



开放融合环境下 NSTL 资源建设的发展思考*

□曾建勋

摘要 国家科技图书文献中心(NSTL)自创建以来从印本资源、数字资源、知识资源等方面持续进行科技资源信息保障体系建设,但是在当前开放融合新形势下,仍然面临着资源建设完整性和系统性欠佳、融合性和开放性不足、不平衡不充分、协同性和安全性不够等问题和挑战。面对开放获取、数据融合、API应用、大数据技术下的开放融合新生态,NSTL在未来应该实施元数据战略,建设立体化资源本土长期保存体系,推进多层次元数据的语义知识组织,构建适应发现服务的数字业务管理体系,不断壮大和完善国家科技信息资源建设和服务体系。

关键词 开放融合 NSTL 信息资源建设 资源发现服务

分类号 G250

DOI 10.16603/j.issn1002-1027.2020.06.009

20世纪90年代中后期,我国科技文献特别是国外科技文献订购量一直低水平徘徊,国外科技文献基本满足率不足1/3,严重制约了科技创新活动。为解决我国科技文献资源严重匮乏和保障率极度低下的问题,2000年6月,经国务院领导批准,科技部联合财政部、原国家经贸委、农业部、卫生部、中国科学院等正式组建国家科技图书文献中心(National Science and Technology Library, NSTL)。NSTL的成立标志着我国科技图书馆共建共享体系正式形成。作为国家科技文献基础条件平台,NSTL是国家科技情报系统文献信息保障的核心组成部分,与中国高等教育文献保障系统(China Academic Library & Information System, CALIS)、公共图书馆文化系统并列成为全国三大文献保障体系。

1 NSTL 资源建设二十年的发展历程

NSTL自成立以来,按照“统一采购、规范加工、联合上网、资源共享”的原则,收集和建设理、工、农、医等各学科领域的科技文献信息资源^[1],充分运用现代先进技术手段和条件,开展国家科技文献信息保障体系建设,面向全国提供文献信息服务。经过二十年的发展,NSTL已经建设成为较为完备的以印本为基础的数字资源保障体系,成为国内权威的

科技文献资源收藏保存中心和科技文献信息服务枢纽,为我国经济社会发展和科技创新发挥了不可替代的支撑作用。

1.1 印本资源的基础保障

NSTL成立之初就坚持国家科技文献保障的初心,以印本收藏为基础开启全面印本保障时代。按照“统一规划,专业分工,统筹协调,避免重复”的运行机制,推进印本文献信息资源的全面建设。目前,每年订购印本期刊1.7万种,印本非刊文献8千多种,覆盖全球绝大部分科技类期刊,SCI期刊收录品种超过90%;不但外文文献特别是外文期刊文献的品种数量迅速增加,而且极大地提高了国家经费投入的使用效益,优化了文献资源配置。

二十年来,NSTL紧密围绕国家科技创新和《国家中长期科学技术发展规划纲要》(2006—2020年)的优先主题、重大领域、基础研究、前沿重点领域,充分发挥文献信息资源支撑保障作用,切实保障单个机构或市场难以有效、公平和长期保障的科技文献,确保特色、核心、重点资源收藏的系统性和连续性;不断强化纸本文献的支撑和保障,有步骤地实现较完整的纸本保障和有效备份,加强国外数字资源对应的纸本备份建设,消除高校图书馆向数字资源转型的后顾之忧,支持其他信息机构由印本文献向数

* 国家社会科学基金重点项目“基于知识组织的图书馆资源发现服务体系研究”(编号:17ATQ002)的研究成果之一。

曾建勋,ORCID:0000-0002-0432-9618,邮箱:zeng@istic.ac.cn。



字资源转型;优先建设我国国外科技文献保障体系中结构性缺失的文献,逐步增加符合地方特点以及经济欠发达地区需求的国外科技文献资源;加大对国外科技会议文献、科技报告以及其他灰色文献的建设力度;加强与我国科技合作的重点国家及新兴工业国家和周边国家与地区(如:韩、日、俄、印、巴西及我国台湾地区等)的科技文献采集,合理拓展文献采集的地域和语种范围;调整和优化文献资源的学科、馆藏和载体布局。

1.2 数字资源的引领建设

随着数字出版的不断深入,传统印本资源逐渐无法适应科研人员的信息行为习惯。2005年,NSTL 全面推进数字资源建设,以国家许可方式订购国外回溯期刊数据库 19 个,涉及回溯期刊 3028 种,回溯年代至创刊年,面向全国开通国外专业性中小学协会数据库 40 个,涉及现刊 696 种;面向西部地区开通中文电子图书 26 万册。同时,以补贴方式支持科研机构 and 高校集团采购外文数据库 146 个,为成员单位开通电子期刊 2.5 万种,电子资源年下载使用量超过 6000 万篇,为我国科技发展提供强有力的信息保障。

数字资源的大量涌现需要创新资源建设模式。尤其是大量数字资源采用租用方式引进,无法长期使用和可靠保存,数字资源在利用和共享上受到很大限制。网络环境下,NSTL 强调以构建适应数字时代的国家科技文献资源战略保障体系为目标,坚持“藏用并举”的原则,推动数字资源国家长期保存体系建设,保障国家科技文献信息资源的可靠拥有和长期使用。NSTL 重点加强结构性缺失的科技文献数字资源建设,对仅电子出版的期刊进行补充建设;加强对低保障率资源、低使用率资源的国家保障,有选择地引进回溯数据库,解决早期重点学科资源的缺失问题^[2];推动“国家许可”模式的科技文献数字资源引进和服务,积极组织和支持成员单位及其他科技文献机构的数字资源集团联合采购,对“国家许可”资源实行直接公共服务,做好数字资源的延伸保障和协调保障。同时,NSTL 强化开放获取资源的保障,顺势而为相继加入“OA2020 计划”和 Scoap3 国际高能物理开放联盟,并于 2014 年开始针对学术类开放资源进行规范化采集加工,制订了《NSTL 开放资源描述元数据规范》《NSTL 开放资源遴选标准》,目前每年采集保存 9000 种期刊、800

册会议录、500 份报告、1 万份课件、1 万份学位论文以及 2000 册图书的相关数据,建设了有上千种科技期刊的中文科技期刊开放平台。

1.3 知识资源的转型奠基

随着网络化的不断深入,人们对于知识的需求已不再局限于传统文本,基于更细粒度的知识组织需求愈发旺盛。2006 年,NSTL 启动了国际科学引文数据库建设,目前“国际科学引文库”提供了 6000 种国际高水平期刊的来源文献检索和引文检索服务,为数据挖掘、交叉揭示提供了新途径。在知识组织方面,叙词表作为一种主题检索语言,在资源检索、数据规范、信息分析方面起着重要的作用,2011 年,NSTL 依托成员单位数十年的长期积累,组织实施“面向外文文献的知识组织体系构建及其示范应用”大型项目,发展知识组织基础设施,先后完成“我国数字图书馆标准规范”“汉语主题词表”“英文超级科技词表”“机构名称规范文档”等一大批基础工具。目前,基于英文科技超级词表的文献自动标注和智能检索系统投入使用,大规模语义计算工具包为科技监测、知识演化、学术关系网络分析和领域科研信息环境建设等提供新手段,从而为构建知识资源体系奠定了必要的基础。

NSTL 是一个基于网络环境的公益性的、虚拟的、面向全国的科技文献信息保障服务机构,建成了从联合编目系统、数据联合加工系统、回溯数据服务系统到文献综合管理系统的一整套规范化、科学化的数字业务管理流程,为 NSTL 面向各类机构及个人用户开展知识服务,提供了数字化的业务管理与服务平台。

2 新形势下 NSTL 资源建设面临的问题与挑战

经过二十年努力,我国科技文献信息保障事业取得显著成绩,资源保障水平和服务能力不断提升。与此同时,也清醒地看到,一方面互联网时代数字化和信息技术的迅猛发展,使科技信息从创建、组织到传播、服务的链条发生根本性变化。多变的文献信息形态、开放的协同科研创新流程、在线的学习交流工具和多元的传播获取路径,正在颠覆图书馆原有业务流程和信息服务模式,挑战业界对海量数据的知识组织和管理能力。另一方面,我国科技人员数量、科技研发投入也已位居世界前列,科技发展已经由原来的跟跑阶段,进入到现在的并跑、领跑阶段,



有些领域已进入“无人区”。既需要强化基础研究,又需要强化产业创新;既需要推进关键技术研发,又需要推进双创发展;既需要强化高技术尖端领域发展,又需要促进社会大众学习革新。把握国际研究前沿和创新萌芽的同时,相应的数字化科研范式、新型知识生产流程和知识吸收方式也发生了变化。现有信息服务模式不能适应科研流程发生的根本性变化,既无法嵌入数字化科研范式和知识化生产流程,也无法嵌入工作平台和科研任务过程。因此,面对创新驱动发展战略需求,应对科技发展和科研环境变化的新形势,NSTL 资源建设仍然存在诸多问题与挑战。

2.1 科技文献信息资源建设完整性和系统性欠佳

数字环境下,资源形式更为丰富,服务方式和手段更为先进。在教育领域,以新媒体为核心的翻转课堂、在线教学等不断推陈出新,学术搜索引擎、移动服务也越发深入便捷。新时代下用户对信息资源的认知程度和相关需求更加多元化和精准化,需在满足用户文献发现需求的基础上,加强对于用户的交流发表需求、存档需求、评价考核需求等的满足。目前,NSTL 针对期刊、会议、学位论文等科技文献资源的采集较为完整、系统,学科覆盖面较全,而针对产业创新需求的统计年鉴、智库报告、市场分析报告、产业报告等科技文献采集则较为薄弱。NSTL 的全文资源也仅包含自身馆藏,尚未充分集成其他单位馆藏资源,无法形成更广泛的知识空间联盟;国家元数据库仅拥有 4 亿条元数据,不能满足资源的全面发现与大数据分析需求;需加强产业创新资源的采集,增加专业性较强的数据库资源开发,推进新兴工具类资源的收集与开发,更好地将文献数据资源转化后嵌入到各类科技创新活动中,以全面加强产业自主创新的支撑。

2.2 科技文献信息组织传播融合性和开放性不足

开放理念正在开创新型传播交流范式,推动不同科研条件要素,如文献、科学数据、软件、工作流/门户、仪器及教育资源等多种资源间的开放链接,也形成科技文献内容、技术和数据融合发展的新趋向。近年来,在信息化建设背景下涌现出众多的科技文献资源共享服务平台,在资源支撑保障方面发挥了很大作用。然而,各科技文献服务平台与其他数据平台各自建设,缺乏相互融合关联的方式和互通共享的机制;数据整合、加工精度不够,对元数据的揭

示利用不够,资源间的关联融合缺失:现刊—过刊论文元数据、论文—引文、电子—开放资源数据,尚未全面集成关联;期刊论文、学位论文、科技报告等各类型资源尚未打破资源类型界限,交叉揭示程度不够,无法进行关联检索;人才、机构、项目、技术等实体的规范、识别、关联不够,不便于计量评价。应用系统集成和数据调用存在障碍,用户界面集成融合度不高,系统之间技术标准不一,互操作不够,加剧了信息孤岛现象。在印本为主流时出现的“低水平重复”,在数字形态主导时代似乎又演绎成新一轮数字资源的重复。需要借助众创、众筹等共享经济理念,进一步强化开放共享意识,优化文献资源配置和统筹协调,加强数据间可操作性和关联揭示,强化开放获取平台和开放知识库建设,促进资源开放集成融合^[3],形成开放协同、互联互通、自主可控的科技文献信息开放生态机制。

2.3 科技文献信息公共服务不平衡、不充分

长期以来,我国科技文献保障在不同地区、不同系统间的发展不平衡、不充分。东部地区普遍优于中西部地区,经济发达地区和经济落后地区的文献保障存在着较大差距;高校、科研机构远优于企业,对于众多的小微企业、孵化企业而言,需要科技文献情报的支撑,但受限于高校、科研院所文献服务范围与企业自身的经济能力,众多企业、高新区很难获得文献服务与情报支撑;系统内同样发展不平衡,以高校为例,重点院校每年文献采购经费可以达到千万,而普通高校则经费严重不足。因此,需要 NSTL 依靠国家平台优势,向各级各类公共服务平台开放馆藏资源数据,带动各地区、各行业、各系统的科技文献信息服务,缓解科技文献供给不平衡带来的矛盾。

我国科技文献供给明显存在资源利用不充分的问题,缺乏基于文献数据的分析型数据库和工具等,所呈现的产品层次和服务功能不强,亟待强化信息深度加工处理,发展具有自主知识产权的信息工具和平台^[4]。同时,因为缺乏用户信息的组织和用户画像,缺少用户需求、用户偏好、用户行为、用户习惯的信息采集利用,我国科技文献供给无法实现用户情景敏感的智能推送,以有的放矢地提高文献资源的利用效率和服务效益,释放文献大数据的真正价值。

2.4 科技文献信息保障体制协同性和安全性不够

随着科技文献从印本向数字转变,文献资源订购从资产购置转变为服务许可,购得的仅是“使用



权”而非“拥有权”，数字信息资源的国家拥有和保存面临挑战。首先，国外科技文献出版商仍处于垄断地位，“核心”资源停售停供情况时有发生，资源的在线访问存在中断风险。中国作为世界学术论文产出大国，2019年发表SCI论文超过50万篇，却仅有4.1%发表在中国自己的期刊上。此外，外文学术分析工具与系统平台多依赖国外，通过互联网在线访问国外数字文献资源所留下的日志信息，有可能暴露我国科研动态和研究方向，用户行为面临着极大的泄密风险；我国文献元数据的加工深度不够，本土原创产品供给不足，制约自身信息服务质量、层次的提升和秘密保护的研发。

随着我国自主创新能力的不断提升，国际范围内的科技竞争必然会表现为科技信息竞争与封锁，资源存放在国外，一旦发生自然灾害、国际纷争等不可抗力事件，我国对外文科技文献的访问和使用将遭遇重大阻碍。科技文献保障的国际形势日益严峻、安全风险日渐增大，国家基础平台保障作用将更突出。国家科技文献保障体系需要在保证印本资源订购的同时，对纯电子出版的数字资源、网络资源、开放获取资源等进行完整采集和本地保存，并实现本土化平台访问和保障。

3 新时期信息资源建设的开放融合新生态

回顾图书馆系统的发展历程，从相互之间无关联的图书馆自动化系统，实现机读目录，到出现联邦检索、OCLC的联合目录，再到现今的资源发现系统，实际上贯穿着一种渐进式开放融合的理念。就资源发现系统来说，其实质是在继承搜索引擎原理的基础上，发扬图书馆联合目录的优势，只不过不仅仅是在图书馆间进行开放协同服务，还扩展到数据库商、出版社之间的开放获取和协同服务；资源发现系统在实现图书馆数据聚合、业务系统整合、服务协同以及工具融合的基础上，将信息检索、文献获取与图书馆用户认证系统、借还书系统、原文传递系统、代借代查等服务协同链接起来，并利用开放应用程序接口(Application Programming Interface, API)服务、开放链接和大数据、云计算、知识组织和人工智能等技术，最终实现集资源发现、文献获取或传递、收录引证、学科分析、科研评价于一体的一站式集成服务。发现服务的兴起标志着一种新型的开放融合生态正在跨界形成。

3.1 开放获取正在形成信息传播新机制

开放获取运动的出现，开放创新、开放科学、开放评价、开放出版、开源社区的蓬勃兴起，致使开放资源、开放数据、开放期刊、开放仓储越来越多。开放存取作为一种全新的学术传播机制，正在推动学术资源和科研成果的自由传播，提升公共信息利用水平。目前，许多国家或组织已经出台强制性开放获取政策，要求公共资金资助科研成果存储到开放知识库，实行开放共享。同时，开放获取期刊和论文资源迅猛增长，不仅涌现出PubMed Central、PLoS等开放获取出版平台，牛津大学出版社等传统出版机构也逐步开展开放获取实践。开放存取模式下，科研人员不需要通过付费订阅就能访问学术信息的全文，阅读、下载、拷贝、传递、打印、检索、超链接该文献，这彻底改变了用户获取和使用学术信息的渠道、方式、需求甚至观念。开放存取资源免费向公众开放，消除价格障碍，降低资源建设和文献获取的门槛，打破资源建设和服务的原有格局，将对图书馆传统的资源建设、资源存储、知识产权管理、资源保障服务模式造成巨大冲击。

3.2 API应用正在推进信息服务新业态

通过API实现跨界的开放式调用，形成一个开放、灵活、自主、优化的合作体系。通过数据交换对价值链进行整合、重塑与创新，实现信息服务业的“互联网+”融合发展，将大大减少数据和应用“孤岛”，释放更多计算能力，实现价值重塑。它意味着将图书馆馆藏数据、数据商出版数据、网络开放数据等数据资产化，把图书馆、数据商和相关单位知识库等拥有的数据、服务和业务能力以API的形式开放给生态系统各参与方，实现业务能力的互联互通，带动数据—软件开发—信息服务整条产业链的创新发展。滴滴平台调动社会车辆提供租车共享服务；维基百科借助社会力量构建百科知识服务于公众；百度文库资源众筹共享，表明无论是基于服务保障的全文长期保存，还是基于登记注册的元数据集成融合，抑或是基于开源的软件工具融合，都可以通过图书馆、出版社、数据库商、开发商之间的合作，借助API应用共同构建，使更多的资源拥有者、标注者、开发者参与到资源共建共享中来。通过实现交流或交易的媒介功能、信息服务功能、产业组织功能和利益协调功能，来拓展新的价值网络，带动信息服务业态创新，加快信息服务转型升级。



3.3 数据融合正在塑造内容产业的新模式

科技文献出版正在从印本出版、印本与电子资源并存,发展到纯电子出版,甚至语义出版,新型出版方式和出版物形式不断涌现,非正式交流形式日益丰富,学术博客、网络社区、即时交流工具等蓬勃兴起,媒体融合呈现立体化出版趋势,既反映出全球科技发展轨迹变化和时代创新特征,又反映出数字信息形态变革和科研范式变化。知识单元不仅包括图片、表格、公式、人物、机构等,还包括知识点、概念、等级、本体关系等复杂关系。这些知识单元通过 XML 系列置标语言进行描述和标记,并在知识组织体系中的公理、命题的基础上,与特定领域以及跨系统的各种知识组织工程相结合,支持可计算、可推理的智能语义检索与语义知识发现。越来越多的出版机构,在推进语义出版、支持各种细粒度的知识单元关联与计算,如 Springer 推出科研图谱(SciGraph)服务,将传统集中式、以文献为中心的出版平台转换为分布式、以事件为中心和以 RDF 为基础的复合语义架构。除 Springer 外,还有一大批国际学术出版商,已经在语义出版方面进行诸多实践,无论是 Wiley 的“Smart Article”,还是 Elsevier 的“Article of the Future”“ClinicalKey”都成果卓著。同时,数字技术和数字环境在颠覆传统资源形态的同时,使文献资源的揭示程度从书目层级逐步深化到篇章、图表、科研实体等知识单元,出版传播手段从文本拓展到了交互文本、可执行论文、音视频等等。各种载体、数据、内容、工具资源的跨界融合越来越强烈,出版商、集成商在拓展原有数据库、元数据应用模式的同时,也在深化文献的资源描述与信息服务,全面改造信息资源建设与服务模式。

3.4 大数据技术正在注入信息时代的新动能

大数据、云计算、人工智能等新兴技术的迅猛发展,给科技文献保障体系建设带来前所未有的技术挑战。大数据环境下,海量数据存储、复杂数据处理、大数据深度挖掘和预测分析服务对图书馆的数据存储能力、计算能力、知识组织能力、数据分析能力提出严峻挑战,并将极大地改变图书馆信息资源建设、资源存储和保护、信息发布和传播、知识化服务的现有模式。云计算作为一种以数据为中心的新型网络计算方式,把资源、计算能力和领域知识组织有机融合,形成认知计算型知识组织工程,并与出版融合应用,形成基于计算、分布存储的动态数据云服

务,将深刻地影响未来图书馆的互联网运作和服务模式,改变图书馆信息获取、存储、检索和共享的方式与理念。当前人工智能已成为引领未来的战略性技术,深度学习、人机协同、自主操控、知识学习、自主智能等人工智能技术在深刻改变人类社会生活的同时,将会给图书情报行业的技术策略、技术体系、技术基础设施、业务系统、智能化服务带来革命性变化。人工智能在信息智能获取、智能参考咨询、智能检索、语义数字图书馆、图像识别等领域的直接应用,将推动科技文献信息保障和服务从“互联网+”进入“智能+”时代。

总之,各类新技术的复杂性和颠覆性,对图书馆的技术应用能力和知识服务能力带来前所未有的技术挑战,必将颠覆图书馆传统业务模式和技术体系,催生图书馆新机制、新业态、新模式和新动能。开放是颠覆性技术,更是颠覆性机制,构建开放融合生态环境支撑的数据融合和语义组织体系,推进基于智能认知和知识组织的开放服务体系是大势所趋。当下,数字图书馆、数字出版、学术搜索引擎在语义环境下走向融合归一,在不断开放融合的生态环境里,图书馆需要认清自身功能定位,塑造科技信息保障核心能力。

4 开放融合环境下 NSTL 资源建设发展策略

科技文献信息的生产、组织、分析、传播和服务等正进入数据化、工程化、智能化、聚合化和可计算的新时代,需要基于开放共享理念,实现跨系统、跨产业、跨领域的协同发展,重组信息资源建设格局,形成从出版社、图书馆到终端用户高度开放融合的有机整体。为此需要资源建设树立“资源即数据”“系统即服务”“用户即资源”的新理念,强化元数据规模化建设,实现大规模语义化、关联化知识组织、工程化数据处理,构建具有知识关联特征的知识服务系统。

NSTL 资源建设总体策略如图 1 所示,需要重构文献信息资源建设渠道,推进元数据战略,适应开放融合环境,建设国家元数据库体系;与此同时,一方面需要构建全文保障协同服务机制,适应国家信息安全保障需要,建设立体化资源本土长期保存体系,另一方面构建多层次知识组织流程和工具体系,适应高层次的科研管理需求,建立知识单元开放关联的结构体系。而这些都是需要重构数字业务管理体系,适应知识发现服务潮流,提升资源建设的效率和效益。

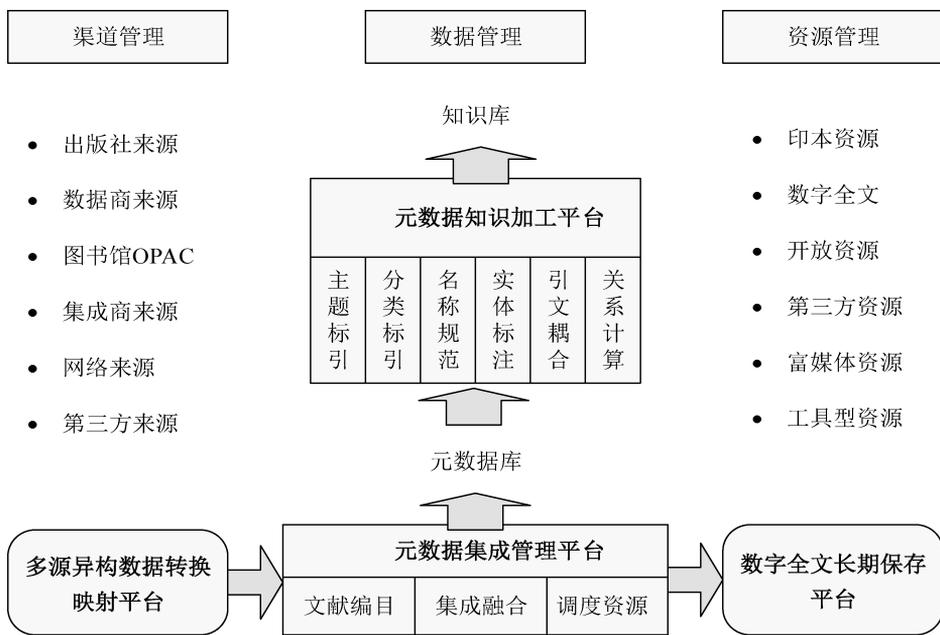


图1 NSTL 资源建设总体策略

4.1 实施元数据战略,建设国家元数据库体系

经过数十年的建设,国内高校图书馆、公共图书馆和科技图书馆均积累了丰富的未经集成的元数据资源。NSTL 作为国家级保障机构,需要全面实施科技文献国家元数据战略,最大限度构建国家元数据库,通过开放国家资源目录,同各公共馆、高校馆、专业馆一道,构建统一的元数据集成联盟。探索基于知识产权管理、市场化激励措施的元数据登记注册集成共享机制,大力拓展元数据的多来源获取渠道,加强对订购资源、开放资源、网络资源等元数据的及时发现、采集、规范和保存,实现图书馆、出版社、集成商等多源异构文献元数据的集成融合,编制外文科技文献联合总目录,建立中国科技信息资源的“大”元数据体系。按照元数据收割/导入、转换、校验、集成、查重、归一等环节设计元数据集成融合流程,制定元数据统一标准规范和描述模型,实行多来源元数据格式的映射登记,建立元数据集成融合系统,实现多源元数据在数据、信息和语义三个层面的集成整合,形成覆盖元数据资源采集获取、集成整合、信息服务等全流程的功能模型,并不断提高元数据采集的完整性和更新的及时性,形成国家元数据库的可持续发展机制。

当前,文献资源的揭示粒度不断细化,元数据的本质内涵不断加深,早已不再局限于书目揭示,转而形成一个包含多层级内涵的体系。在统一标准框架

下,构建集多来源、多载体数据融合的元数据战略体系,集国家印本资源、国家数字资源、第三方馆藏、第三方全文保存于一身的全文协同保存体系,集原文传递、代查代借、链接获取于一体的资源调度体系,集分类法、叙词表、机构文档、母体文档于一体的语义组织体系,以及面向用户本体和用户画像的用户实名注册管理系统。

4.2 基于全文协同保障,建设立体化资源本土长期保存体系

构建我国自主可控的科技文献信息开放生态环境,需要调动全社会相关力量拓展资源采集渠道,建立元数据注册登记的集成仓储机制、数字全文分布保存与协同保障机制;在构建联合目录的基础上,联合各家实现原文多途径的协作保障,大范围地推进馆际互借,构建更广泛的共享空间。构建基于公共资金权益保障的数字资源本土长期保存格局,通过分工合作,吸引高校、科研院所共同参与,推动国际重要科技文献数字资源全部实现本土长期保存,形成集中与分布相结合、数字全文长期保存和原文获取服务于一体的数字资源长期保存体系与协同服务保障格局;针对西方限制的科技文献品种,巩固商业化市场渠道的同时,开拓民间交流获取渠道,强化卡脖子资源的长期保存;提升数字资源长期保存的自我管理能力,实现我国原生数据资源、国外引进数字资源、网络开放资源等资源的长期保存和可持续更



新,增强抵御信息风险的能力。

在保障中外文期刊、会议、学位论文、科技报告等一次、二次文献的基础上,增加智库资源、年鉴、进展、综述、述评等三次文献的采集;规划新媒体、科学数据、可视化图谱等新兴资源建设,适时推进富媒体学术资源、事实型数据资源、术语型组织资源、软件型工具资源等建设和保障,形成媒体融合时代的立体化资源保障体系;在巩固完善面向学术研究的资源建设基础上,大力开拓面向产业创新的资源建设,加强市场报告、产业分析、统计手册等资源采集,形成学术信息资源与产业信息资源相融合、适应知识增长和创新需求变化的一体化创新资源体系,有序地向多元化载体和新型信息内容资源建设拓展。

4.3 推进多层次元数据的语义知识组织,提升大数据增值效应

通过人工智能技术与语义知识组织的优势互补和深度融合,强化元数据集成、数字知识表示、语义组织关联和用户认知模型研究,加强知识组织体系及其工具的开放利用,深化多粒度的科技文献信息深度组织与知识揭示,大力支持第三方应用服务开发;强化科技文献资源与科学数据、科技项目、科研机构等资源的关联,探索对文献目录、章节知识点、内容提要、实体、表格、图片、公式、概念等进行细粒度加工和关联,提升对现有资源的揭示力度,不断提高文献大数据的知识处理能力,构建高附加值资源库;通过可泛化的知识学习与计算工具引擎,采用新一代科技文献智能知识服务关键技术与产品,实现大规模语义知识库自动构建、知识计算推理与知识服务可视化交互等,形成统一集中的科技文献资源聚合管理服务体系。推进文献资源的深度聚合和集成可计算,按照预先设计的语义框架揭示术语表示的资源语义关系,逐步形成可检索、可关联、可计算的一体化开放知识体系。

4.4 构建适应发现服务的数字业务管理体系,引领资源建设新模式

大数据环境下,图书馆的业务流程正在发生翻天覆地的变化,未来的服务管理平台不是一种简单的统一集合,而是一个多元工具、算法、应用的融合集成系统,需要实现知识组织体系的开放应用,如自动标引、识别和分类、导航、评估、链接、计量等新业务,以及分析工具协同开发与开放应用,如需求感知、主题监测、用户画像、个性推送、精准服务等新服

务的开拓与深化。需要构建适应大数据的数字业务新流程,在海量元数据集成的基础上,通过书目一体化管理系统、元数据集成管理系统和知识加工处理系统,完成对元数据的规范统一、语义化处理;通过高度模块化、可扩展、可定制、个性化应用的信息服务平台,采用 APP 方式实现图书馆各业务管理和服务功能模块的按需安装使用。需要强化文献资源建设业务系统间的相互协同和衔接,在对 NSTL 现有的综合管理系统、联合目录系统、开放资源管理系统、非析出加工系统、数据加工系统、引文加工系统进行功能增强的优化改造基础上,按照元数据统一标准规范和描述模型,建立多源异构元数据集成融合与规范加工系统,构筑支持各类数据的加工、融合、计算和管理的大数据基础设施,提升数据管理能力。这样,加快推进 NSTL 资源建设从采购管理向渠道管理转变,推动数据管理业务从简单数据加工向数据集成融合转变,推动文献征订遴选、联合编目、资源加工、数据仓储等业务系统重组整合,构建新型的基于数字资源、兼顾印本资源管理的全新数字业务管理体系。

5 结语

随着开放融合新生态的形成,原有的图书馆体系在重构之中也会焕发新生,将打破现有公共图书馆、科技图书馆、高校图书馆的格局,集合国家馆、公共馆、高校馆、出版社及网络资源,形成“全覆盖”馆藏,转型为由研究规划型图书馆、资源建设型图书馆、技术开发型图书馆和信息服务型图书馆组成的宝塔形的新格局,推进嵌入团体的相关文献服务、知识服务、智库服务、分析服务、数据服务等。走进新时代, NSTL 将继续坚守国家科技文献资源保障的初心,完善可持续发展的开放创新机制,坚持国家科技文献元数据战略,实施国家科技文献本土化长期保存战略,加强三次文献、新媒体资源、开放资源、科学数据、工具算法资源的集成融合,佐之以实名认证的发现系统,拓展基于情景敏感的知识化、智慧化、个性化服务方式,支持非营利性服务,孵化市场化服务,打造一个融合多源异构大数据的本土化公益服务平台。随着越来越多的图书馆元数据、数字资源及其业务系统、服务系统逐步实现云端化, NSTL 将在开放融合新生态中携手内容行业各类主体,不断壮大和完善国家科技信息资源发现服务体系,形成



互利共赢的科技文献保障命运共同体,携手建立科技文献保障服务的“新业态”,开创科技信息资源保障事业发展的新篇章。

- 3 张晓林,孙坦,刘细文,等.数字时代国家科技文献中心的战略选择[J].图书情报工作,2009,53(1):42-46.
- 4 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,2011,37(2):30-35.

参考文献

- 1 彭以祺,吴波尔,沈仲祺.国家科技图书文献中心“十三五”发展规划[J].数字图书馆论坛,2016(11):12-20.
- 2 沈仲祺.构建开放协同的服务机制[J].医学信息学杂志,2010,31(1):5-10.

作者单位:中国科学技术信息研究所,北京,100038

收稿日期:2020年5月9日

修回日期:2020年10月29日

(责任编辑:支娟)

Reflection on the Development of NSTL Resource Construction in an Open and Fusion Environment

Zeng Jianxun

Abstract: Since its establishment, National Science and Technology Library (NSTL) has continuously carried out the construction of an information resource system of science and technology including printed resources, digital resources, and knowledge resources. However, under the current environment of open integration, NSTL still faces problems and challenges, which are resources construction is far away from comprehensive and systematic, integrated, open, balanced, coordinated and safe. Facing the new ecology of open access, data fusion, API application and big data technology, NSTL should implement metadata strategy, build a three-dimensional local long-term preservation system of resources, promote the semantic knowledge organization of multi-level metadata, develop a digital business management system suitable for discovery services, and constantly expand and improve the construction and service of national science and technology information resources system.

Keywords: Open Convergence; NSTL; Information Resource Construction; Resource Discovery Service

=====

(接第 54 页)

An Analysis of Statistical Data of Chinese Academic Libraries in 2019

The Secretariat of Steering Committee for Academic Libraries of China

Abstract: In 2019, the average of document resources purchase fee decreased in Chinese academic libraries. While the print version document resources purchase fee continuously showed a dynamic decline trend, the total purchase fee and the proportion of electronic resources exceeded that of the print version document resources, and the average purchase fee of libraries' electronic resources was less than that of the previous year for the first time. The library building acreage has stopped growing. The professional title of library directors mainly composed of professors and associate professors. The total number of contract library staffs and student assistants tended to be stable.

Keywords: Academic Libraries; Document Resources Purchase Fee; Building Acreage; Human Resources; Digital Resources; Information Resources Sharing

2020年第5期
大学图书馆学报