制定有利于学术交流的知识产权保护等政策,规范IR建设中的技术标准,从而促进机构知识库的发展。

学术机构应该根据自身特色、专业设置、重点学科等,从IR的运行特点入手,制定建设IR的相关政策。机构知识库建设者可以参照相关的网络知识产权的立法内容,参考国内外其他学术机构知识库的知识产权保护政策和国外开放获取方面关于知识产权保护的政策,制定本机构知识产权

保护政策,对机构知识库中信息资源的收集、传播和利用进行法律和政策上的引导。利用相关网站检索出版商对作者自我典藏所制订的版权政策,收集与提供各出版社对自我典藏式版权转让协议的信息。在向公众开放信息的同时尽可能地保护知识产权所有者的利益,至少须确保 IR 不是有意侵犯版权或其他知识产权^[1]。

与机构知识库相关的建设政策是 IR 发展的重要推动力^[9],直接影响到机构知识库的学术价值和持续发展。机构知识库建设者应该根据机构的性质和机构库的建设目的通盘考虑:制订机构知识库的组织政策,确定其组织结构及相应的职责,赋予参建的部门和人员管理权力,构建机构知识库的基本框架和运行规则;制订学术资源收录政策,明确可以收录的资源范围、种类,确保机构知识库的学术质量,保障机构知识库建设稳步、长久地发展;制订学术资源提交政策,对学术资源提交者、审核者、提交程序、提交内容的格式、不同类型产品的提交标准等做出相应的规定,保证提交的学术资源达到一定的规范,做好学术资源的质量控制工作;制订标准化政策,实施学术资源提交的元数据标准、互操作协议标准以及文档格式标准等,便于资源的共建共享与交流利用;制订权限控制与管理政策,根据资源提交者的要求和不同的资源类型对用户的权限设置控制,为不同的用户设置不同功能的权限级别,实现权限分级及其控制与管理^[10];制订激励政策,在职称职务评定、项目申请等工作中,承认发表在已被学术界认可的网络载体上的学术成果,机构知识库中开放获取的科研论文可以纳入年终考核,定期通过机构知识库平台发布奖励信息,鼓励科研人员将科研成果提交给机构知识库,激发机构成员呈缴学术资源的积极性和主动性。

完善的政策既有助于解决 IR 建设过程中遇到的各种性质不清、责任不明等问题,也有助于 IR 建设中、建成后的各种权利与义务关系问题的解决。制订政策时越科学具体,构建机构知识库时就越顺利^[1]。

4.5 采取有效的质量控制措施

质量是机构知识库建设的生命线。机构知识库未来要持续发展,必须针对不同类型的资源,采

取与之相适应的行之有效的质量控制措施。

机构知识库的收录政策决定了其所收录的学术资源的内容既有经过同行评议过程的,也有未经同行评议过程的。对此,机构知识库可以实行不同的文件类型政策,将经过同行评议的论文和还没有经过评议的论文分开保存,以确保在一定程度上上的质量控制^[14]。

机构鼓励作者自存档于机构知识库,但对作者提交到机构知识库的学术资源,机构知识库只做存档,只有通过了评审的学术资源,才被机构知识库所承认并提供开放获取。评审的内容应包括作品是否符合该机构服务对象的需求,作品是否属于所要存档的文献类型,作品是否有版权许可,作品是否符合一定的格式,作品的元数据是否符合一定标准等,以此来减少作者对质量控制的担忧。

对通过评审存储到机构知识库提供开放获取的学术资源,机构知识库还应建立反馈机制,通过用户评议,及时删除冗余并及时更新,以保证机构知识库的整体质量^[11]。

4.6 加大宣传力度

机构知识库要顺利构建与运作,宣传问题不容忽视。工作团队应利用各种场所,借助各种途径,在本机构大力宣传 IR 对于教学、科研和学术传播的重要意义与作用^[12],引起大家的关注和讨论。

具体来说,机构可以通过编印简明手册、不断发送电子邮件、在图书馆网站开设专栏、日常用户培训、文献检索课程或专题讲座以及向作者演示其文献可被搜索引擎所寻获等行之有效的方式扩大对机构知识库的宣传。通过大力宣传,普及 IR 知识,使更多科研人员了解开放获取的运作模式和优点,明白机构知识库既能够提升其在学科领域的知名度,提高其文章的引用率,扩大其学术成果的传播和影响范围,又能够方便地访问和免费获得,有效降低获取障碍^[14]。一旦作者认识到他们的文献存放在一个可开放获取的机构知识库,能够对其产生很大的影响并从中受益,高校教师或研究机构的科研人员就会积极参与,主动、及时地把他们的文献存放在机构知识库中,支持和推动机构知识库的建设。

5 结 语

基于开放理念的机构知识库是以用户为中心的一种新的学术资源组织形式和学术交流方式。在数字环境下,对于保存学术资源、促进学术传播、拓宽学术交流渠道、共享利用信息资源等方面具有重要的作用。综合国外机构知识库的发展态势,我们可以预见机构知识库将成为数字化时代学术研究的一项关键性的基础设施^[15],而且作为基础设施的机构知识库之间的跨国的、跨地区的联合将越来越多。在我国,机构知识库的建设刚刚起步,相关管理部门应出台一些相关政策,规范、支持和推进机构知识库的建设工作。大学和科研机构也要积极行动起来,借鉴国外机构知识库的建设经验,构建我国大学和科研机构的机构知识库,促进国内学术研究机构之间的学术交流,并参与国际间的学术交流和信息共享^[15]。

参考文献

- [1] 陈韬. 北邮机构知识库原型系统设计研究 [D]. 北京: 北京邮电大学, 2007.
- [2] 陈钦琳. 学术交流与知识共享的新平台——机构知识库[J]. 现代情报, 2006(9): 150-151, 156.
- [3] 赵继海. 机构知识库: 数字图书馆发展的新领域 [J]. 中国图书馆学报, 2006(2): 33-36, 50.
- [4] 常唯. 机构知识库: 数字科研时代一种新的学术交流与知识共享方式[J]. 图书馆杂志, 2005(3): 16-19.
- [5] 曾英姿. 机构知识库属性与功能之理论探讨 [J]. 现代情报, 2007(10): 16-18, 21.
- [6] 常唯. 数字学术环境中的机构知识库探析 [J]. 图书情报工作, 2006(7): 45-50.
- [7] 郭淑艳, 王东艳. 机构知识库构建研究 [J]. 图书馆学研究, 2007(5): 9-12, 8.
- [8] 马漫江. 机构知识库: 学术交流与资源共享新模式[J]. 高校图书馆工作, 2007(1): 10-13.
- [9] 邓君. 国内机构知识库研究进展 [J]. 图书与情报, 2007(5): 37-42.
- [10] 范亚芳, 渠芳. 对我国高校机构知识库建设的几点思考[J]. 情报杂志, 2007(9): 132-134.
- [11] 洪梅, 马建霞. 机构知识库建设机制初探 [J]. 情报杂志, 2007(8): 37-39.
- [12] 尤春玲. 高校图书馆机构知识库构建初探 [J]. 科技情报开发与经济, 2007(28): 83-84.
- [13] 柯平. 机构知识库大学图书馆的新平台 [J]. 新世纪图书馆, 2007(1): 5-8, 99.
- [14] 柯平, 王颖洁. 机构知识库的发展研究 [J]. 图书馆论坛, 2006(6): 243-248.
- [15] 李枫林, 赵雪芹, 胡吉明. 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.

Speculations on the Construction of Institutional Repositories

Li Xia

(Library, Nanjing University of Technology, Nanjing 210009)

Abstract: This article expounds the definition of institutional repositories under the background of open access, analyzes the role of institutional repositories in academic institution. It further discusses the problems that need to be solved in the establishment of institutional repository, and introduces some strategies to set up it. Institutional repositories are founded by academic organizations. The repositories take down various digital academic resources and some relevant outer resources which have been restored in these organizations. Institutional repositories are opening accesses, communicating and working Services Databases by interaction protocols to provide the data for opening accesses from outer users. It's a totally new way for academic resources organization and academic communication oriented users.

Keywords: institutional repository, academic resources, academic institution, open access