

国内外学位论文开放获取平台的网络调研与分析

张雅男

(天津社会科学院图书馆, 天津 300191)

摘要: 学位论文开放获取是开放获取研究的重要方面, 许多国家都在发展学位论文OA工程。本文以网络调研法, 对美国、英国、澳大利亚、德国等十几个国家及我国港澳台地区的学位论文OA平台进行实证研究, 归纳总结出这些平台的一些建设特征, 并结合我国实际, 提出加强以图书馆为核心的学位论文OA建设主导力量、强化OA资源的区域性和学位性、碎片化切割管理元数据等多方面建议, 以期为我国学位论文开放获取建设提供借鉴。

关键词: 学位论文; 开放获取; 网络调研

中图分类号: G250

DOI: 10.3772/j.issn.1673-2286.2016.3.009

1 引言

开放获取(Open Access, OA)始于20世纪90年代, 是国际科技界、学术界、出版界、信息传播界为打破商业出版者对科研信息的垄断, 实现学术资源共享的运动^[1]。以arXiv(由PualGinsparg建立的物理学、数学、计算机科学和非线性科学领域遵守OAI协议的电子预印本资源库, 是免费的联机学术成果最早的资源库之一)为代表的早期免费联机学术资源库的发展促使开放获取运动得到全球重视^[2]。2001年布达佩斯开放获取宣言(Budapest Open Access Initiative, BOAI)明确给出开放获取的定义:“文献可以在公共网络上被免费获取, 允许任何用户对该文献的全文信息进行阅读、下载、复制、分发、打印、检索、超链接, 支持爬行器收割并建立本地索引、用作软件的输入数据、用于其他任何法律允许的用途”^[3], 被认为是国际开放获取的开始。之后各国纷纷建立全国性开放获取计划, 普遍采取强制性OA自存储政策, 推动开放获取的实践发展^[2]。

学位论文是著者提出其研究和发现, 并自荐报请授予学位或专门职位资格的文献^[4], 博、硕士学位论文更是了解国内外科技动态和学科前沿的重要信息源。由于学位论文绝大多数不公开发表或出版^[5], 科研人员在获取国内外学位论文方面存在困难。开放获取运动以来, 国

际学位论文OA平台迅速发展, 截至2015年7月, 全球已有1607个机构知识库存有博硕士学位论文。学位论文的开放获取, 加快了学术交流共享, 极大地方便了科研人员, 正成为开放资源的重要组成部分^[6]。在此背景下, 为进一步了解国际学位论文OA平台的建设情况, 本文对美国、英国、澳大利亚等十余国家和地区的学位论文OA仓储展开了网络调研, 就其基本情况和建设特征进行阐述, 以期为我国学位论文开放获取平台建设提供启示。

2 研究综述

国内外开放获取研究涉及政策法规、产权保护、组织机制和技术结构等多方面。

2.1 政策法规

政策法规的制定是开放获取有效推进的重要动力, 许多国家相继制定并实施了一系列相关政策。2005-2008年, 美国国家卫生研究院(NH)、国家自然科学基金、麻省理工大学、加州大学和哈佛大学相继颁布了开放获取政策草案, NH明确要求研究人员在论文公开发表6个月后, 将论文提交给PubMed Central供公

众免费使用^[7]；英国研究理事会（RUCK）提出了公共资助研究成果应面向公众接受有效使用、通过有效机制确保所发表研究成果的质量、研究成果应长期保存等四项原则；德国研究基金会2006年提出开放获取政策草案，要求受其资助的研究人员将研究成果提交至开放获取机构保存^[8]；中国科学院、国家自然科学基金委员会2014年发布政策声明，中科院所各类公共资助科研项目所产生的论文、国家自然科学基金会全部或部分资助科研项目产生的论文，将在论文发表后12个月内实施开放获取。就学位论文而言，各国普遍实施的学位论文呈缴制度是其开放资源建设的政策保障^[4]。ROARMAP (Registry of Open Access Repositories Mandatory Archiving Policies) 平台显示，全世界已有107家机构实施了学位论文强制性开放政策。其中，美国、英国、澳大利亚和意大利的机构数量占到总量的73%。我国的香港大学和香港理工大学也实行了强制性政策^[9]。

2.2 产权保护

开放获取带来信息共享的同时，也引发知识产权保护问题。目前国际上普遍采取的产权保护方法有两种，一是借鉴知识共享协议（Creative Commons, CC协议），在便于使用者获得作者授权的同时实现对网络作品著作权的保护；二是技术处理法，即通过用户时间限制、软件使用限制、文档加载水印等计算机网络技术实现开放获取与知识产权的有效平衡^[10]。具体到学位论文，国外多数大学的开放获取平台都在CC-BY协议框架或更严格的CC-BY-NC-ND协议框架下运行，对学位论文元数据收割、再利用的授权或限制性条款没有明确规定，也有一些大学对平台使用者提出了严格的限制性要求，例如下载全文需要注册登录等。另外，许多大学在建设学位论文开放获取平台时，也会注意不要侵犯他人的知识产权，它们会对成果提交者提出必须满足的知识产权要求^[11]。

2.3 组织机制

长期以来，国际开放获取知识库建设形成中心资源和分布式知识库大型合作的组织模式^[12]。例如：爱尔兰国内所有大学自建机构知识库，同时创建一个全国性研究平台，实现对各大学机构知识库内容的统一管

理；英国大型开放获取项目EThOS (Electronic Theses Online Service) 完全集成了国家所有电子博士学位论文，数据达40多万条；荷兰的NARCIS项目囊括了全国所有大学和一些研究机构的电子学位论文；新西兰多家国家级研究机构和政府部门联合创办了联盟机构库。此外，拉丁美洲的科学文档机构知识库联盟也覆盖巴西、哥伦比亚和墨西哥的多个机构知识库^[13]。

2.4 技术结构

开放获取技术的关键环节之一是开放存档计划元数据收割协议（Open Archives Initiative Protocol for Metadata Harvesting, OAI-PMH）。该协议明确规定了建立开放获取资源库的指南、协议，提供相应的工具链接，实现了检索系统对开放获取资源元数据的跨库检索和集成。2008年后，OAI-PMH升级为OAI-ORE^[2]。软件方面，最常用的两大开源软件是DSpace和Eprints^[13]。目前开放获取平台大都采用DSpace或Eprints软件，依照OAI协议（有些还通过必要的插件如wiki、EPrints等）建设而成。学位论文OA平台也主要通过OAI-PMH等开放协议自动收割参与机构库或访问机构FTP地址的方式实现数据采集。就学位论文的系统架构而言，主要包括服务器、存储、操作系统、持久化层、虚拟化五方面。服务器多选择开源的PostgreSQL、MySQL，存储架构可以是具备RAID5的直连型存储，操作系统则选择具备高稳定性的UNIX系统^[14]。

目前国内有关学位论文开放获取的研究涉及平台建设、现状研究、政策障碍等方面，但缺乏对国际学位论文开放获取平台的调查分析。本文对国际主要学位论文开放获取平台的概况、区域性特征、产权声明和权益管理、检索设置和题录信息等方面进行了详细研究，具有创新性。

3 网络调研与分析

3.1 调研对象

本文选取美国、英国、德国、澳大利亚、巴西等国及我国港澳台地区的共13个学位论文OA平台作为调研对象（详见表1）。其中，结合学位论文开放平台层次划分的观点^[4]，美国的NDLTD和巴西的BDTD属于全球性平台，英国的EThOS、瑞典的DiVA和澳大利亚的TROVE等属于国家或地区性平台，丹麦的E-thesis、日

表1 国际主要学位论文OA平台概况表

序号	国家	平台名称	数据量	全文获取情况	知识产权要求
1	美国	NDLTD (Networked Digital Library of Theses and Dissertations, 网络博硕士学位论文数字图书馆)	收录全球200多家大学图书馆、7个图书馆联盟、29个专业研究所的学位论文	文摘多全文少,分为“可无限下载”“有限制下载”“不能下载”三种。可下载论文支持一键下载,速度快	有知识产权声明
2	英国	EThOS (Electronic Theses Online Service, 电子论文在线服务)	集成英国121所高校图书馆,完全集成国家所有电子博士学位论文,包含30多万条数据,回溯到1800年	部分免费全文下载。有些论文可即时下载,有些需到成员机构库预约。需先将目标论文添加到“篮子”再下载。提供论文打印或CD/DVD等数码产品邮寄“增值”服务	有知识产权声明,明确要求不允许替他人下载论文
3	德国	Theses Online (学位论文在线)	收录德国国家图书馆、化学、计算机科学、教育学、物理和数学几个学习团体及5所德国大学的学位论文	点击目标论文题目即可获取PDF全文,方便快捷	有知识产权声明
4	荷兰	NARCIS (National Academic Research and Collaborations Information System, 国家学术研究和协作信息系统)	包括荷兰所有大学和一些研究机构的电子学位论文	部分可开放获取的论文标有“open access”,未收录全文的论文也不提供文摘,只提供来源链接	有知识产权声明。“当打印、解释、或者出版本组织数据,用户被要求写上如下内容:‘这些信息来源于NARCIS收集的信息,NARCIS是DANS的一个服务,DANS是KNAW和NOW的一个组织’。”
5	丹麦	E-thesis (Electronic Publications at the University of Helsinki, 赫尔辛基大学电子论文系统)	收录赫尔辛基大学80%的学位论文,约有4500篇博士学位论文	所有论文实现了在线获取,可直接快速查看目标论文的摘要和下载全文,快速便捷	有知识产权声明。“任何人都可以通过网络阅读和打印这些论文供他们自己所用。”
6	瑞典	DiVA (Digitala Vetenskapliga Arkivet, 学术档案数据库)	收录瑞典2万多篇文献,多数为博士学位论文,也有研究报告等其他文献。提供北欧28所大学自19世纪以来的部分文献	点击目标论文题目即可获取PDF全文,方便快捷	有知识产权声明。“是一个免费的全文存储系统,不论是谁,只要能利用互联网的就有权利阅读、下载、打印和复制DiVA出版的资源,再版或再次发行则需要作者许可”
7	西班牙	DIALNET	收录西班牙国内46所高校和研究机构的学位论文	全文为主。获取全文时,一些论文需链接到所在学校机构库才能下载全文,一些则可直接点击“full text”即时下载	有知识产权声明
8	澳大利亚	TROVE (澳大利亚国家图书馆特罗韦数据库)	收录澳大利亚40多所高校博硕士论文和部分新西兰大学学位论文	分为无限制网上免费获取“View online”和有限制网上获取“View online (access conditions)”。前者可直接快速获取全文,后者只能看到题录信息。获取速度快	有知识产权声明
9	肯尼亚	The union list of theses and dissertations (博硕士学位论文联合数据库)	收录肯尼亚37所大学和多个研究机构的研究生论文和学位信息等	以文摘为主,论文全文较少	有知识产权声明

续表

序号	国家	平台名称	数据量	全文获取情况	知识产权要求
10	日本	KUDD (Kyoto University Doctoral Dissertation Database, 京都大学博士学位论文数据库)	该数据库包含京都大学4.1万篇博士学位论文	多数学位论文支持开放获取, 以日语为主, 英语较少	有知识产权声明
11	巴西	BDTD (Biblioteca Digital de Tesese Dissertações, Digital Library for Theses and Dissertations, 博硕士学位论文数字图书馆)	截至2014年共有文献52 388条, 学位论文52 018条, 教授论文370篇。是仅次于NDLTD的全球第二学位论文存储机构	在目标论文上可即时点击获取PDF文献, 操作方便速度快	有知识产权声明
12	我国港澳台地区	HKLIS DTC (HKLIS Dissertations and Theses Collections, 港澳博硕士论文跨校集成系统)	收录7所香港高校和1所澳门高校的学位论文信息	论文自动链接到各大学机构知识库。部分论文提供无限制开放获取, 部分论文只限本校用户使用	有知识产权声明
13	我国台湾地区	National Digital Library of Theses and Dissertations in Taiwan (台湾博硕士论文知识加值系统)	截至2015年4月3日, 该系统共收集了台湾地区71所大学的学位论文。全文授权的博硕士论文达884 482篇, 其中326 835篇支持全文开放获取	部分论文标明可开放获取, 有的标注为不支持开放获取。一些论文有明确的开放获取过渡期提示	有知识产权声明。“同意遵守智慧财产权, 不得擅自复制、传输、改作、编辑或以其他任何形式、基于任何目的加以使用本网站提供的内容, 否则应负所有法律责任”

本的KUDD及德国的Theses Online属于大学或学术机构性平台, 选取对象具有一定代表性。

3.2 建设特征

3.2.1 发展模式呈现多样性

由表1可知, 目前英国、德国、荷兰、丹麦、西班牙、肯尼亚、日本等国的学位论文OA平台建设多面向本国高校或研究机构, 体现为国家性。瑞典DiVA“提供北欧28所大学自19世纪以来的部分文献”; 澳大利亚TROVE除本国40多所高校博硕士论文外还“收录部分新西兰大学学位论文资源”; 我国港澳地区的HKLIS DTC集成了香港和澳门两地区的博硕士学位论文。这些OA平台带有明显的区域性。而美国NDLTD“收录全球200多家大学图书馆、7个图书馆联盟、29个专业研究所的学位论文”, 巴西BDTD“是仅次于NDLTD的全球第二学位论文存储机构”, 实现的是全球学位论文的开放获取。所有这些国家和地区表明了目前国际学位论文OA仓储建设呈现出区域性、国家性和全球性的多层次、多模式发展态势。

3.2.2 图书馆扮演重要角色

一方面, 许多国家的学位论文开放获取平台都是在各高校或研究机构原有的图书馆学位论文资源建设的基础上整合而成, 即多以集成建设图书馆学位论文资源为主要形式。例如, 美国NDLTD收录有全球200多家大学图书馆、7个图书馆联盟的学位论文; 英国ETHOS集成国内121所高校图书馆学位论文; 德国Theses Online含有国家图书馆及5所德国大学图书馆的学位论文资源; 我国港澳地区的HKLIS DTC收录有香港和澳门8所大学图书馆目录和机构知识库里所有博硕士学位论文。另一方面, 多数高校学位论文开放获取强制性政策的实施主要由图书馆负责。图书馆同时要负责提交平台建设、操作指南提交、学位论文收割、开放获取政策解释、一般版权问题咨询、实施效果监控和反馈等事宜^[9]。由此可见, 图书馆在全球学位论文开放获取建设中扮演着重要角色。

3.2.3 元数据加工精细化

调研发现, 国际主要学位论文OA平台的元数据加工都非常精细, 主要体现在论文题录信息和检索项目

的详细设置上。各国学位论文题录信息除包含作者、题目、授予单位、论文时间、全文获取性、关键词、有无摘要、论文类型、语言等基本要素外,还涉及所属学校、作者现属单位、作者邮箱、管理员、监督员、学科、总页数、答辩日期及使用条款规定等内容;检索项则主要有作者及其ID、监管者及其ID、研究机构及ID、题名、主题、短语、关键词、摘要、参考文献、URL、语言种类、论文年限、论文类型、学科领域、基金、提交日期、签发日期、所属图书馆及可获取范围等多种。同时,许多学位论文OA平台设置兼顾个性化检索项目,如美国NDLTD的“全球资源定位器(URL)”、荷兰NARCIS的“数据集”、瑞典DiVA的“最后更新日期”、丹麦E-thesis的“签发日期”等。对元数据进行深加工,进而实现开放获取平台良好的检索功能和详细的题录信息展示,是国外学位论文OA平台的共性。

3.2.4 重视产权保护,使用权益各异

由于一些学位论文包含专利、创意等信息,或具有公开出版的需求^[9],给予明确的知识产权声明是各国学位论文OA平台建设的共识,但内容表述各有不同。例如,英国EThOS明确要求不允许替他人下载论文;荷兰NARCIS要求“当打印、解释或者出版本组织数据,用户被要求写上如下内容:‘这些信息来源于NARCIS收集的信息,NARCIS是DANS的一个服务,DANS是KNAW和NOW的一个组织’”;丹麦E-thesis则声明“任何人都可以通过网络阅读和打印这些论文供他们自己所用”;我国台湾博硕士论文知识增值系统明确标注“同意遵守智慧财产权,不得擅自重制、传输、改作、编辑或以其他任何形式、基于任何目的加以使用本网站提供的内容,否则应负所有法律责任”。

此外,全文、文摘等的可获取性差别较大。美国NDLTD以文摘为主,只有少量全文,分为“可无限下载”“有限制下载”“不能下载”三种级别,网站支持免登录使用;英国EThOS文章部分免费、部分需支付费用,“检索浏览无需登录,下载全文需登录”;澳大利亚TROVE分为无限制网上免费获取“View online”和有限制网上获取“View online (access conditions)”,前者可直接快速获取论文全文,后者只提供论文基本题录信息,但“需各机构专属用户名和登录口令”;我国港澳地区的HKLIS DTC“部分论文提供无限制开放获取,部分论文只限本校用户使用”;台湾地区博硕士论

文知识增值系统则“部分论文标明可开放获取,有的标注为不支持开放获取。一些论文有明确的开放获取过渡期提示”。据有关研究,学位论文开放获取的权益分为四种:完全公开、限制性公开、保密论文和混合级别^[9],本调研基本体现了这一点。

4 启示与建议

4.1 加强以图书馆为核心的学位论文开放获取建设主导力量

学位论文数据主要产生于各高校及其他研究生培养机构,这些机构近年来逐渐完善的学位论文呈缴制度保障了学位论文OA仓储的主要数据来源,而高校图书馆及其他机构图书馆普遍承担着学位论文的收录工作,这就使得图书馆具有了建设学位论文OA平台的首要条件。我国学位论文开放获取建设应加强图书馆的核心作用。图书馆也要明确这一点,一方面做好学位论文元数据加工工作,尽可能多地实现本校、本机构学位论文开放获取化;另一方面要积极搜集更多国际国内学位论文OA网址导航,从资源建设和信息导航两方面为用户的学位论文OA资源利用提供服务。

4.2 内容建设扩大区域性,强化学位性

目前我国学位论文开放获取建设表现出数量少、主要以本国学位论文为主的特点,而国际学位论文OA建设正向着突破本国、面向区域的大趋势发展。因此,我国的学位论文开放获取工程也应放眼国际,面向全球。例如,可以收录东亚各国或其他区域的学位论文进行开放获取。在内容方面,多数国家的学位论文OA平台以博硕士学位论文为主,如英国的EThOS和日本的KUDD主要收录博士学位论文,美国的NDLTD和我国港澳地区的HKLIS DTC及台湾博硕士论文知识增值系统同时收录博硕士学位论文。也有一些国家的学位论文OA仓储收录有非学位论文资源,如巴西的BDTD除包含学位论文外,还收录教授论文,瑞典的DiVA收录有研究报告、文章等学位论文以外的其他资源。值得注意的是,包含学位论文以外文献资源的OA平台,在检索时会给以获取学位论文资源为目的的用户造成干扰,从而降低平台效率。因此,我国在建设学位论文开放获取平台时,应对内容进行严格控制,只收录博硕

士学位论文,强化平台的“学位”性,提高使用效率。

4.3 碎片化元数据切割,细化题录信息和检索项

元数据加工是学位论文开放获取平台建设的基础性工作,对检索功能、论文信息的展示起着决定性作用。由前文可知,主要国际学位论文OA平台大都对学位论文元数据进行了碎片化切割,实现了多层次多角度的检索路径及详尽的论文题录信息展示,极大地方便了用户对学位论文OA资源的使用。我们在进行学位论文开放获取建设时,要尽可能详细地对元数据进行切割,提供作者、题名、主题、研究机构、论文年限、论文类型等基本元数据信息及全球资源定位器、学科领域、导师信息等其他内容,实现检索项和论文题录信息的细化。

4.4 强化产权意识,加强使用权益管理

学位论文开放获取建设必须重视知识产权问题。目前国际上大多数学位论文开放获取网站都有明确的产权声明,其内容大致包括开放获取学位论文来源解释、对论文免费使用形式的规定、尊重知识产权的承诺及禁止行为的规定等几个方面。不同国家、不同开放获取平台的规定各不相同,但都鼓励更多的论文作者积极上传论文并支持开放获取。除产权申明外,对开放获取对象权益的管理也十分重要,使用权限分为多个级别。我国在OA平台建设首先应根据各平台实际情况,制定并显著标识知识产权声明,其次要依据作者意愿或其他开放获取政策,对学位论文进行完全公开、限制性公开、保密论文和混合级别四个级别的权益设置,利用IP限制、加载水印、知识产权保护期到期提醒等多种方式进行

管理,实现学位论文开放获取平台的健康有效运行。

参考文献

- [1] 肖冬梅.开放存取运动缘何蓬勃兴起[J].图书情报工作,2006(5):128-131.
- [2] 初景利.开放获取的发展与推动因素[J].图书馆论坛,2006(12):238-242.
- [3] Budapest Open Access Initiative [EB/OL]. [2005-11-28]. <http://www.soros.org/openaccess/read.shtml>.
- [4] 丁道劲,沈丹.我国学位论文开放资源建设方式研究[J].数字图书馆论坛,2015(8):32-36.
- [5] 何怡.中外网上学位论文数据库的检索与利用[J].图书馆工作与研究,2011(2):44-45.
- [6] Search or Browse for Repositories [EB/OL]. [2015-06-19]. <http://www.opendoar.org/find.php>.
- [7] 宛福成.开放获取运动、政策与服务综述[J].情报科学,2006(11):1746-1751.
- [8] 付晚花,肖冬梅.国际开放获取政策及其研究进展综述[J].图书馆杂志,2010(3):23-27.
- [9] 李国俊,邱小花,季淑娟,等.国外高校硕博学位论文强制性开放获取研究[J].大学图书馆学报,2015(1):85-89.
- [10] 倪代川,任树怀.信息共享空间环境下开放获取与知识产权探析[J].情报理论与实践,2009(7):61-64.
- [11] 李秋实,宋仁君,张雅男,等.国际开放获取学位论文资源的整合利用研究[J].大学图书馆学报,2015(5):44-49.
- [12] Suber P. Open Access in 2008 [J/OL]. The Journal of Electronic Publishing, 2009, 12(1) [2013-03-25]. <http://quod.lib.umich.edu/cgi/text/idx/jjep/3336451.0012.104/-open-access-in-2>.
- [13] 李武,梁小建,杨琳.近五年来开放获取运动的国际进展分析[J].科技与出版,2013(8):71-75.
- [14] 刘丹,李欣,于亚秀,等.面向社会的学位论文信息开放获取与共享平台建设[J].图书馆,2014(1):129-133.

作者简介

张雅男,女,1988年生,硕士,天津社会科学院图书馆馆员,研究方向:知识服务,E-mail:2858871627@qq.com。

Research on Open Access of International Dissertation Based on Internet Survey

ZHANG YaNan

(Tianjin Academy of Social Sciences Library, Tianjin 300191, China)

Abstract: The Open Access of thesis and dissertation is an important aspect of Open Access research. Various kinds of dissertation OA programs have been developed in many countries. Based on internet survey, this article makes an empirical study between many countries' dissertation OA institutional repository, such as America, Britain, Australia and Hong Kong as well as Taiwan in China. After summarizing some characteristics of international dissertation Open Access Repository, suggestions of how to develop Chinese dissertation Open Access programs have been given.

Keywords: Thesis and Dissertation; Open Access; Internet Survey

(收稿日期:2016-02-19)