

期刊规范文档建设方案与应用场景研究*

王军辉, 方安, 任慧玲, 葛红梅, 吴思竹, 胡佳慧, 梁芳, 胡铁军
(中国医学科学院医学信息研究所, 北京 100020)

摘要: 期刊规范文档是由期刊名称规范描述记录组成的文档, 可以实现对期刊信息的规范控制和统一管理。以NSTL为例, 从两个方面概括NSTL数字业务工作中各系统对期刊规范文档的应用需求, 提出以NSTL联合目录系统中的期刊目录数据为基础, 以外部数据资源为补充的NSTL期刊规范文档建设方案, 描述期刊规范文档在NSTL文献综合管理系统、联合目录系统、联合数据加工系统及前台用户检索与服务系统中的应用场景。

关键词: NSTL; 规范文档; 联合目录系统

中图分类号: G203

DOI: 10.3772/j.issn.1673-2286.2015.07.002

1 引言

期刊资源是各类文献信息服务机构收藏的重要对象, 然而各种原因所导致的期刊名称及相关信息的变化, 不仅给期刊管理工作带来不便, 同时也影响文献服务的效果。规范文档(Authority File)是指由规范记录组成的计算机文档。规范记录包含有各个检索点的规范标目、参照关系以及参照根查等信息^[1]。规范文档作为一种工具, 通过实行规范控制来保证文档标目的准确性和一致性, 可以引导用户根据多种标目形式找到相应的规范标目, 进而完成针对相应目录的操作和管理。期刊规范文档则是由期刊名称规范描述记录组成的文档, 可以实现对期刊信息的规范控制和统一管理。

在期刊规范文档建设方面, 国内相关信息机构针对特定应用目的建立了内部使用的期刊规范库, 如中国科学技术信息研究所面向知识评价建立的期刊规范名称数据库^[1], 而国内规模较大的规范数据库, 如中文名称规范联合数据库^[2]和国家图书馆规范名称数据库^[3]

则主要针对个人名称、团体名称及会议名称等进行规范, 基本不涉及期刊名称规范。国外著名的美国国会图书馆规范数据库^[4]和VIAF数据库^[5]虽然涉及期刊名称的规范, 但数据量很少; 收录期刊数量庞大的乌利希期刊指南^[6]则是以“旧见新”的参照方式提供最新的期刊书目信息, 由于不涉及规范标目的设置, 因此并不算是严格意义上的期刊规范文档; 美国国立医学图书馆的locatorPlus系统^[7]专门针对期刊信息建立规范文档并和书目系统进行关联, 但限于学科领域, 仅包含生物和医学领域的期刊规范信息。

国家科技图书文献中心(National Science and Technology Library, NSTL)是基于网络环境的科技信息机构^[8]。截至2014年底, NSTL联合目录系统中收录的24,000多种期刊仅刊名有变化的就达6,500多种, 所占比例近30%。除了期刊更名问题, 还存在不同数据来源中的期刊别名和刊名缩写不一致问题。因此, 建设NSTL期刊规范文档对于加强NSTL数字业务工作中期刊信息的集中管理, 提高资源利用的效益和文献

* 本研究得到国家科技图书文献中心委托课题“NSTL期刊规范文档建设方案研究”(编号: 2014XM040)资助。

信息服务的质量有重要意义。

传统意义上的规范文档一般用于编目和检索两个方面^[1]，而面向NSTL全数字业务工作平台的期刊规范文档的利用则要贯穿于期刊遴选采购、期刊信息编目、期刊数据加工、信息检索服务的整个业务流程中，这就对规范文档建设提出更高要求。本文从满足NSTL各业务系统对期刊规范文档需求的角度，阐述NSTL期刊规范文档的建设方案及后续可能的应用场景。

2 期刊规范文档建设的需求分析

目前NSTL已经形成了涵盖文献资源采集、数据加工、文献信息服务等各项工作流程的数字化业务管理与服务平台，主要包括文献综合管理系统、联合目录系统、联合数据加工系统及前台用户检索服务系统^[9]。结合各业务应用系统的实际运行情况，以及在整个NSTL范围内涉及到的期刊资源状况，NSTL数字业务工作平台对期刊规范文档的需求可以概括为以下两个方面。

2.1 资源描述与检索

期刊是一种连续出版物，随着时间的推移，相关期刊之间会发生吸收/被吸收、部分吸收/被部分吸收、继承/被继承等多种复杂的沿革关系。最新的编目标准《资源描述与检索》(RDA)^[10]界定的连续作品之间的关系达10种之多。与传统图书馆一样，NSTL需要建立期刊规范文档来应对上述问题对期刊管理工作的挑战。基于时间轴对期刊沿革信息进行纵向规范控制，对内可以指导期刊采编工作，对外可以为用户提供更好的检索服务。

2.2 数据统计与分析

NSTL根据不同的业务需求，分别购买或采集了大量的外部数据资源，如汤森路透的论文数据和美国国立医学图书馆的PMC数据等。而在对这些外部数据资源进行本地化的过程中，会遇到同一种期刊在不同数据来源中有不同的刊名和刊名缩写的问题。另外，NSTL联合数据加工系统在对订购期刊的引文数据进行加工处理的过程中，由于各期刊对引文标注要求不尽相同，因此会产生引文数据和基础数据无法匹配的情况。上述问题均可以通过期刊规范文档的映射来解决。

对同一种期刊的不同刊名表达形式及缩写方式进行横向的规范控制，对内可以辅助数据加工工作，对外可以提供更完善的引证分析服务。

3 建设方案

NSTL期刊规范文档建设以NSTL联合目录系统中的期刊数据为基础，基于一定规则进行期刊规范文档的自动抽取；同时以外部数据源中的期刊信息为补充，来扩展丰富期刊规范库内容，最终通过人工编辑、审核后形成正式规范文档。

借鉴FRBR/FRAD模型理念^[11]，将NSTL联合目录系统中每条期刊书目对应的期刊均视为一个物理品种实体，利用期刊历史变革的描述信息，将彼此有沿革关系的单个期刊实体对应的物理品种书目记录汇聚成期刊逻辑品种，即一个作品(Work)。针对纵向和横向两个维度的规范控制需求，实际处理过程分为两个阶段。

第一阶段：将NSTL联合目录系统中关于期刊历史变革元素集中的先前款目对应的书目视为父节点，后继款目对应的书目视为子节点，通过对期刊沿革形成的树形结构进行遍历来确定所有相关书目记录，针对每条书目记录分别生成候选规范记录，同时记录下每个书目记录的父节点书目ID和子节点书目ID。所有相关书目记录对应的期刊集合视为一个逻辑品种，由此实现基于时间轴的期刊沿革信息的纵向汇聚。涉及到的NSTL联合目录系统中的期刊元数据项见表1。另外，针对NSTL联合目录系统中有关期刊沿革的信息可能存在部分缺失的问题，采用外部期刊资源库中的沿革信息作为有效补充。

第二阶段：针对第一阶段生成的每条初始规范记录，分别集成来自购买外部数据源中的期刊别名信息，实现对同一种期刊的不同刊名表达形式及缩写方式的横向汇聚。最后经编目人员人工编辑审核后，生成正式的NSTL期刊规范记录。

NSTL期刊规范文档的具体建设流程见图1，其中：

(1) 外部期刊资源库沿革信息的利用(虚线框)以乌利希期刊指南为例进行展示。

(2) 刊名规范化处理依据《NSTL联合目录系统建设项目基本元数据规范》(2013年版)和《国家科技图书文献中心连续出版物编目手册》(2013年修订版)，对期刊正题名和并列题名进行规范化处理。具体包括题名选取、针对题名首冠词、题名中的数字及特殊

符号等所进行的一系列处理过程。

(3) 外部数据源期刊别名API中的数据为NSTL数

表1 NSTL联合目录系统中期刊沿革相关元数据项

先前款目	
名称	标签
先前书目ID	record_id_preceding
先前题名	title_preceding
先前ISSN	issn_preceding
正编正刊书目ID	record_id_parent
正编正刊题名	title_parent
正编正刊ISSN	issn_parent
后继款目	
名称	标签
后续书目ID	record_id_succeeding
后续题名	title_succeeding
后续ISSN	issn_succeeding
补编增刊书目ID	record_id_suppl
补编增刊题名	title_suppl
补编增刊ISSN	issn_suppl

据研究管理中心基于购买的多种外部数据源, 通过特定的匹配算法抽取后获得^[12]。

4 应用场景和更新机制

期刊信息的变更会持续发生, 因此期刊规范文档的建设也必须动态进行。上节中的建设流程只是NSTL期刊规范文档建设的第一步, 还需要在期刊遴选、编目和引文加工过程中进行动态更新和维护。图2设想了未来NSTL期刊规范文档在整个NSTL数字业务工作平台中的应用场景及后续可能的动态更新机制。

4.1 文献综合管理系统

NSTL文献综合管理系统通过对多种订购模式对应的订购过程的控制, 来实现各类文献资源购买及相关活动的流程化管理。系统具体数据的更新与维护, 主要由NSTL各成员单位采集人员操作完成, 分散性维护数据的运作机制容易导致数据质量良莠不齐。NSTL期刊规范文档建成后, 中心层面在审核确定相关期刊信息或新增资源品种时, 可通过期刊规范文档进行查重确认, 以规范订单数据。

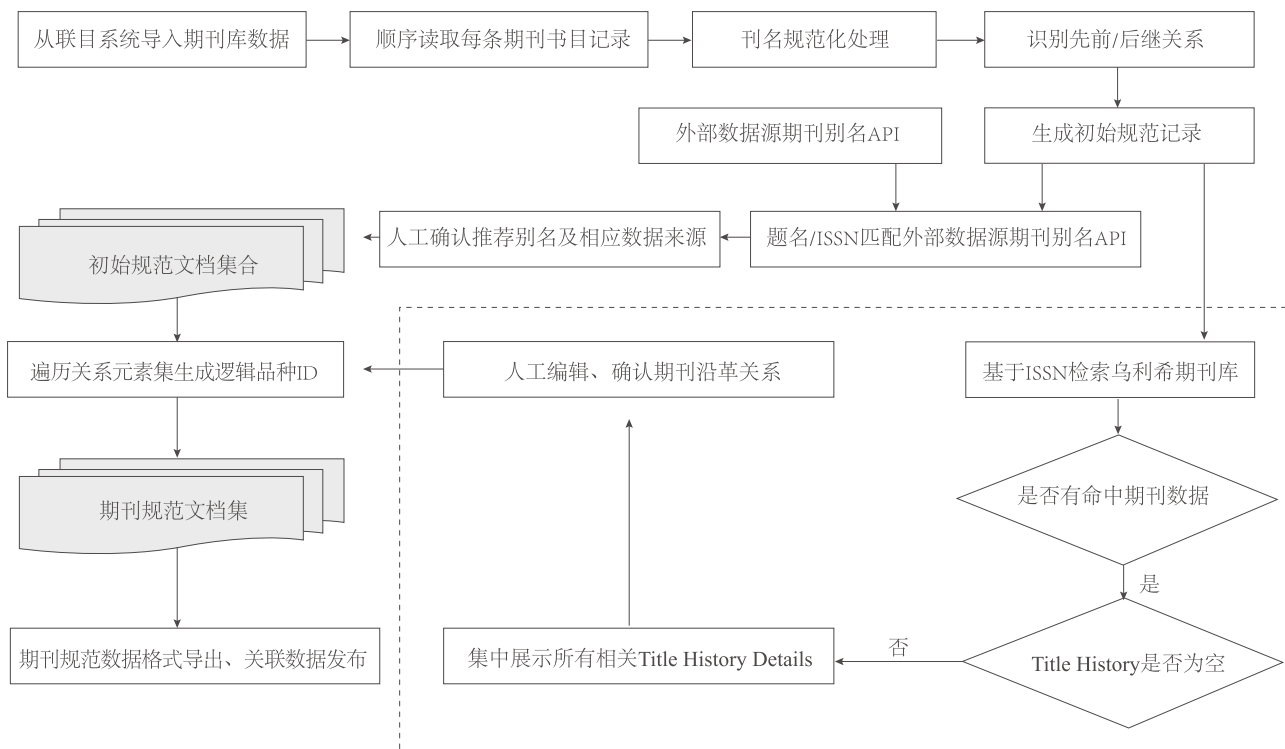


图1 NSTL期刊规范文档建设流程图

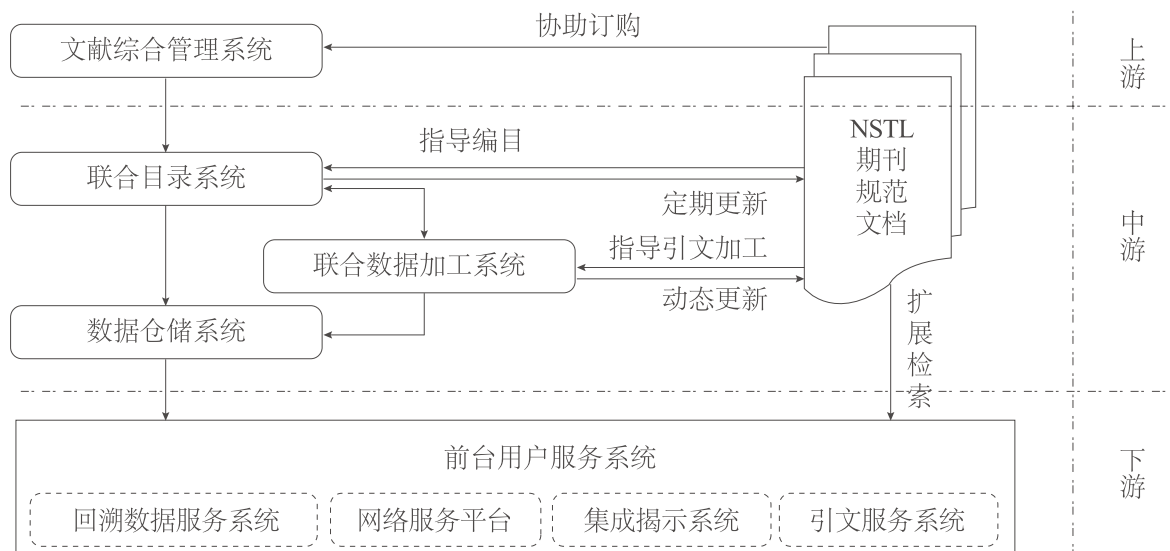


图2 期刊规范文档在NSTL数字业务工作平台中的应用

4.2 联合目录系统

NSTL各成员单位分别有独立的本地编目系统，NSTL联合目录系统定期收割各成员单位的书目数据信息。9家成员单位的编目人员在编目过程中，可通过NSTL期刊规范文档的参照信息来判定某一种期刊是否为新刊，或是联合目录系统中已有期刊的变更期刊，进而采取不同的编目处理方式。

反之，当联合目录系统收割到成员单位的新增期刊书目时，编目人员可在系统辅助基础上判断期刊书目的类型，并进行相应的期刊规范文档更新操作。如果新增期刊书目和联合目录系统中的已有期刊没有任何沿革关系，则生成一条新的规范记录，赋予全新“逻辑品种ID”；如果是由联合目录系统中已有期刊变更而来，则在生成一条新的规范记录的同时，记录其与父节点“书目ID”的关系，并沿用已有“逻辑品种ID”。

4.3 联合数据加工系统

NSTL联合数据加工系统包括中心管理系统和各分中心管理子系统，加工对象包括期刊论文、会议录、文集汇报、学位论文、科技报告、专利和标准等文献类型^[12]。NSTL期刊规范文档一方面可以应用于NSTL联合数据加工系统分中心的引文数据加工工作，辅助引文刊名的识别和拆分；另一方面，可以应用于NSTL联合数据加工系统中心管理系统的引文数据规范工作，支持

引文归一处理。同时，在引文数据加工过程中，基于特定算法识别出的期刊新的别名及缩写形式，在经过人工审核后亦可以补充进规范文档，对规范文档进行动态更新。

4.4 前台用户服务系统

面向前台用户提供检索服务的回溯数据服务系统、网络服务平台、集成揭示系统和引文服务系统均可以基于NSTL期刊规范文档，集中展示同一种期刊不同历史阶段的所有沿革信息，以方便用户选择；或者通过自动汇聚输出该品种下所有沿革期刊相关的命中文献，提高用户检索的查全率和查准率。

5 结语

为满足NSTL数字业务工作平台中各系统对期刊规范文档的应用需求，本文初步提出以NSTL联合目录系统中的期刊目录数据为基础，以外部数据资源为补充的NSTL期刊规范文档建设方案。下一步需要强化相关研究，具体包括：对多种外部期刊资源库进行深入分析，并在数据测试的基础上确定一种或多种期刊资源库，作为对NSTL联合目录系统中已有期刊沿革信息的必要补充；在《NSTL元数据标准规范指南》指导下，遵从NSTL元数据框架进行NSTL期刊规范文档元数据方案的设计工作；进行样例数据测试，根据测试过程中发现

的问题对期刊规范文档建设流程中的部分环节进行适当调整, 在此基础上撰写系统需求说明书, 开发NSTL期刊规范文档建设与管理系统。

参考文献

- [1] 曾建勋, 王立学. 面向知识评价的规范文档建设方法[J]. 图书情报工作, 2012(10):101-106.
- [2] 中文名称规范联合数据库检索系统[DB/OL]. [2015-06-19]. <http://cnass.cccna.org/jsp/index.jsp>.
- [3] 卜书庆, 郝嘉树. 国家图书馆中文书目规范控制现状及研究[J]. 图书馆论坛, 2010(6):209-213.
- [4] Library of Congress Authorities[DB/OL]. [2015-06-19]. <http://authorities.loc.gov/>.
- [5] VIAF: The Virtual International Authority File[DB/OL]. [2015-06-09]. <http://www.viaf.org/>.
- [6] Ulrichsweb: The Global Source for Periodicals[DB/OL]. [2015-06-09]. <http://ulrichsweb.serialssolutions.com/>.
- [7] NLM LocatorPlus[DB/OL]. [2015-06-09]. <http://www.locatorplus.gov/>.
- [8] 国家科技图书文献中心[DB/OL]. [2015-06-09]. <http://www.nstl.gov.cn>.
- [9] 曾建勋, 邓胜利. 国家科技图书文献中心资源建设与服务发展分析[J]. 中国图书馆学报, 2011(2):30-35.
- [10] RDA Toolkit[EB/OL]. [2015-06-09]. <http://access.rdatoolkit.org/>.
- [11] 规范数据的功能需求[EB/OL]. [2015-06-09]. http://www.ifla.org/files/assets/cataloguing/frad/frad_2009-zh.pdf.
- [12] 张建勇, 孟连生, 刘筱敏, 等. NSTL联合数据加工系统的功能框架设计[J]. 图书情报工作, 2011(3):64-69.

作者简介

王军辉, 男, 1982年生, 中国医学科学院医学信息研究所助理研究员, 研究方向: 医学信息资源管理与利用, E-mail: informatics@163.com。

Construction Scheme and Application Scenes of Journal Authority File in NSTL

WANG JunHui, FANG An, REN HuiLing, GE HongMei, WU SiZhu, HU JiaHui, LIANG Fang, HU TieJun
(Institute of Medical Information, Chinese Academy of Medical Sciences, Beijing 100020, China)

Abstract: Journal authority file, which consists of authority records of journal title, could control and manage journal related information. The paper summaries its application demands in NSTL digital business platform, proposes its construction scheme based on the records in NSTL union catalog system and supplemented by external journal resources, describes its application scenes in literature management system, union catalog system, joint data processing system and user retrieval systems in NSTL.

Keywords: NSTL; Authority file; Union catalog system

(收稿日期: 2015-06-20; 编辑: 雷雪)