

我国数字图书馆标准规范建设内容及特点分析^{*}

王秀香

(国家图书馆, 北京 100081)

摘要: 本文从数字图书馆标准规范内容的角度来梳理我国数字图书馆标准规范建设的主要成果, 涵盖数字资源建设与加工标准, 数字资源内容描述、组织和整合标准, 数字图书馆服务标准, 数字图书馆系统互通互联的互操作标准, 数字资源长期保存标准, 以及数字图书馆体系化发展标准六方面, 并总结我国数字图书馆标准规范建设的四个主要特点: 积极采用已有研究成果和标准, 结合项目自身需求研制标准规范, 体系化建设特点明显, 注重合作与共享。

关键词: 数字图书馆; 标准; 指南

中图分类号: G250

DOI: 10.3772/j.issn.1673-2286.2016.9.003

标准规范建设是保证数字图书馆有序、稳定、高效并可持续发展的基础, 也是数字图书馆能够长期发挥作用的必要条件。它对数字资源建设的规范化和标准化, 实现资源共建共享, 促进信息资源整合, 降低数字图书馆建设成本, 提升数字图书馆服务效能等方面具有重要意义。

1 我国数字图书馆标准规范建设的发展脉络

20世纪90年代, 我国开始研究与建设数字图书馆, 数字图书馆建设大致经历数字资源原始积累阶段(以文献信息资源的数字化加工为核心)、关键技术攻关研发阶段(围绕数字图书馆体系建设、海量数字资源的存储与检索、数字资源的发布与服务、多媒体和异构资源的整合处理、大规模并发访问控制等核心技术问题进行研发)和数字服务集成应用阶段(强调知识发现与知识交流, 并试图构建以数字图书馆为核心的知识网络)三个发展阶段^[1]。

数字图书馆标准规范内容建设与数字图书馆的发

展进程紧密相关, 其内容建设也为满足三个发展阶段的需求而逐步完善与健全起来。为适应数字图书馆发展初期对馆藏资源大规模数字化加工的特点, 数字图书馆标准规范建设从倾力于数字资源建设、加工与描述的标准规范建设, 到数字图书馆服务的相关标准, 再到数字图书馆标准规范体系构建, 数字图书馆标准规范建设的发展脉络可分为六个阶段: (1) 围绕数字资源建设与加工的标准规范建设; (2) 围绕数字资源内容描述、组织和整合的标准规范建设; (3) 围绕数字图书馆服务的标准规范建设; (4) 围绕各数字图书馆系统互通互联的互操作标准规范建设; (5) 围绕数字资源长期保存的标准规范建设; (6) 围绕数字图书馆体系化发展的标准规范体系建设^[2]。

数字图书馆标准规范建设是随着数字图书馆的发展与需求而不断演进的, 从单纯的资源数字化到数字资源的组织与揭示, 从各系统单独的服务标准到互联互通环境下数字图书馆体系化服务标准。可以说, 数字图书馆标准规范建设是一个由点到面、由浅入深不断完善的过程。

^{*} 本研究得到国家社会科学基金项目“我国数字图书馆标准规范体系及其构建机制研究”(编号: 12BTQ013)资助。

2 我国数字图书馆标准规范建设的主要内容

数字图书馆标准规范有其自身发展规律和特点。本文按照数字图书馆标准规范发展脉络,从数字图书馆标准规范建设内容来梳理各阶段的特点,包括数字资源的加工、组织、服务、互操作、长期保存与体系化建设六个方面^[2]。

2.1 数字资源建设与加工相关标准规范建设

对馆藏文献的大规模数字化加工是数字图书馆发展初期的核心任务,数字资源建设与加工的标准规范研制成为当务之急。数字图书馆标准规范建设初期,在内容上主要是对象数据加工标准规范的制定,包括各种类型数字对象,如文本、图像、音频、视频等在创建与加工过程需要遵循的标准规范,如数字文件编码、数字文件格式、数字文件标识等标准以及数字资源选择与加工程序等操作规范。

数字文件编码是将文字、图形、图像、语音及各种属性信息转换为计算机能够处理、存储和传输的数字编码(包括字符编码和结构化信息编码)。国际普遍要求基本编码标准应遵循ISO/IEC10646/UNICODE;我国汉字编码处理有关字符编码的国家标准主要有GB 2312—1980(《信息交换用汉字编码字符集——基本集》)和GB 18030—2000(《信息交换用汉字编码字符集基本集的扩充》)。数字文献结构编码涉及如何定义文献结构,按照国际普遍要求采用XML DTD/XML Scheme定义。数字文件格式包括保存格式、浏览格式和预览格式等。数据格式方面的标准非常多,不同资源类型、不同应用级别需要采用不同的格式标准。目前在数字资源建设中主流的对象数据格式标准主要有ISO 19005-1(PDF/A)、ISO/IEC 10918-1(JPEG)、ISO/IEC 15444(JPEG2000)、ISO/IEC 11172(MPEG-1)、ISO/IEC 13818(MPEG-2)、ISO/IEC 14496(MPEG-4)及一些工业标准(如TIFF等^[3])。数字内容标识方面的标准主要有URI、DOI、SICI、BICI等。其中,URI是所有标识互联网资源的地址和名称的通用集合,包含URN和URL;DOI用来标识数字环境中的内容对象;SICI是标识期刊和期刊包含文章的标识符;BICI是标识图书和图书包含内容的标识符^[3]。

国内大型的数字图书馆建设项目,大都围绕数字资源建设与加工制定系列标准。20世纪90年代,“中国试验型数字图书馆”项目和中国高等教育文献保障系统(China Academic Library and Information System, CALIS)数字图书馆项目开展存档标准、数字加工标准等内容的研制^[4];大学数字图书馆国际合作计划(China Academic Digital Associative Library, CADAL)围绕数字资源加工制定数字对象制作规范(如《文档数字对象制作规范》《图片数字对象制作规范》《音频数字加工标准与操作规范》《视频数字对象制作规范》等)和数字对象编码标准(如《数字内容编码与内容标记规范》)^[5];国家数字图书馆工程研制了汉字处理规范、数据(如文本、图像、音频、数据等)加工规范和对象标识规范(如《数字资源唯一标识符规范》)。

国家标准有《图书馆馆藏资源数字化加工规范 第2部分:文本资源》(GB/T 31219.2—2014)、《图书馆馆藏资源数字化加工规范 第3部分:图像资源》(GB/T 31219.3—2014)、《图书馆馆藏资源数字化加工规范 第4部分:音频资源》(GB/T 31219.4—2014)、《图书馆馆藏视频资源数据加工》(未结项)。全国图书馆标准化技术委员会颁布的行业标准:《数字对象唯一标识符规范》(WH/T 48—2012)、《图像数据加工规范》(WH/T 46—2012)、《音频数据加工规范》(WH/T 49—2012)和《文本数据加工规范》(WH/T 45—2012),这些标准互为补充,共同规范并促进图书馆资源的数字化加工工作。

2.2 数字资源内容描述、组织和整合相关标准规范建设

随着数字图书馆的进一步发展以及数字化信息处理能力的逐步提升,对数字资源的加工开始深入到文献内容,重视对文献内容的组织与揭示,数字图书馆标准规范建设内容也逐步拓展到对数字资源内容的描述、组织和整合领域。

数字资源内容的描述是对资源的结构化定义,涉及的标准规范主要包括元数据标准、元数据置标、元数据互操作等;数字资源组织与整合涉及的标准规范主要包括对象管理、知识组织等^[1]。

元数据标准指描述某些特定类型资料的规则集合,一般包括语义层次上的著录规则和语法层次上的规定。语法层次上的规定包括描述所使用的元语言、文

档类型定义使用的语法、具有内容的元数据格式(含内容数据)及其描述方法。元数据标准实现的功能包括对资源的描述、管理和定位,以及对资源的评估^[6]。目前,国际图书馆界常用的元数据标准:用于文献资源描述的MARC、Dublin Core;用于档案资料描述的EAD;用于视听资料描述的LC-Audiovisual Metadata;用于数字图像描述的MOA2 metadata、CDL metadata、VRA core等。除描述元数据外,技术元数据、管理元数据、保存元数据、元数据互操作等相关标准也逐渐发展与日益完善。知识组织的标准化是实现资源利用、整合与共享的基础,包括本体、关联数据、叙词表编制标准等。

为促进数字资源利用与管理,国内各大型数字图书馆工程均围绕数字资源描述与组织制定系列标准。国家数字图书馆工程制定了核心元数据规范、专门元数据规范、管理元数据规范、保存元数据规范、元数据编码、元数据映射等系列元数据规范,并专门制定知识组织规范^[7];CADAL制定了多维度标签分类标准、知识组织服务标准和数字图书馆知识标引规范,如知识元抽取规范、学科分类标准规范、学术水平等级切分标准等系列知识组织标准规范^[8];军队院校数字图书馆制定了分类法、主题词表、本体、知识组织体系元数据描述等知识组织标准规范;全国文化信息资源共享工程编制《文化共享工程视频资源建设格式(V1.0)》《全国文化信息资源共享工程视频资源编目规范》等。各图书馆围绕自身数字资源建设需求,建立一系列元数据规范,如上海图书馆的《名人手稿馆元数据方案》、北京大学图书馆的《拓片元数据标准》等。

全国信息与文献标准化技术委员会(SAC/TC4)于2010年发布国家标准《信息与文献都柏林核心元数据元素集》(GB/T 25100—2010);全国图书馆标准化技术委员会组织制定《网络资源元数据规范》(WH/T 50—2012)、《图像元数据规范》(WH/T 63—2014)、《管理元数据规范》(WH/T 52—2012)、《音频资源元数据规范》(WH/T 62—2014)、《视频资源元数据规范》(WH/T 63—2014)、《电子连续性资源元数据规范》(WH/T 64—2014)、《电子图书元数据规范》(WH/T 66—2014)、《古籍元数据规范》(WH/T 66—2014)、《期刊论文元数据规范》(WH/T 67—2014)、《学位论文元数据规范》(WH/T68—2014)等系列文化行业标准,这些标准对规范全国元数据加工与管理工作起到积极作用。

2.3 数字图书馆服务相关标准规范建设

21世纪初期,数字图书馆建设从试验性阶段进入建设与服务并进的阶段,数字服务成为图书馆服务的重要内容。数字资源服务区别传统服务的许多问题开始凸显,因此,在进行针对性研究和探索的基础上,相关服务及管理类标准也成为数字图书馆标准规范的重要内容,包括数字资源的检索标准、网络服务标准、应用服务标准、版权管理标准等^[1]。

数字图书馆推广工程制定的服务标准,包括资源发布规范、检索服务规范、应用服务规范三类;并制定管理标准,包括长期保存规范、版权管理规范、数据管理规范、用户管理规范、其他相关规范(数字资源统计、评估等)五类;CADAL制定数字版权管理标准规范集,包括《第一部分:数字资源版权加工许可导则》《第二部分:数字资源版权安全管理技术导则》《第三部分:数字资源传输版权管理导则》《第四部分:数字资源发布版权维护导则》《第五部分:数字资源用户服务版权保护导则》^[9];军队院校数字图书馆制定了业务活动标准规范 and 信息服务其他标准规范,业务活动标准规范包括参考咨询服务技术标准规范、馆际互借标准规范、信息报道和推送服务标准规范、开放流通服务技术标准规范,信息服务其他标准规范包括用户身份统一认证、用户管理规范、安全和隐私保护标准/知识产权协议、计费与结算标准规范、资源调度接口标准规范、信息环境服务等级应用纲要;CALIS委托北京大学图书馆起草《CALIS虚拟参考咨询服务工作》。

国家标准有《信息与文献 开放系统互连 馆际互借应用服务定义》(GB/T 23269—2009)、《信息与文献 开放系统互连 馆际互借应用协议规范 第1部分:协议说明书》(GB/T 23270.1—2009)、《信息与文献 开放系统互连 馆际互借应用协议规范 第2部分:协议实施一致性声明(PICS)条文》(GB/T 23270.2—2009)、《信息技术 数字版权管理 术语》(GB/T 30247—2013)等。全国图书馆标准化技术委员会发布了《图书馆参考咨询服务规范》(WH/T71—2015)文化行业标准,以规范并推动参考咨询服务的发展^[10]。

2.4 数字图书馆互操作相关标准规范建设

在数字图书馆资源和服务不断发展的同时,各类型数字图书馆系统间的互操作标准也开始引起重视。

在数字图书馆建设初期,各数字图书馆有分布性、自主性和异构性的特点。分布性指物理上分布在不同地方,自主性指由多个不同厂商开发或提供数字资源,异构性指数字资源的存储结构、数据模型、表现形式等方面的不同。这些特点阻碍数字图书馆的互联互通、资源共享。互操作指两个及以上系统或组件进行信息交换和信息利用的能力,因此,提高互操作性是数字图书馆网络化、体系化发展的关键。

数字图书馆的互操作性主要体现在系统间数据交换和服务协作两个层面。具体到研究领域,包括元数据的研究和中间件研究,如互操作协议研究^[11]。目前应用较为成熟的互操作标准有Z39.50、Open Archive Initiative和Z39.88。

国内大型数字图书馆工程对互操作问题非常重视,并有针对性地制定相关标准。数字图书馆推广工程制定元数据收割协议等互操作标准;军队院校系统制定用户服务接入标准规范、数据应用标准规范、数据传输标准规范、数字资源检索协议标准规范、网络集成服务标准规范、Web基础信息服务标准和协议、Web站点标准规范等系列互操作标准规范;为确保CADLIS各应用系统间顺利集成,CALIS针对主要应用系统制定相关接口规范,还为保证第三方软件编写接口参考实现方案,如统一认证接口规范、数字对象唯一标识符本地解析规范、数字对象唯一标识符参考实现方案、异构资源元数据检索接口(ODL)规范和开发指南、OpenURL资源调度规范等。

2.5 数字资源长期保存相关标准规范建设

随着数字资源的快速增长和快速湮灭,图书馆在进一步加强对这些资源的采集、加工、整理和服务的同时,开始重视资源的长期保存问题。因此,与长期保存相关的文件格式、描述格式、数据转换、平台技术、框架结构、长期保存元数据等标准也被纳入数字图书馆标准规范建设内容体系。

在数字资源长期保存领域,最有影响力的标准是1999年美国空间数字系统咨询委员会提出的OAIS,该标准被国内外众多图书馆数字资源保存项目采用,国家图书馆制定的《中文元数据方案》也采用此标准^[12]。此外,《基于文档的电子信息的长期保持》(ISO/TR 18492)也为数字资源的长期保存提供较好的方法和建议^[1]。

在数字资源长期保存元数据方面,目前还没有统

一通用的数字信息资源长期保存元数据规范格式,一些国际或国家级的数字资源保存项目都有各自的元数据格式。由联机计算机图书馆中心(Online Computer Library Center)推出的PREMIS是数字资源长期保存领域的一个准行业标准,现已被许多项目采用。

20世纪90年代,我国陆续制定和颁布一些电子文件管理标准和办法,如《电子文件归档与管理规范》(GB/T 18894—2002)和《电子公文归档管理暂行办法》等^[13]。我国数字图书馆建设相关机构在积极吸纳借鉴国外已有成果的基础上,结合我国实践,在数字资源长期保存领域进行深入研究,取得较好成效。科技基础性专项资金重点项目我国数字图书馆标准与规范建设二期研究计划,将数字资源长期保存有关标准的研制作为重点内容。国内各大型数字图书馆建设项目也研制了专门的数字资源长期保存标准,如国家数字图书馆工程的《国家图书馆数字资源长期保存规范》;数字图书馆推广工程的《长期保存元数据规范》;CADAL的“数字资源仓储标准规范集”,包括《资源保存与管理规范》《保存元数据标准》等相关标准^[9]。

为规范我国数字资源长期保存的格式,全国信息与文献标准化技术委员会已将ISO 28500立项为国家标准。为规范并促进我国数字资源的长期保存,2011年全国图书馆标准化技术委员会立项《图书馆数字资源长期保存元数据规范》(WH/Z1—2012),该标准采用PREMIS2.0,结合我国图书馆等文化行业机构对数字资源长期保存元数据描述、揭示及使用的实际需求,做出相应规定;2013年立项国家标准《图书馆数字资源长期保存信息封装规范》。

2.6 数字图书馆体系化发展相关标准规范建设

数字图书馆间的互联互通,使数字图书馆体系化发展已成为必然趋势。而标准规范是数字图书馆体系化发展的基础,体系化的标准规范是保障数字图书馆网络化、体系化发展的基本需求。数字图书馆相关标准规范的涵盖范围从单纯数字资源加工与描述标准,逐渐扩展到数字资源生命周期全过程的标准(涵盖数字资源建设、服务、保存、技术、管理等多方面的标准规范),再发展为涵盖网络环境下数字信息系统互操作、数字图书馆安全、数字图书馆评估等多方面的标准和规范。

我国数字图书馆标准规范建设项目从数字图书馆标准规范体系建设的角度进行规划建设,涵盖数字资源加工、元数据、唯一标识符、长期保存等方面的标准规范研究与制定^[14],并对我国数字图书馆标准规范发展战略、建设机制进行分析,形成包括我国数字图书馆标准规范总体框架与发展战略、数字图书馆标准规范开放建设机制等七方面,涵盖数十个子项目的庞大规范体系。国家数字图书馆工程从2005年开始依据数字资源生命周期基本框架进行项目标准的设计,规划建设30余项标准规范项目,涵盖数字资源创建、加工、描述、组织、服务、管理及保存等多个环节。中国高等教育文献保障系统于2004年发布《中国高等教育数字图书馆技术标准与规范》,包括CADLIS总体架构和基本技术标准与规范、各子项目专用技术标准与规范、有关产品认证和项目管理等方面相关标准,随后不断修订完善相关标准,形成一系列资源、服务与技术等方面的标准规范^[14]。

3 我国数字图书馆标准规范建设的特点

3.1 积极采用已有研究成果或标准

各数字图书馆项目在建设过程中对一些基础性标准大都采用已有国际标准、国家标准或行业标准等,如《国际标准书目记录》《中国文献编目规则》《普通图书著录规则》《中文机读目录格式》《古籍著录规则》《汉语文古籍机读目录格式》等。标准规范研制成果在数字资源揭示与组织中得到广泛应用,如国家图书馆牵头制定的《中文元数据方案》参考美国国会图书馆、澳大利亚国家图书馆、CEDARS项目、*Dublin Core*、NEDLIB项目等较为成熟的元数据集标准规范。而OAIS作为国际标准被普遍接受为数字资源长期保存系统的基本框架,国家数字图书馆元数据应用方案、《中文元数据方案》都将OAIS作为参考模型。

3.2 结合项目自身需求研制标准规范

我国数字图书馆标准规范建设的一个显著特点是项目拉动,大部分标准规范都是项目成果。在标准规范建设过程中,各工程都是结合项目需求及自身特点进行研制,如国家数字图书馆的“汉字处理规范和专门元数据规范——古文献”子项目,就针对中文和古籍制定系列标

准规范,包括拓片、舆图、甲骨、古籍、家谱五个标准。

3.3 体系化建设特点明显

国内各数字图书馆项目在标准规范研制过程中,大都遵循数字资源生命周期环节进行建设,包括数字内容标准规范、数字对象描述标准规范、数字资源组织管理标准规范、数字资源服务标准规范,即用体系化的思路进行标准规范建设。由张晓林等提出的将数字信息资源建设涉及的标准规范分为内容创建、描述、组织、管理、服务、长期保存和项目建设等^[15],也是从数字资源生命周期的角度进行划分。

3.4 注重合作与共享

标准规范建设是一项共同参与研究、制订与修订的工作,是一个通过逐步完善,最终形成体系的过程。我国数字图书馆标准规范在建设过程中,秉持开放合作的理念,共同研制标准规范。如“我国数字图书馆标准规范建设”项目是由中国科学技术信息研究所、国家图书馆、CALIS管理中心等约20个图书情报机构共同参与建设的。此外,对经过项目实践检验的成熟标准,应依托全国图书馆标准化技术委员会等行业标准化组织将其上升为文化行业标准或国家标准,供国内业界共同参考使用。

4 结语

在我国数字图书馆二十余年发展过程中,研究制定了一系列标准规范,为促进我国数字图书馆建设起到重要作用。虽然数字图书馆标准规范建设成效显著,但仍存在一些问题(如标准规范的体系化程度不够,内容有待丰富);各数字图书馆项目的标准规范差异较大,互使用性较差;标准规范的宣传和使用效果不佳等。近年来随着信息技术的发展,全媒体服务成为数字图书馆服务的发展重点,针对全媒体的标准规范建设将成为数字图书馆标准规范建设的重点任务。

参考文献

- [1] 魏大威.数字图书馆理论与实务[M].北京:国家图书馆出版社,2012.
- [2] 韩永进.中国图书馆事业发展报告(数字图书馆卷)[M].北京:国家图书馆

出版社,2016:352-353.

- [3] 赵悦. 数字资源标准规范建设研究[J].现代情报,2009,29(3):71-73.
- [4] 张晓林,肖珑,孙一刚,等.我国数字图书馆标准与规范的建设框架[J].图书情报工作,2003(4):7-11.
- [5] 数字对象框架[EB/OL].[2014-05-29].<http://www.cadal.cn/bzgf/index.htm>.
- [6] 冯项云,肖珑,廖三三,等.国外常用元数据标准比较研究[J].大学图书馆学报,2001,19(4):15-20.
- [7] 国家数字图书馆工程.标准规范[EB/OL].[2014-07-24].<http://www.nlc.gov.cn/newstgc/gjsztsggc/bzgf/>.
- [8] 知识组织框架[EB/OL].[2014-06-03].<http://www.cadal.cn/bzgf/zszz.htm>.
- [9] 资源集合框架[EB/OL].[2014-08-06]. <http://www.cadal.cn/bzgf/zyjh.htm>.
- [10] 国家数字图书馆工程.图书馆参考咨询服务规范[EB/OL].[2014-08-25].
<http://www.nlc.gov.cn/tbw/tbwbzxd/hybz/hb23.pdf>.
- [11] 张海涛,郑小惠,张成昱.数字图书馆的互操作性研究:Z39.50和OAI协议的比较[J].现代图书情报技术,2003(2):13-15.
- [12] The Open Archival Information System (OAIS)[EB/OL].[2014-08-06].
<http://www.searchtechnologies.com/OAIS-open-archival-information-system.html>.
- [13] 张美芳.数字信息长期保存标准建设的国际化经验及其借鉴[J].数字与缩微影像,2008(3):4-8.
- [14] 赵悦,申晓娟,胡洁,等. 数字图书馆推广工程标准规范体系建设规划与实践[J].国家图书馆学刊,2012(5):46-53,59.
- [15] 张晓林,曾蕾,李广建,等.数字图书馆建设的标准与规范[J].中国图书馆学报,2002(6):6-15.

作者简介

王秀香, 1985年生, 馆员, 研究方向: 图书馆标准化, E-mail:wangxiuxiang@nlc.cn。

Analysis on the Content and Characteristics of Digital Library Standards

WANG XiuXiang

(National Library of China, Beijing 100081, China)

Abstract: The article sorts out the main results of digital library standards from the perspective of standard content, covering the construction and processing standards of digital resources, digital resource description, organization and integration standards, digital library service standards, digital library system interconnection interoperability standards, digital preservation standards and the digital library systemically development standard etc. Then, the author summarizes the main features of digital library standards: learning and using the existing research, combing the demands of the project to develop its own standards, pay attention to the systemic thinking when constructing digital standard content, focusing on cooperation and sharing when develop and use digital library standard.

Keywords: Digital Library; Standard; Guideline

(收稿日期: 2016-08-17)