

NSTL开放获取资源组织揭示方式研究^{*}

任慧玲¹, 葛红梅¹, 刘蕾¹, 赵艳²

(1. 中国医学科学院医学信息研究所, 北京 100005; 2. 中国科学院文献情报中心, 北京 100190)

摘要: 开放资源已成为图书馆资源建设的重要组成部分, 其分散性、动态性和多样性为组织揭示带来了挑战。本文分析开放资源组织揭示的需求特点、借鉴国内外主要科技信息出版商、开放仓储和开放信息资源集成商对开放资源组织揭示的经验, 结合NSTL已有的资源和信息服务基础, 从增加资源内部关联、明确权益信息、实现开放资源与订购资源整合揭示三个方面提出NSTL开放获取资源揭示方案, 为下一步整合利用开放资源提供参考。

关键词: 开放获取; NSTL; 组织揭示

中图分类号: G250

DOI: 10.3772/j.issn.1673-2286.2015.08.005

开放资源已逐渐逼近主流学术资源, 正在成为图书馆资源建设不可或缺的组成部分^[1]。与传统资源相比, 开放获取资源在出版形式、开放政策方面具有多样性、动态性的特点, 其物理分散性进一步增加了图书馆对其进行组织揭示和再利用的难度。如: 开放协议不同导致的揭示层次不一, 开放政策不稳定, 开放资源的元数据标准各异, 数据接口不同, 统一采集揭示难度较大等。因此, 通过分析不同类型及来源开放资源的特征、元数据结构等特点, 设计面向用户信息发现的开放资源组织揭示方式, 已成为开放资源建设研究的重要内容。

文、科技报告等传统资源类型外, 还包括开放课程以及开放数据等。不同类型开放资源的描述元数据各不相同, 如“学位”元数据是“学位论文”类型资源的特征属性, 对应的属性值是“硕士”、“博士”等。开放资源除了类型多样以外, 还具有“视频”、“音频”、“程序”等多种数据格式以及“网络”、“硬盘”等多种载体形态, 并且一个开放资源数据对象中往往是多种数据格式和载体形态的组合, 如一篇学位论文中既有文本, 又包括图像和音视频, 并且可能还有网络和硬盘等多种载体形态。

1 开放资源信息组织揭示特点分析

传统的资源组织主要包含了各种分类表或主题词表, 以及相应的标引规则系统等。开放资源的组织在传统资源组织揭示的基础上, 还要考虑其多样性、权益复杂性等特点, 进行多层级的知识关联揭示。

1.1 开放资源类型的多样性

开放资源与传统印本文献、订购网络文献相比, 文献类型更加丰富。除了期刊、图书、会议论文、学位论文、科技报告等传统资源类型外, 还包括开放课程以及

1.2 开放资源组织的关联性

随着数字出版、结构化出版、语义出版技术的普及, 开放资源的信息组织越来越表现出细粒度、关联化的特点。开放资源的数字出版对信息的组织揭示更多的从“一本书”、“一种期刊”深入到图书章节、论文片段、一个图表、引文等更加细粒度的“知识单元”, 不同对象之间的包含、继承、衍生等关系也更加复杂。此外, 由于开放资源的动态特性, 订购资源可以转换为开放资源, 开放资源也可以转换为订购资源, 开放资源本身也有“关停并转”, 因此, 不同开放资源类型之间、开

^{*} 本研究得到国家科技图书文献中心(NSTL)资助项目“NSTL开放资源描述元数据规范”资助。

放资源与订购资源之间的关联揭示更为复杂。

1.3 开放资源信息的可获取性和使用权益的复杂性

可获取性指可获取的信息内容、获取方式、获取时差等。开放资源可获取的信息内容包括元数据或不同版本的全文数据; 获取方式包括人工下载或者提供标准API下载接口, 也可分为阅读、下载、分享等; 获取时差即是指即时开放获取、延时开放获取等获取时间限制。除CC协议外, 不同出版社对数据下载、使用的约定各不相同, 这导致开放资源的权益问题更加复杂。“权益信息”是开放资源组织揭示中必须考虑的问题^[2]。与传统的订购资源不同, 开放获取资源的“可获取信息内容”、“获取方式”、“获取时差”、“获取网址”、“权益信息”等属性是开放资源的专有特征, 对于用户获取和使用开放信息资源十分重要, 也是开放资源信息揭示的对象。

2 重要学术资源平台中开放资源的组织揭示

开放资源主要的信息来源既包括开放获取平台, 如DOAJ、SciELO、BioMed Central、PMC等, 也包括

商业性数据库平台, 如Science Direct, SpringerLink和SpringerOpen等^[3]。其中, DOAJ、PMC和Science Direct平台在信息组织与资源揭示方面比较具有代表性, 分别代表了开放获取期刊集成网站、开放获取资源数据库、综合性商业数据库平台三种类型。通过对三个平台中开放资源的组织揭示分析, 可以基本了解当前开放资源平台的信息组织特征及规律。三个平台中资源组织的各个指标对比分析结果见表1。

通过上述对比可知: Science Direct平台整合了开放资源和订购资源, 并拥有期刊的所有元数据, 能够实现统一检索和发现, 在数据关联性方面也做的较好。此外, Science Direct用不同图标标识出了不同等级开放度及使用协议的文献, 有利于用户的发现和利用。PMC通过将开放资源与PubMed数据相关联, 并且通过用户定制的筛选功能来选择不同资源。由于其文摘数据与PubMed/Medline共用, 展现了强大的主题检索、智能检索功能。DOAJ虽然收录了多学科、多语种、经同行评议的学术OA期刊及其元数据信息, 但由于期刊和论文的数据均由OA出版社提供, 文摘、全文元数据完整性无法保障, 只有60%的期刊能够提供论文检索。

3 NSTL开放资源的组织揭示框架

2011年起, 国家科技图书文献中心(NSTL)就启动

表1 三个重要学术资源平台中的开放资源信息组织揭示对比

资源名称	Science Direct	PMC	DOAJ
组织方式	开放资源和订购资源统一在Science Direct平台一站式检索功能中	按完全参与期刊、NIH资助论文发表期刊、选择性存缴期刊分类	期刊按所属学科、二级或三级主题分类; 提供刊名字顺索引
元数据标准	Journal Article Tag Suite	Dublin Core、NLM Journal Archiving and Interchange XML、Journal Article Tag Suite	Dublin Core
全文格式	HTML/PDF	PDF/HTML/ePub/PubReader/XML	HTML/PDF
数据获取	OAI-PMH	OAI-PMH/API/全文FTP	OAI-PMH/未直接提供篇级全文下载, 只提供期刊链接
使用许可协议	CC协议/Elsevier协议	CC协议/其他类似允许再利用的协议	CC协议/Attribution-Share Alike License
检索功能	期刊浏览/期刊检索/文章检索/学科导航/新刊导航/首字母导航	期刊浏览/快速检索/期刊检索/文章检索	期刊浏览/期刊检索/文章检索/学科导航/新刊导航/首字母导航
个性化	社交媒体分享	参考文献与PubMed数据库系统内链接/收藏/订阅	社交媒体分享

了开放资源的采集、组织揭示与服务工作,但对开放资源揭示的层次仅限于资源名称导航,不能提供检索浏览、分面聚类以及个性化的基于社交网络的信息服务,并且尚未与NSTL的订购资源进行整合,进而为用户提供一站式的资源发现服务。

随着开放资源以及发现服务的蓬勃发展,为了向用户提供便捷易用的开放资源和订购资源集成发现服务,亟需结合NSTL本身的资源和服务特点,设计并提出NSTL新的开放资源组织揭示框架,如图1所示。

3.1 增加资源关联,实现多粒度组织揭示

根据开放资源数字化的特点,制定统一的元数据

标准,以期刊论文、图书章节、会议论文、科技报告、课件内容作为开放资源信息揭示粒度,即以“单篇”为单位,描述信息的外部特征、内容特征、开放信息特征、关联信息属性、管理信息属性等,对于开放课件、开放会议、开放学位论文等资源,依据资源本身特点,增加相应的描述和管理元数据,比如“课程使用对象”、“论文导师”、“会议举办者”等信息。提供更多浏览检索、分面聚类、外部链接、知识发现功能,同时根据“读图时代”用户需求特点,可视化展示信息分布和检索结果,提高系统的易用性。

通过细化信息揭示粒度,进行跨类型资源关联。比如“一篇期刊论文”与“一个图书章节”可以因为“主题相关”或“作者相关”建立关联,“会议论文”与“学术

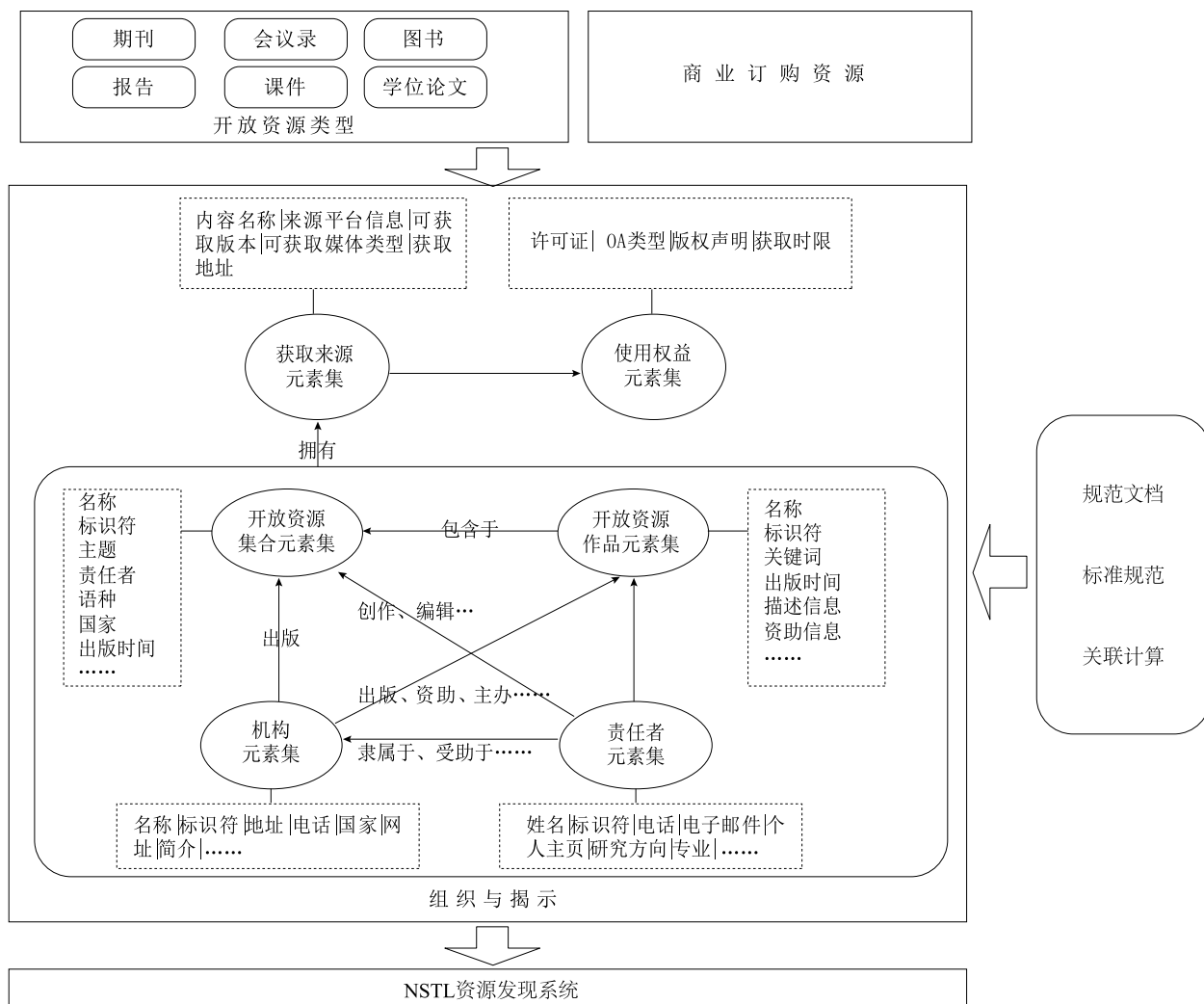


图1 NSTL开放资源组织揭示框架

来源: NSTL开放资源描述元数据标准规范课题组.NSTL开放资源描述元数据规范(征求意见稿).2015年4月

机构”可以通过“会议主办机构”建立关联。NSTL开放资源描述元数据标准规范课题组通过制定文献品种的唯一标识符以及开放资源“历史信息”和“唯一标识信息”的元数据描述,实现相同资源的唯一定位以及文献出版过程中“母体信息”变更之间的关联,见图1。

3.2 重视开放资源权益信息,明确全文使用权限

首先,在元数据框架中即设立“使用权益”、“链接地址”信息等元数据内容;其次,在数据采集过程中,采集开放资源的使用权益和链接地址,作为进一步传播利用这些开放资源的遵循依据;最后,在结果展示和客户服务界面,展示不同开放资源的使用权益信息,指导用户“署名”使用、“非商业性”使用、“禁止演绎”使用和“相同方式共享”使用,既保护创作者和出版商的权益,又促进开放资源的广泛传播和可持续利用^[4]。

3.3 实现开放资源与订购资源的整合揭示

兼容NSTL原有资源类型和元数据标准,建立相关资源之间的关联。NSTL已有的订购资源包括期刊、会议文集、科技报告等多种类型的文摘、引文以及通过原文传递服务获取全文的印本馆藏。现在即将遴选加入的大量多种类型的开放获取资源,将整合在统一平台

上。借鉴美国国立医学图书馆开放资源仓储PMC的揭示方法,将订购资源和开放资源融为一体,提供统一的检索入口。用户在一个检索平台上即可一站式检索“开放资源”和“订购资源”,统一获得摘要和题录,根据订购资源、免费资源、开放资源等文献来源的不同,分别通过“原文传递”、“直接全文链接”、“直接开放资源链接”三种方式获得全文。

同时针对开放获取资源,提供明确的信息标识和方便的全文链接。在数据结构设计上,既考虑NSTL原有的元数据标准,又根据开放资源的特点,扩展相应的“资源类型”、“使用权益”、“地址链接”、“数据管理”、“格式和载体”等元素,适应开放资源揭示的特点和需求。

参考文献

- [1] 张晓林,李麟,刘细文. 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] 许可协议说明[EB/OL][2015-08-04]. http://creativecommons.net.cn/licenses/licenses_exp/2014-09-15.
- [3] 开放资源建设操作指南[EB/OL][2015-08-04]. <http://open-resources.las.ac.cn/drupal/?q=node/3025>.
- [4] 陈怡帆. 图书馆在开放获取期刊再利用中的权益保障研究[D].北京: 协和医科大学,2015.

作者简介

任慧玲, 女, 1971年生, 中国医学科学院医学信息研究所研究馆员, 研究方向: 信息资源建设, E-mail: ren.huiling@imicams.ac.cn。
葛红梅, 女, 1979年生, 研究方向: 图书馆信息组织、图书馆元数据建设。

Research on the Ways of Organization and Revelation about NSTL Open Access Resource

REN HuiLing¹, GE HongMei¹, LIU Lei¹, ZHAO Yan²

(1. Institute of Medical Information, Chinese Academy of Medical Sciences, Beijing 100005, China;

2. National Science Library of Chinese Academy of Sciences, Beijing 100190, China)

Abstract: Open resource has become an important part of the construction of library resources, its dispersion, dynamic and diversity brings challenges for organization and revelation. This article analyzes the characteristics of open resource organization and revelation needs, draws lessons from domestic and international experience of the major scientific and technological information publishers, open storage and open information resources integrators on organizing and revealing open resource, combined with NSTL existing resources and information services infrastructure, from the three aspects of the increase in internal resources association, the clear of interest information, the achievement of the integration and revelation of open resource and subscription resources to propose reveal scheme of NSTL open access resources, in order to provide a reference for further integrate and use of open resource.

Keywords: Open Access; NSTL; Organization and Revelation

(收稿日期: 2015-08-04; 编辑: 雷雪)