

国内外语义出版实践研究^{*}

李娇, 寇远涛, 黄永文, 薛欢欢, 鲜国建
(中国农业科学院农业信息研究所, 北京 100081)

摘要: 本文介绍多种国内外重要语义出版机构/平台的实践进展, 并从数字资源描述模型、语义知识增强形态、语义出版发布形式三方面对其进行比较分析, 对比国内外语义出版的发展现状, 总结现有实践研究的共性特征及其存在问题, 并探讨我国科技期刊语义出版模式的发展对策, 以期为我国语义出版的研究和实践提供参考。

关键词: 语义出版; 描述模型; 语义知识增强; 发布形式

中图分类号: G274

DOI: 10.3772/j.issn.1673-2286.2017.12.004

1 引言

数字网络环境下, 作为数字出版、复合出版的高级形态, 尤其在科学、技术和医学等领域进行了一系列试验并逐步推出正式服务后, 语义出版成为学术期刊的主流出版形式。语义出版最早由Shotton于2009年提出^[1], 他将语义出版定义为一种语义增强的期刊出版形式, 指出语义出版是一种增强已出版期刊文献内涵意义的过程, 通过语义标记来丰富出版物的知识内容和表现形式, 提高出版物信息的可操作性、交互性和关联性, 最终实现智能化出版; 他还系统化地提出科技期刊论文语义出版的表现形式, 分析了DOI、超链接、支持排序的参考文献, 以及可语义标注的术语、基于上下文引用关系等语义出版形式的可行性^[2]。

语义出版实践者通常借助语义技术、网络服务协议(如语义出版及引用本体^[3]), 从出版平台、出版物、阅读终端三个层面增强语义^[4]。语义出版对文献知识内容的结构化发布和呈现, 为科研用户提供更加精确、高效的阅读体验, 同时也给信息服务商、图书馆等机构的传统服务流程带来冲击, 因此, 近年来语义出版已成为国际学术界在学术出版与信息领域管理领域的研究热点。

目前, 出版机构、信息服务商、图书馆等建设主体正积极开展语义出版实践, 虽然成果较丰富, 但关于国

内外语义出版实践对比分析的研究很少。本文选取目前国内外语义出版实践成果相对成熟的出版机构/平台Nature、PMC、RSC、Elsevier和中华医学会杂志社等, 结合学科领域发展状态从数字资源描述模型、语义知识增强形态、语义出版发布形式等方面分别进行比较分析, 总结国内外语义出版发展现状、现有实践研究的共性特征及其存在的问题, 以期为我国语义出版的研究和实践提供参考。

2 国内外代表性语义出版实践

2.1 Nature

Nature在语义出版方面的尝试始于关联数据, 2013年, Nature发布一款科学学术术语动态页面产品^[5], 使其作为语义出版架构的简单概述, 并构建了依据出版核心工作流的关联数据架构, 形成用owl本体定义的通用元数据模型。此后, Nature开启了基于本体实践语义出版的新篇章。2015年, Nature发布自然本体门户“nature.com ontologies portal”, 提供支撑Nature语义出版的核心模型、领域模型等主要语义模型, 以及文章和贡献者等实例数据集^[6]。

Nature语义出版模型架构如图1所示, 基础层由

^{*} 本研究得到公益性科研院所基本科研业务费项目(编号: JBYW-AII-2016-17)和中国农业科学院科技创新工程项目(编号: CAAS-ASTIP-2017-AII)资助。

RDF语言族提供,用于对核心本体进行编码。该语义本体模型通过继承SKOS模型的标准化语义,定义多个领域层次类别的基础概念,由于它们是对企业内部特定应用或领域知识的编码,故称为领域本体。

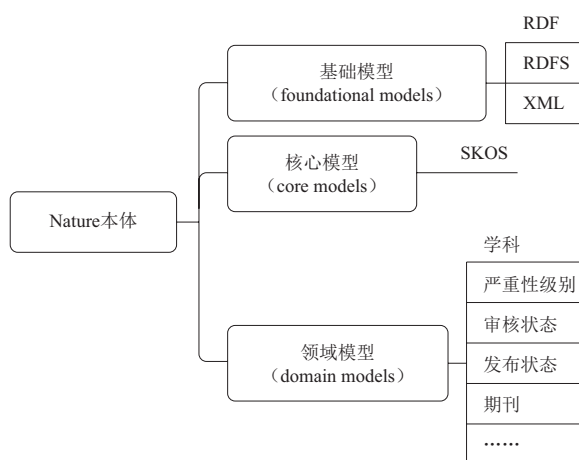


图1 Nature语义出版模型架构

2015年, Nature出版集团与Springer合并后在语义出版方面的探索进一步深入, 2017年Springer·Nature推出SciGraph (科研图谱) 服务, 整合科研界的各种信息^[7], SciGraph的数据集由Springer·Nature和Digital Science共同合作完成, 包含1.5亿—2亿条学术界关注对象的信息 (三元组)。Nature将传统集中式、以文献为中心的、XML为基础的企业出版平台转换为分布式、以事件为中心和RDF为基础的复合语义架构, 保证语义出版产品的质量和稳健性。

2.2 PubMed Central

PubMed Central是由NIH/NLM创建并维护的生物医学和生命科学文献资源整合中心, 隶属于美国国家生物技术中心, 其在OA期刊的数字化出版技术远超过其他平台^[8]。PMC在语义出版过程中的主要做法是实现文献全文XML格式转换及动态呈现^[9]。

2012年, 美国国家生物技术中心基于PMC和各出版商的实践开发期刊文档标签套件 (the Journal Article Tag Suite, JATS), 一组定义XML元素、期刊文献标记属性的NLM DTD (National Library of Medicine Document Typing Definitions) 标准^[10], 用于对文献格式进行统一描述。JATS规定期刊文章分为四部分: 前置部分 (front), 存储期刊论文的元数据内容; 主体部分 (body), 存储文章正文部分的内容信息, 细分为内容、

章节及签名栏三部分, 由与数学公式、表格、图片、文本内容、其他辅助信息相关的26种元素组成; 后置部分 (back), 存储术语表、参考文献或附录等辅助信息; 浮动部分 (floats-group), 包括文章主体和后置部分引用的图和表、对文章的评论、编辑的总结、读者反馈、作者对同行评议内容的反馈、文章的次级论文等^[11-12]。同时, JATS标准兼容多方使用需求, 根据不同的应用场景对标签集进行分类, 包括存档和交换标签集^[13]、期刊出版标签集^[14]和文章作者文档标签集^[15]。

另外, JATS提供了一种通用的期刊数据交换文档格式, 是推动PMC语义出版发展的核心技术, 由于其良好的适用性与可操作性, 目前广泛应用于出版商、科技期刊、知识库和图书馆等机构, 并且已成为美国的国家标准和科技期刊界的行业标准^[11]。

2.3 RSC

RSC是全球知名的化学科学出版社之一, 目前已出版44个同行评议期刊, 提供一系列在线数据库和文献更新服务, 是语义出版领域的重要范例^[16]。2007年, RSC提出“Science Come Alive”并进行积极探索^[17], 设立的RSC Prospect项目旨在通过语义增强实现期刊文章的机器可读, 丰富RSC期刊在线出版的功能。RSC语义出版模式的关键是利用本体和唯一的化合物标识符, 使文章可被计算机识别。RSC的技术编辑在文章中标识出化合物、概念和数据, 将其链接至相关的学术环境, 极大增强了RSC文章 (HTML格式) 对学科知识的揭示和关联能力。

由图2可见, RSC语义出版的核心要素包括对化学学科知识环境的关联、科技期刊论文的结构化描述与标记以及全文嵌入式HTML的呈现方式^[18]。

(1) 关联化学学术环境。为更好地实现化学学科相关内容和数据的读取、对比, RSC将本体与自建的基于Web的化学结构数据库 (ChemSpider) 关联, 形成RSC语义出版自有的基础信息系统。

(2) 结构化描述与标记。基本流程包括XML预处理、XML编辑与修正、结合语境和主题领域对应的XML标注。

(3) 嵌入式HTML的呈现。RSC语义出版的学术期刊以Rich HTML形式发布, 且标记为Rich HTML的文章可实现全文嵌入式HTML标记, 并以不同颜色突出显示定义的词汇, 实现语义分类。

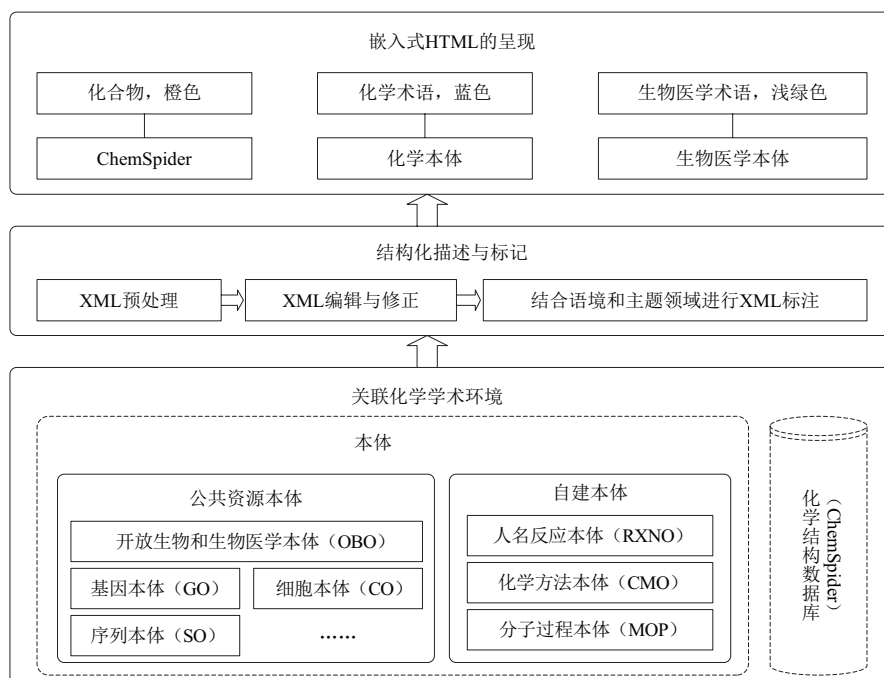


图 2 RSC语义出版核心要素

2.4 Elsevier

Elsevier作为全球领先的多媒体出版集团和科学、技术、医学信息产品和服务提供商^[19], 提出语义出版的发展路线图^[20], 阐释语义出版的内涵和实现路径, 通过一系列技术对期刊论文进行加工, 使得原始文本附加值得以提升, 最终变成富含语义知识的智能内容。

Elsevier在语义出版方面进行积极的尝试与开发, 2009年首次在Cell出版社实施“Article of the Future项目”^[21-22], 从呈现形式、内容和上下文三个方面创新传统的学术文献服务。随后, Elsevier旗下FEBS Letters与MINT合作, 对发表在FEBS Letters上的论文提供结构化摘要^[23]。2012年, Elsevier发布医学信息平台ClinicalKey^[24], 将“智能内容”引入临床领域, 实现检索人性化、可定制的内容服务。ClinicalKey拥有丰富的学科信息资源, 包括国际性综合生物医学信息书目数据库、期刊、图书、图片、医疗操作等13种门类^[23], 其与Healthline合作开发爱思唯尔合并医学分类法(Elsevier's Merged Medical Taxonomy, EMMeT), 并根据EMMeT进行深度标引, 将信息资源转换成核心医学概念、同义词、等级关系、本体关系的分类系统体系和本体库, 支持语义检索, 为检索请求提供具有针对性且具体的答案。Elsevier作为知名出版机构, 其在语义出版方面的实践

真正意义上呈现了国际语义出版发展全历程, 尤其是ClinicalKey作为语义出版在医学领域的经典应用实例, 为数据库语义化发展及学科语义出版平台构建提供了宝贵经验。目前, Elsevier旗下的ScienceDirect平台已经在7个学科领域建立语义出版模型^[25]。

2.5 国内语义出版发展现状

随着语义出版在知识点聚类上的优势逐渐显现, 我国出版领域也进行积极的尝试与开发。2009年彭希珺等率先提出Journal 3.0模型, 从丰裕化结构化语义化内容(Enriched Content)、关联与融汇发现技术(Linked & Mashed-up Discovery)、开放与交互的传播利用方法(Open & Collaborative Communication)三个维度揭示期刊数字化发展模型^[20]。

随后, 国内大的数据库商(如万方数据库)、期刊采编系统提供商(如玛格泰克)和部分期刊也开始逐步尝试语义出版模式。2014年, 中华医学会杂志社在研究JATS的基础上推出中华医学会期刊文档交换和存储标准CMA JATS, 依此指导期刊数据的生产, 迈出资源整合关键的一步, 推动全文数字出版^[26]。2016年, 乐小虬等开发了一种面向语义出版的机构化论文写作工具DPaper, 实现论文在写作阶段的结构化、对象化^[27]。

2017年,国内首个学术期刊动态语义出版与知识服务重点实验室成立,定位以应用为导向促进学术期刊动态语义出版和知识服务技术及产业发展。

目前国内语义出版仍处于探索阶段,由于语义技术与检索技术在具体领域实践方面的欠缺,及出版体制、资源权属等问题,与国外还存在一定差距,语义分类与检索系统还不够成熟^[28]。目前国内语义出版主要形式是排版软件开发商与各期刊采编系统提供商合作,实现编辑平台到排版软件的直接对接,部分期刊/机构与其达成三方合作,最终以中文全文HTML或Rich HTML的形式发布。其中,依托于中国作物学会和中国农业科学院作物科学研究所的期刊《作物学报》是由Rich

HTML发布的典型案例,其技术支持归属于玛格泰克。

3 多角度对比分析

3.1 数字资源描述模型

李楠等在大量的语义出版研究文献基础上归纳出语义出版的两条基本技术路线:(1)借助本体技术实现文献对象及其知识内容的语义描述;(2)采用关联数据为出版物连接更多外部开放的数据资源提供技术框架^[25]。由于学科领域发展和技术路线的差异性,上述各机构/平台对数字资源内容描述模型不尽相同(见表1)。

表1 数字资源描述

机构/平台名称	描述语言	描述模型	概念、属性数量
Nature	XML/RDF/RDFS	Nature本体门户	类(50个),属性(140个)
PMC	XML	JATS	元素(>250个),属性(>130个)
RSC	XML	RSC本体	OBO: 本体(74个),术语(>89.9万个)
			GO: 术语(约2.37万个)
			RXNO: 术语(849个)
ClinicalKey	XML	EMMeT	核心医学概念(25万个)
			同义词(>100万个)
			等级关系同义词(>100万个)
			本体关系的分类系统和本体库(100万个)
中华医学会杂志社	XML	CMA JATS	元素(168个),属性(88个)

注:以上数据均来自网络与科学文献,部分数据截至2016年。

从描述语言来看,语义出版描述语言常用的是XML和RDF,两者在一定程度上是互补的。XML是一种完全面向数据语义的标记语言,具有易控制、易扩展、易综合等特性,但在数据含义交换正确性等方面还存在一定问题。RDF以XML作为编码和传输的语法,其模型描述能力非常强大,可以全面地描述任意复杂资源,但使用技术门槛较高^[29]。应用场景上,XML主要用于出版内容的结构化、碎片化和形式化描述,RDF三元组更适用于后期发布和应用阶段。

从描述模型来看,Nature、RSC语义出版均采用本体技术,极大地提高了出版物检索效率,使信息服务更加具有针对性和专业性。PMC采用的JATS已成为应用较广泛的文献资源描述模型之一,三种针对不同应用场景的标签集可供使用者选择以完成文档的转换、存储及管理,优化工作流程。

3.2 语义知识增强形态

2012年,Shotton归纳了语义出版的实践形式,并定义了语义出版的8项语义增强功能^[30]:(1)丰富对在线论文有机内容的描述,如交互式图形、数据表格、参考文献列表等;(2)增强对论文内容描述的语义标注,如通过命名实体的语义标记,链接术语和概念的描述性定义及有关实体的附加信息;(3)提供与文献相关的其他信息来源链接,如文献作者个人主页、相关的国际机构网址等;(4)提供参考文献的链接;(5)以可操作形式提供文献实验数据的访问,如可下载的电子表格或CSV等;(6)提供对支持该文献研究的全数据集的访问;(7)实现语义相关文献的信息集成;(8)发布文献信息采用通用的数据描述规范,实现开放访问,如内容摘要、文献的详细题录信息,以及参考文献题录信息等。

作者分别从Nature、Elsevier、RSC等选取部分期刊进行调研, 语义增强实现情况如表2所示。

四种期刊均实现了(1)(4)(8)功能, 其他功能发展则相对不平衡, 这表明针对文献的章节图表、基本题

录信息等内容的结构化、规范化描述已成为业界公认的语义出版基本形式, 而涉及部分语义特征描述的语义化功能则有待加强。

表 2 语义增强功能实现情况

语义增强功能	Nature	Elsevier	RSC	国 内
	<i>Nature Chemistry</i>	<i>Cell</i>	<i>Molecular BioSystems</i>	《作物学报》
丰富对在线论文有机内容的描述	√	√	√	√
增强对论文内容描述的语义标注	√	-	√	-
提供与文献相关的其他信息来源链接	√	√	-	-
提供参考文献的链接	√	√	√	√
以可操作形式提供文献实验数据的访问	√	√	-	-
提供对支持该文献研究的全数据集的访问	-	-	-	-
实现语义相关文献的信息集成	√	√	√	-
发布文献信息采用通用的数据描述规范, 实现开放访问	√	√	√	√

此外, 本文对上述期刊/机构已实现的语义出版基本形式范围内的多模态数字资源发布程度进行对比分析, 发现国内外存在区别。(1) 国外期刊/机构提供文本、图、表的形式更加多元化。如Nature允许将图像以PPT文档或其他格式导出使用, 而《作物学报》的图、表仅提供图片格式的下载方式。(2) Nature、RSC等提供多媒体、计算机算法等附加资料, 而国内目前未有实践案例。究其原因, 附加资料的存储、发布、传播、管理涉及复杂的技术、知识产权、标准以及管理机制^[20], 国内外发展水平不一, 其本身也存在一系列有待研究的问题, 如链接机制, 存档及长期保存机制, 复杂技术环境下的封装、交换和传递机制等^[31]。

3.3 语义出版发布形式

语义出版发布形式是语义检索、展示、存储和传播等知识服务层面的直观产物, 不同的发布形式各有侧重。如表3所示, Nature、Elsevier、RSC和玛格泰克的语义出版发布形式是Rich HTML (也被称为Enhanced HTML), 指以HTML标准, 即网页形式呈现文章内容, 并在文章原文基础上进行内容分析、知识标引, 以达到富媒体出版的目的, 为读者提供便利、轻量级、具有互动性的阅读方式。值得指出的是, RSC是最早推出Rich HTML形式语义出版服务的。Rich HTML是科技期刊出版的重要发展方向, 国际上一些优秀期刊出版平台已

相继开展Rich HTML风格及功能特点的完善工作。

表 3 语义出版发布形式

机构/平台名称	语义出版发布形式	特 点
Nature	Rich HTML	网页形式, 富媒体出版, 提供各种便利和各种附加信息
Elsevier		
RSC		
玛格泰克		
PMC	PubReader和ePub	屏幕自适应/支持复杂格式

PMC的发布形式是基于HTML5及CSS3技术的PubReader^[32]和基于XML的开放格式ePub。PubReader可适应任意尺寸的显示屏, 增强用户阅读体验; ePub则对复杂格式支持较好。

除上述发布形式外, 语义出版物发展过程中还存在两个基于知识单元的关键模型——纳米出版物模型和微型出版物模型。纳米出版物模型在开展知识资源的自动搜集、分析和语义检索、过滤时, 可以定位到某一点、结论/实验数据的相关信息或资源, 粒度更小, 层次更深, 可以提高知识的共用、共享程度^[33]; 微型出版物模型具有明显的模块化、结构化、形式化和网络化特点, 但在知识表示能力和内容组织架构上仍存在不足^[34]。

4 结语

在语义技术和数字出版行业的推动下, 语义出版

正稳步向实例化和成熟化迈进,这一趋势不可逆转,传统出版服务将成为学术期刊数字化生存形态的一种补充。

本文对国内外重要语义出版机构/平台的发展实践进行梳理,通过多角度对比分析得出结论:(1)从语义技术和本体技术的发展来看,国内语义出版应用实践相对滞后于国外,作为语义出版的基础,对中文本体库的建立显得尤为迫切;(2)本体技术对语义出版的发展有非常重要的作用,但由于学科领域本体发展的不均衡性,目前语义出版应用主要分布在语义技术应用广泛且实践成果丰富的医学、生物、化学等领域;(3)语义增强功能方面,绝大多数期刊/文献资源已基本实现题录信息、章节图表、参考文献等内容的规范化、结构化描述,涉及语义特征自动抽取的技术内容需进一步研究实践。

总之,语义出版改变了文献出版和学术传播的形态,为学术资源带来全新的组织发布方式,其在高效利用数据集、提高自动化程度及增强用户体验等方面面临的挑战是适应语义及大数据时代发展的新契机。只有针对性地解决语义网、出版和学科领域现有问题,才能更好地促进语义出版发展。本文认为,在推动语义出版研究与实践过程中,可在多模态数字资源内容结构化、碎片化和形式化等知识表示技术,细粒度知识单元识别与抽取、知识组织,多维多源知识语义关联等语义增强方面继续深入研究。同时,还需要出版商、科技期刊和图书情报服务机构在出版机制、知识产权和开放共享等方面开展研究协作,共同推动语义数据驱动的新型学术交流体系的形成。

参考文献

- [1] SHOTTON D.Semantic publishing: the coming revolution in scientific journal publishing[J].*Learned Publishing*,2009,22(2):85-94.
- [2] SHOTTON D,PORTWIN K,KLYNE G,et al.Adventures in semantic publishing: exemplar semantic enhancements of a research article[J].*PLoS Computational Biology*,2009,5(4):1-17.
- [3] SPAR-semantic publishing and reference[EB/OL].[2017-10-19].<http://sempublishing.sourceforge.net/>.
- [4] 王晓光,陈孝禹.语义出版的概念与形式[J].*出版发行研究*,2011(11):54-58.
- [5] HAMMOND T,PASIN M.Linked data experience at Macmillan: building discovery services for scientific and scholarly content on top of a semantic data model[C]//13th International Semantic Web Conference,2014.
- [6] HAMMOND T,PASIN M.The nature.com ontologies portal[C]//5th Workshop on Linked Science 2015,Colocated with International Semantic Web Conference 2015.Bethlehem:2015.h
- [7] Springer Nature SciGraph: A linked open data platform for the scholarly domain[EB/OL].[2017-10-20].<http://www.springernature.com/gp/researchers/scigraph>.
- [8] 邹强,袁庆,康林,等.PubMed Central的数字化出版简介[J].*中国科技期刊研究*,2014,25(2):240-242.
- [9] National Center for Biotechnology Information.PubReader? view of articles[EB/OL].[2017-10-20].<http://www.ncbi.nlm.nih.gov/pmc/about/pubreader/>.
- [10] JATS.Main Page[EB/OL].[2017-10-20].http://webservices.itcs.umich.edu/mediawiki/jats/index.php/Main_Page.
- [11] 包靖玲,李敬文,沈锡宾,等美国NLM DTD3.0期刊存储和交换标签集中文章正文部分标记解读[J].*中国科技期刊研究*,2014,25(4):515-519.
- [12] 康宏宇,侯震,李姣.基于JATS数据标准的全文文献管理[J].*中国科技期刊研究*,2015,26(11):1171-1175.
- [13] National Center for Biotechnology Information.Journal Archiving and Interchange Tag Set[EB/OL].[2017-10-20].<http://jats.nlm.nih.gov/archiving/rationale.html>.
- [14] National Center for Biotechnology Information.Journal Publishing Tag Set[EB/OL].[2017-10-20].<http://jats.nlm.nih.gov/publishing>.
- [15] National Center for Biotechnology Information.Article Authoring Tag Set[EB/OL].[2017-10-20].<http://jats.nlm.nih.gov/articleauthoring/rationale.html>.
- [16] The Royal Society of Chemistry's[EB/OL].[2017-10-20].<http://pubs.rsc.org/>.
- [17] Molecular BioSystems Group.Science come alive[J].*Molecular BioSystems*,2007,3(1):B8-0.
- [18] 翁彦琴,李苑,彭希琚.英国皇家化学会(RSC)——科技期刊语义出版模式的研究[J].*中国科技期刊研究*,2013,24(5):825-829.
- [19] Elsevier.Take on challenges facing the world's medical students[EB/OL].[2017-10-20].<https://www.elsevier.com/about>.
- [20] 彭希琚,张晓林.国际学术期刊的数字化发展趋势[J].*中国科技期刊研究*,2013,24(6):1033-1038.
- [21] RYLEY J.Article of the future[J].*American Journal of Orthodontics & Dentofacial Orthopedics*,2015,148(6):888-889.
- [22] SOUZA E P,CABRERA E M,BRAILE D M.The article of the future[J].*Revista Brasileira De Cirurgia Cardiovascular Orgao Oficial Da Sociedade Brasileira De Cirurgia Cardiovascular*,2010,25(2):141.

- [23] 翁彦琴,彭希琨.爱思唯尔(Elsevier)语义出版模式研究[J].中国科技期刊研究,2014,25(10):1256-1261.
- [24] ClinicalKey[EB/OL].[2017-10-20].<https://www.clinicalkey.com/#/>.
- [25] 李楠,孙济庆,马卓.面向学术文献的语义出版技术研究[J].出版科学,2015,23(6):85-92.
- [26] 沈锡宾,李鹏,王红剑,等.中华医学会系列期刊全文电子文档交换和存储标准初探[J].中国科技期刊研究,2015,26(5):475-479.
- [27] 乐小虬,王子璇,张晓林,等.DPaper:一种面向语义出版的结构化论文写作工具设计与实现[J].现代图书情报技术,2016,32(11):76-81.
- [28] 胡译文.基于WordNet和SUMO本体集成的自动语义检索及可视化模型[J].国家图书馆学报,2012,21(2):23-32.
- [29] 马福诚,刘保良,张明亮,等.XML与RDF的比较分析[C].天津:海军海洋测绘研究所,2008.
- [30] SHOTTON D.The Five Stars of Online Journal Articles,an article evaluation framework[J/OL].D-Lib Magazine,2012,18(1/2)[2017-10-20].<http://www.dlib.org/dlib/january12/shotton/01shotton.html>.
- [31] MARTINSEN D.Strategic Integration of Article Content: Managing Supplemental Materials(part B)[EB/OL].[2017-10-21].http://www.stm-assoc.org/2011_04_26_Spring_Conference_Martinsen_Strategic_Integration_of_Article_Content_Part_B.pdf.
- [32] U.S. National Library of Medicine.NCBITools/PubReader[EB/OL].[2017-10-21].<https://github.com/NCBITools/PubReader>.
- [33] 吴思竹,李峰,张智雄.知识资源的语义表示和出版模式研究——以Nanopublication为例[J].中国图书馆学报,2013,39(4):102-109.
- [34] 王晓光,宋宁远.语义出版物的内容组织架构研究——基于纳米出版物和微型出版物的比较分析[J].出版科学,2017,25(4):20-27.

作者简介

李娇,女,1989年生,硕士,助理馆员,研究方向:知识组织、关联数据、语义检索。

寇远涛,男,1982年生,博士,副研究馆员,硕士生导师,研究方向:数字图书馆理论与技术、信息管理与信息系统。

黄永文,女,1975年生,博士,副研究馆员,研究方向:语义检索、关联数据。

薛欢欢,女,1994年生,硕士研究生,研究方向:信息资源管理。

鲜国建,男,1982年生,博士,副研究馆员,通讯作者,研究方向:知识组织、关联数据、语义出版、信息系统开发,E-mail:xianguojian@caas.cn。

Research on Semantic Publishing Practice

LI Jiao, KOU YuanTao, HUANG YongWen, XUE HuanHuan, XIAN GuoJian
(Agricultural Information Institute, Chinese Academy of Agricultural Sciences, Beijing 100081, China)

Abstract: This paper introduces the practice progress of a few major semantic publishing institutions or platforms, compares them from three aspects of digital resource description model, semantic enhancement and semantic publication form, and then summarizes the gaps in semantic publishing situation between domestic and foreign, the common characteristics and problems of existing practice research. Moreover, this paper discusses the development strategy of semantic publishing model of sci-tech periodicals to provide reference for the research and practice of semantic publishing in China.

Keywords: Semantic Publish; Description Model; Semantic Enhancement; Publication Form

(收稿日期: 2017-10-31)