

科技部科技基础条件平台工作重点项目 研究成果

项目名称：数字图书馆标准与规范建设

子项目名称：我国数字图书馆标准规范发展战略与基本框架

项目编号：2003DEA4T035

研究成果类型：研究报告

成果名称：我国数字图书馆标准规范建设与应用的实施指南

成果编号：CDLS-S01-008

成果版本：总项目组推荐稿

成果日期：2005年7月

撰写人：张晓林（中国科学院文献情报中心）

王惠临（中国科技信息研究所）

牛振东（国家图书馆）

鞠文红（中国科学院文献情报中心）

项目版权声明

本报告研究工作属于科技部科技基础条件平台专项资金项目《我国数字图书馆标准规范建设》的一部分，得到科技部科技基础条件平台专项资金资助，项目编号为 2003DEA4T035。按照有关规定，国家和《我国数字图书馆标准规范建设》课题组拥有本报告的版权，依照《中华人民共和国著作权法》享有著作权。

本报告可以复制、转载、或在电子信息系统上做镜像，但在复制、转载或镜像时须注明真实作者和完整出处，并在明显地方标明“科技部科技基础条件平台专项资金项目《我国数字图书馆标准规范建设》资助，项目编号 2003DEA4T035”的字样。

报告版权人不承担用户在使用本作品内容时可能造成的任何实际或预计的损失。

作者声明

本报告作者谨保证本作品中出现的文字、图片、声音、剪辑和文后参考文献等内容的真实性和可靠性，愿按照《中华人民共和国著作权法》，承担本作品发布过程中的责任和义务。科技部有关管理机构对于本作品内容所引发的版权、署名权的异议、纠纷不承担任何责任。

《我国数字图书馆标准规范建设》课题组网站 (<http://cdls.nstl.gov.cn>) 作为本报告的第一发表单位，并可向其他媒体推荐此作品。在不发生重复授权的前提下，报告撰写人保留将经过修改的项目成果向正式学术媒体直接投稿的权利。

我国数字图书馆标准规范建设与应用的实施指南

目 录

1. 前言	1
2. 数字图书馆标准规范建设的环境与原则	2
2.1 数字图书馆的开放建设环境.....	2
2.2 数字图书馆标准规范的开放建设与应用环境.....	3
2.3 数字图书馆标准规范建设与应用的的基本原则.....	4
2.4 数字图书馆系统建设和应用标准规范的实施要求.....	5
3. 数字内容创建标准规范建设与应用的实施要求	6
3.1 数字内容创建的标准规范结构.....	6
3.2 内容编码标准规范的实施要求.....	7
3.3 数据格式标准规范的实施要求.....	8
3.4 复合对象格式标准规范的实施要求.....	9
3.5 数字对象标识标准规范的实施要求.....	9
4. 数字对象元数据标准规范建设的实施要求	10
4.1 元数据格式的基本建设框架.....	10
4.2 元数据格式标准规范建立与应用的实施要求.....	11
4.3 代码与词汇体系应用的实施要求.....	12
4.4 元数据格式标准规范描述和公开发表的实施要求.....	13
5. 数字信息服务标准规范建设与应用的实施要求	14
5.1 网络服务易用性标准规范的实施要求.....	14
5.2 检索服务标准规范应用的实施要求.....	15
5.3 应用服务标准规范应用的实施要求.....	16
5.4 系统服务描述标准规范应用的实施要求.....	16
6. 组织管理机制的标准规范建设与应用的实施要求	17
6.1 资源集合描述标准规范应用的实施要求.....	18
6.2 资源集合组织机制标准规范应用的实施要求.....	18
6.3 资源集合管理机制标准规范建设的实施要求.....	19
7. 数字资源系统长期保护标准规范建设的实施要求	20
参考文献	21

1. 前言

(1) 根据对数字图书馆及其相关领域标准规范现状和发展趋势的系统研究,在我国数字图书馆标准规范发展战略^[1]和总体框架^[2]的指导下,针对我国数字图书馆标准规范建设与应用的需要,为指导和约束《数字图书馆标准与规范建设》项目各个具体标准规范的建设,为指导和约束各个具体数字图书馆系统的标准规范建设与应用实践,特提出本指南。

(2) 本指南所指的数字图书馆是在公共领域、利用公共资源、提供公共服务的数字信息资源系统。标准规范是数字图书馆系统可广泛使用性、可互操作性和可持久利用性的基础。每个数字图书馆系统都应该在我国数字图书馆标准规范整体发展战略、总体框架和实施指南的指导下,明确建立自己的标准规范应用原则和实施指南。

(3) 本指南根据我国数字图书馆建设的当前需要和国际数字图书馆标准规范的应用现状与发展趋势,按照数字资源生命周期,提出我国数字图书馆标准规范建设与应用的总体原则,并对生命周期各个主要环节的标准规范建设应用提出要求。

(4) 本指南针对数字资源生命周期的各个环节,在阐述该环节标准规范逻辑关系的基础上,对标准规范的建设与应用提出指导性意见。其中,对于根据总体框架的要求必须予以考虑、需要严格一致、而且相应标准规范比较成熟、业界基本形成共识或本项目组已经明确规定的内容,本指南将提出具体的实施要求,这些要求将用“必须”(Must)表示;对于根据总体框架的要求必须予以考虑、需要保持一致、已有基本的原则性共识、但相应标准规范还在进一步发展的内容,本指南将提出原则性的指导意见,这些意见用“应该”(Should)表示;对于根据总体框架的要求应该予以考虑、但相应的标准规范还处于发展过程中的内容,本指南将提出方向性的原则意见,这些意见用“可以”(May)表示。

(5) 《数字图书馆标准与规范建设》项目的各个子项目组,将提供数字资源生命周期各个环节中的标准规范具体应用的技术性指南(以下简称技术指南)。本指南和各个技术指南构成我国数字图书馆标准规范建设与应用的系统化指南。请参照在《数字图书馆标准与规范建设》项目网站(<http://cdls.nstl.gov.cn/>)上随时公布的各个技术指南。

(6) 各个数字图书馆系统应该直接采用本指南和各个技术指南所提出的“必须”要求,或者按照这些要求,根据各自系统情况进行细化或扩展;应该根据本指南提出的“应该”要求和技术指南提出的实施建议,根据各自系统情况,对所涉及的环节或层次制定具体的实施要求;可以根据本指南提出的“可以”要求和技术指南提出的实施选择,根据各自系统情况,对所涉及的环节或层次提出具体

的实施意见或要求。

(7) 本指南分为前言、数字图书馆标准规范建设与应用的环境与原则、数字资源建设中标准规范建设与应用实施要求、数字资源描述中标准规范建设与应用实施要求、数字资源服务中标准规范建设与应用实施要求、数字资源集合组织管理中标准规范建设与应用实施要求、数字资源长期保存中标准规范建设与应用实施要求。本指南当前版本为总项目组征求意见稿(2004年3月31日)。随着数字图书馆技术和标准规范的发展,随着我国数字图书馆标准规范建设的进展,本指南将适时作出修改。

(8) 本指南是在《我国数字图书馆标准规范总体框架与发展战略》子项目组其它研究成果的基础上完成的,并参考了《数字图书馆标准与规范建设》其他子项目组的有关研究成果,特此向有关研究人员在本文撰写过程中提供的帮助表示感谢。

2. 数字图书馆标准规范建设的环境与原则

2.1 数字图书馆的开放建设环境

建设和应用数字图书馆标准规范,需要了解数字图书馆的开放建设环境和动态发展趋势,把握对数字图书馆系统的可广泛使用性、可互操作性和可持久利用性的要求。

(1) 数字图书馆系统是开放的数字信息系统。它自己建设的资源与服务必须适应多样化、分布、甚至是动态的用户需求,必须能够被用户在复杂的网络环境下方便地利用;它往往需要集成别的(往往是)分布、异构和自主的资源和服务来为自己的用户提供服务;它还需要与数字教育、数字科研、电子政务、电子商务、数字传媒等领域的资源与服务进行有机整合来支持用户的集成应用。

(2) 从整个信息环境的角度和用户信息利用的角度,任何一个数字图书馆系统实际上是更大范围的数字图书馆体系和更大范围的数字信息体系中的一个模块。它可以作为独立的系统支持特定的资源或服务,但也应该能被方便集成到别的系统中成为其中的有机部分,或者方便地集成别的资源与服务,从而支持用户根据自己需要来动态构造所需要的服务。

(3) 而且,随着信息技术和用户需要的不断发展,数字图书馆体系结构不断变化(成为所谓不确定的“移动目标”),使得人们必须通过一种开放的机制,灵活容纳现在和未来的技术、资源和服务,健壮地支持它们的动态集成与定制,有效保障整个数字图书馆体系和各个具体数字图书馆系统的可伸缩性、可扩展性和可持续性。

(4) 因此,从保障对用户的服务效率的角度,从保障所建设的资源与服务

的可广泛使用性、可互操作性和可持久利用性的角度，任何数字图书馆系统都必须将开放性和互操作性作为自己系统设计、应用建设和服务管理的基本原则。实际上，这已经不是技术选择的自由，而是对用户和历史的责任。

2.2 数字图书馆标准规范的开放建设与应用环境

标准规范是数字图书馆系统可广泛使用性、可互操作性和可持久利用性的基础。但是，在复杂的信息环境和迅速发展的信息技术面前，标准规范也面临开放建设的要求。

(1) 数字图书馆本身涉及了多个相互支撑的技术层次，例如计算机、通信、网络服务、信息资源与内容、应用服务、管理等，涉及了多个相互交叉的应用领域，例如出版发行、图书情报、教育、科研、商务、政务、媒体与娱乐等。这些领域和层次都在建设或应用自己的标准规范。即使在所谓“单纯”的数字图书馆领域，也有许多应用领域或应用系统，根据不同的资源类型、服务功能、技术机制或历史过程，已经（并将继续）制定和应用多个涉及数字资源与服务的标准规范。

(2) 在上述标准规范空间，有些层次由于其在技术结构中对于其他层次的支撑地位（称之为技术支撑领域，例如通信协议对于网络服务），有些领域由于其在资源生产和业务应用的优势地位（称之为优势业务领域，例如出版发行对于内容对象、远程教育对于课件等），使得它们的标准规范成为相应技术层次、资源或业务的主流标准规范。辨别、支持和应用这些主流标准规范，是促进开放性和互操作性的有效做法。但同时，也有许多应用涉及众多的领域、资源和业务，例如描述内容对象的元数据标准、对内容对象进行检索和搜索的规范协议等，因此在跨领域、跨资源和跨系统应用上存在标准规范的互操作问题。而且，由于新技术和应用不断出现，还会要求或产生新的标准规范，还存在标准规范的迁移与兼容问题。

(3) 为了理解标准规范的复杂关系空间对标准规范互操作的要求，需要构建标准规范的框架体系来描述不同层次或方面的标准规范之间的关系以及对互操作的要求，从而对整体标准规范体系和具体标准规范的建设与应用界定明确的要求。本项目基于数字资源生命周期，并参考了信息系统互操作框架结构，按照数字资源的产生与建立、内容对象的描述、数字资源的集合组织、数字资源的利用服务、数字资源系统及其服务的管理、数字资源长期保存的流程，建立了基本的数字图书馆标准规范总体框架^[2]，由此反映数字图书馆领域标准规范本身的复杂关系空间。

(4) 在标准规范的复杂关系空间中，仅仅建立和坚持某种标准规范是远远不够的。标准规范的生命力在于其本身的开放性，从而支持广泛的应用；在于其

本身的可互操作性，从而支持与（使用不同标准规范的）其他领域和其他系统的有机互操作；在于其本身的可扩展性，从而支持在技术、应用和互操作机制动态发展情况下的兼容与发展。因此，标准规范的建设与应用，也需要通过开放性要求（可互操作性和可扩展性可以看成开放性的有机部分）来保障标准规范本身的可广泛使用性、可互操作性和可持久利用性。

2.3 数字图书馆标准规范建设与应用的基本原则

标准规范的开放性，一方面通过有关技术机制予以支持和保障，另一方面通过开放的建设与应用机制予以支持和保障。这种开放的建设与应用机制的基本原则是：

（1）开放性原则。应该坚持或优先考虑直接采用那些被广泛接受和应用、得到有效标准化过程支持和市场支持的标准规范，避免重新编制已经存在的标准规范，避免与已经存在并广泛使用的标准规范形成冲突。应该坚持采用或者优先采用开放标准，即那些知识产权明确属于公共领域、可被公共获取、采用开放语言和标准格式描述、有可靠的公共登记和持续维护机制、有可靠的开放转换和开放扩展机制、公开发布详细技术文件的标准规范。而且，应该优先采用那些得到开放源代码软件支持、采用开放建设过程的标准规范。

（2）国际化原则。应该遵循国际标准优先的原则，尽量直接采用国际标准、国际范围通用的行业标准或该领域（行业）先进国家的国家标准（以下用“国际标准”代表），并根据本地需要制定明确的技术性应用指南。在国际标准不能有效满足本地需要的情况下，应该遵循国际性的规范、在国际标准的逻辑框架基础上进行定制和扩展本地标准，从而保持与国际标准的可互操作性。无论哪种情况，都应该持续跟踪国际标准的发展，及时更新应用指南或本地标准。同时，对于具有明显的本地特色和应用先进性、具有广泛的国际化应用前景的本地标准，应该积极推进它们的国际化。

（3）互操作原则。应该在标准规范的设计中优先考虑和充分支持其互操作能力，而不是如何满足一个具体应用或具体系统的内部功能要求。而且，这种互操作能力应该支持互操作层次体系中多层次间的互操作，应该支持多个不同应用领域或资源领域之间的互操作，应该支持应用流程中多个逻辑模块间的互操作，应该支持已有或未来的多个标准规范之间的互操作。具体应用或具体系统的特殊功能要求，可以通过在标准规范基础上，按照规范方式进行定制或扩展来支持。

（4）互支撑原则。数字图书馆作为网络环境的应用系统，必须直接利用技术支撑领域的标准规范。而且，数字图书馆处于多种资源、应用和技术交叉交互的环境中，应该积极利用优势业务领域的标准规范（优势标准规范）；如果优势标准规范不能有效满足本地需要，应该在它们的基础上，按照规范方式来定制和

扩展成本地标准，从而保持与优势标准规范的互操作性。有些领域并不构成优势业务领域，但其资源和应用与数字图书馆的资源和应用发生广泛的交叉（称之为合作领域或交叉领域），应该主动促进、积极组织与这些领域之间的标准规范合作建设，积极促进数字图书馆标准规范的跨领域应用。

（5）模块化原则。每个标准规范都是数字图书馆标准规范体系中的一个逻辑模块，各自主要满足整体体系中的一定纵向层次或横向环节的功能或应用上的互操作要求，不能够也不试图去通过自身的“完善”或“扩充”来解决属于整体或属于其他模块的问题；但同时，它通过对其他标准规范的开放应用来支持更大范围或更复杂条件下的互操作，包括采用共享底层标准规范、嵌入相关标准规范、复用其他标准规范、嵌入到其他标准规范、根据相关标准规范进行定制或扩展等。模块化可以使每个标准规范具备即插即用（Plug-and-play）的能力、增加不同标准规范之间的互操作性、提高每个标准规范和整个标准规范体系的可扩展性。因此，每个标准规范应该清楚描述自己在整体体系中的定位和与其他标准规范的关系，应该明确规定自己对其他标准规范（包括其他领域的标准规范）的开放应用机制，充分支持层次型模块化（例如 TCP/IP 结构），充分支持核心型模块化（例如在核心集基础上扩展专门集和本地应用集），支持模块间和模块与系统间的一致性和互操作。

（6）完整的应用工具流原则。标准规范需要一系列应用工具（例如定义与标记语言、应用规则、词表与代码表、最佳实践指南、处理模板、转换方式、应用接口等）的支持，而为保证标准规范的开放性，建设与应用标准规范时应该该系统描述和建设所需要的应用工具体系，同时按照开放性要求建设这些应用工具，支持开放和规范的标准规范应用 workflow。

2.4 数字图书馆系统建设和应用标准规范的实施要求

每个数字图书馆系统应该在我国数字图书馆标准规范建设的发展战略、总体框架和实施指南的指导下做到：

（1）必须坚持开放地进行标准规范建设与应用。应该广泛吸纳各种利益相关者参加自己的标准规范建设或应用决策。应该建立标准规范的公开建设或应用程序，公开建设计划和要求标准，公开征集标准规范草案，公开对所征集的标准规范草案的审查、修订程序和规则，公布标准规范的编制过程和各个阶段性成果。应该公布自己的标准规范建设或应用情况，吸引其他单位参加自己的标准规范建设，建立可靠的公开的意见反馈机制，建立开放试验机制。应该坚持标准规范作为公共领域的开放标准规范，公开发布利用标准规范所需要的详细技术支持文件，支持通过开放源代码方式建立标准规范的各种应用工具，在公共登记系统中登记标准规范、技术支持文件和应用工具。应该积极支持第三方复用、定制或扩

展标准规范。

(2) 建立标准规范应用的总体战略和总体应用指南。应该根据自己系统的需要,按照可广泛使用性、可互操作性和可持久利用性的要求,明确规定自己应用标准规范的战略、原则和管理机制。应该按照数字资源生命周期,明确规定自己对于各层次和环节标准规范的实施要求,明确规定自己系统对这些标准规范进行定制或扩展的原则和实施要求。

(3) 建立具体标准规范的应用指南。应该根据具体标准规范和我国数字图书馆标准规范项目组制定的应用指南,明确规定自己对于每个标准规范的应用规则,明确规定对于每个标准规范所涉及的应用工具体系(例如定义与标记语言、词表、代码表、转换语言、应用接口等)的实施要求和使用规则。

(4) 建立可靠的标准规范开放应用机制。应该建立专门的公共渠道(例如在主页明显位置、专门网页、甚至专门网站)来发布自己系统应用标准规范的战略、原则、管理机制、总体应用指南、具体应用指南、以及在各个层次和环节所建立或应用的实际标准规范。应该将所公布的文件以规范描述文件形式置于自己系统公共服务器根目录下的公知位置,供第三方系统自动搜索和调用;应该按照规定的格式在公共登记系统登记所公布的文件,支持第三方公开搜寻和调用。应该及时更新所公布的标准规范文件,保持可靠的版本信息。

(5) 建立规范的标准规范开放描述机制。应该采用开放定义语言和规范形式定义和描述具体的标准规范,应该采用开放标记语言、链接与指向机制、转换语言等来定义和标记自己对标准规范的定制或扩展,应该采用规范格式来描述相应的政策文件。应该积极支持第三方系统基于开放语言来识别和解析自己的标准规范及其应用工具。

(6) 建立完善的长期应用管理机制。应该建立和公布明确的管理机制,包括各种政策文件、标准规范、应用指南、应用工具等的版本信息、登记制度、维护责任、意见反馈渠道、修订历史等,形成可靠的应用管理链。

3. 数字内容创建标准规范建设与应用的实施要求

3.1 数字内容创建的标准规范结构

(1) 在数字资源建设中,数字内容包括由传统载体(印本、图片、音频、视频等)经过数字化加工形成的数字对象,或者是原生数字形态的对象(Born-digital resources,例如直接的数字文本、图像或录音文件等),由数字对象派生或衍生的数字对象。这些数字对象可以是数字图书馆系统自行建设的或者是由第三方资源拥有者提供的资源。

(2) 数字内容创建过程涉及多个环节。在宏观上,它涉及:a. 对需要数字

化处理的内容（可能本身已经是数字对象）的选择、收集（采集）和整理，b. 对这些内容进行数字化处理或采集，c. 对数字化所生成的数字对象的组织。本指南主要涉及第二个环节，即对内容进行数字化编码（内容编码），将经过编码的内容按照一定格式组成数字对象（对象格式），对数字对象进行标识（对象标识），根据应用需要对多个内容对象及其具体内容进行组合形成复合对象（复合对象格式），以及对数字化编码、格式化、标识和组合的技术处理过程（数字化加工过程）。括号中表示的内容编码、对象格式、对象标识、复合对象格式、数字化加工过程，构成了本指南所指的数字内容创建过程中主要的标准规范要求。

（3）在某些情况下，一个环节被结合在另一环节中（例如在图像数字化过程中，内容编码和格式处理被结合在扫描过程中），一个环节与另一个环节被分割在不同的应用系统中（例如数字对象的复合组织可能由另一个“完全不相关的”应用系统来完成）。但是，数字内容创建过程的各个标准规范相互关联和共同地决定着数字内容的可广泛使用性、可互操作性和可持久利用性，因此一个数字图书馆系统，即使是完全使用第三方提供的数字信息资源，也必须充分把握数字内容创建的全过程、把握这个过程中所涉及的标准规范类别，从而清醒地和关联地分析、选择和规定数字内容创建的标准规范。

3.2 内容编码标准规范的实施要求

内容编码是制约数字信息可使用性乃至可持续性的最基本条件。内容编码主要涉及文字编码、特殊信息编码、数字文献结构编码（当然，从形成结果角度，也可以将数字文献结构编码归到格式标准）。音频视频信息的内容编码主要通过相关的格式标准来体现。

（1）数字图书馆系统必须对数字内容创建中涉及的内容编码标准予以明确的规定。

（2）数字图书馆系统在文字内容编码标准上必须遵循国际和国家标准。在特殊信息编码和文献结构编码上，采用顺序应该是国际标准、中国国家标准、国际范围通用的行业标准或该领域（行业）先进国家的国家标准、该领域（行业）的主要标准化团体的推荐标准。

（3）对于文字编码标准，必须支持 ISO 10646/Unicode UTF-8 标准，同时应在应用层应该支持 GB2312-1980 和 GB18030-2000，应该兼容 Big5 和 CCCII 标准。

（4）对于数学符号和公式、化学符号、矢量信息、地理参数等特殊信息编码，应该采用支撑技术领域的国际或国家标准，在没有国际或国家标准时应该采用支撑技术领域或优势业务领域的通用标准。应该采用基于XML^[3]的编码标准，在没有国际或国家标准的情况下，应该优先采用W3C^[4]和OASIS^[5]的有关标准，例如矢量数据可以遵循SVG^[6]，多媒体信息可以遵循SMIL^[7]，语音信息可以遵循

SSML^[8]，数学数据可以遵循MathML^[9]，地理信息可以遵循GML^[10]，化学信息可以遵循CML^[11]。

(5) 对于数字文献结构编码，应该采用基于XML的编码标准，应该采用支撑技术领域或优势业务领域的文献结构编码。除非特殊情况，应该采用XML DTD^[12]或XML Schema^[13]来定义文献结构，并尽量采用XML Schema方式，而且定义应经过XML语法验证。

3.3 数据格式标准规范的实施要求

数据格式涉及文本、图像、音频、视频、多媒体等数据内容，需要根据资源建设、使用、长期保存等方面的需要建立格式体系，并根据对不同格式的要求确定格式标准。

(1) 数字内容创建中需要多种数据格式来承担不同责任，例如保存格式(Preservation / Archiving Formats)，有时又称原版格式，要求在数字化状态下真实地保存原始数据形式(例如图像、录音、录象等)的内容及其表现，要求采取非压缩或无损压缩格式；服务格式(Service / Access / Viewing / Reference Formats)，作为正常存储和使用的格式，要求保证数据内容的使用质量、降低传输成本，可采用适当压缩比的压缩格式，可从保存格式中派生；预览格式(Previewing Formats / Sampling Formats)，作为预览或预演信息，提供粗略的内容表现，可采用大压缩比的格式，可从保存格式或服务格式中派生。

(2) 数字图书馆系统必须针对自己的资源范围和应用对数字内容的要求，明确描述自己的数据格式体系，明确规定数字内容创建涉及的数据格式标准。

(3) 数字图书馆系统必须遵循支撑技术领域基本数据格式的国际和国家标准，应该采用支撑技术领域或优势业务领域的通用格式标准，应该采用开放标准。

(4) 对于文本数据格式，在文本数据作为文本文件时，应该采用HTML、XHTML或XML格式；XML格式定义必须遵循XML DTD或XML Schema方式，尽量采用XML Schema方式。在不能有效处理HTML/XML的环境下，应该采用纯文本格式。如果文本数据本身是专门格式文本，必须提供将这些格式转换为HTML/XML文本或纯文本的公开方法。在文本数据作为图像文件时，应该采用TIFF^[14]、JPEG^[15]或PDF^[16]格式。

(5) 对于图像数据格式，应根据保存、服务或预览的不同要求采用不同格式标准。保存格式应该采用非压缩或无损压缩的TIFF格式；服务格式可以采用JPEG格式或压缩TIFF格式；预览格式可以采用GIF^[17]或PNG^[18]格式；线图图像可以采用JPEG、PNG或GIF格式。

(6) 对于视频和音频数据格式，应该优先选择开放标准，也可以采用在国际范围和应用领域广泛接受的通用格式。对于视频数据的保存格式，应该优先采

用数字录像格式，也可以采用MPEG格式^[19]；对于视频数据的服务格式，应该采用MPEG格式，也可以采用通用的格式，例如Apple Quicktime^[20]和MS Real Video^[21]。对于音频数据的保存格式，应该优先采用数字录音格式，也可以采用MP3^[22]格式；对于音频数据的服务格式，应该采用MP3格式，也可以使用通用的WAV^[23]、Apple Quicktime^[24]和MS Real Audio^[25]格式。

3.4 复合对象格式标准规范的实施要求

复合对象是由一系列底层数字对象（子对象）按照一定结构组合而成、并能够对各层子对象进行操作的数字对象，其中子对象的数据类型可以是文字、图像、音频、视频、二进制程序等以及由它们组成的多媒体对象。复合对象可以是将有子对象物理地组织在一起、封装在一个文件之中的实际对象，也可以是通过链接参照机制将分布在不同位置的多个底层对象组织在一起的虚拟对象，还可以是可调整封装或链接对象的动态对象。

(1) 在采用复合对象时，必须针对特定的应用需要，明确建立复合对象的层次化逻辑结构，明确定义和描述复合对象内的子对象类型、关系、结构、结构与关系的选择组织机制、子对象及其结构和关系的描述机制等。

(2) 应该采用优势业务领域的通用的复合对象格式标准，在没有优势业务领域（或图书情报领域本身是优势业务领域）时尽量兼容密切相关领域的复合对象格式标准。应该采用开放标准。

(3) 应该建立或选择开放的定义语言来定义和描述复合数字对象的内容与结构，而且这个定义语言应该是基于XML语言。

(4) 复合对象定义语言应该支持任意类型的子对象，应该同时支持内嵌对象和外链对象，应该支持将另一复合对象作为子对象，应该支持对复合对象和各层子对象的元数据描述。

3.5 数字对象标识标准规范的实施要求

数字对象唯一标识符是在特定命名域（特定的数字对象类别和范围）内，按照统一命名规则和命名程序授予数字对象的唯一和永久的名字。数字图书馆系统通过一个命名登记机构存储命名登记数据，通过一定的地址解析机制将数字对象的唯一标识符转换成相应的物理存放地址，支持基于唯一标识符的跨系统调用数字对象。

(1) 必须采用唯一标识符机制来标识自己的数字对象。

(1) 必须明确规定自己系统所建立或采用的数字对象唯一标识符命名体系及其命名标识规则，命名规则必须保证唯一标识符本身是逻辑的、与物理地址无

关、永久性的。如果需要标识的数字对象是复合对象，应该分别对子对象和复合对象进行唯一标识。

(2) 数字对象唯一标识符应该优先采用URI或Info命名域^[26]中的标识符命名体系，可以采用相应应用领域的通用标识符命名体系。在存在国际标准和开放标准的情况下应该采用国际标准和开放标准，还可以积极参照信息资源富集国家的国家标准。在我国的数字对象唯一标识符标准建立后应该采用该标准。

(3) 如果选用非公共的唯一标识符命名体系，必须建立相应的标识符解析机制，保证标识符体系本身的可解析性，保证解析机制的正常运转。必须公布自己标识符命名体系的命名规则、相应的解析机制和解析接口，支持第三方系统对这些标识符的解析。

(4) 无论什么情况，必须公布自己系统的唯一标识符命名体系。在全国的唯一标识符命名体系登记系统建立后，必须登记自己系统的唯一标识符命名体系。

(5) 在跨领域、跨对象类型和跨应用系统情况下，应该支持多个唯一标识符命名体系的应用和互操作。

(6) 在全国的唯一标识符解析机制建立起来后，应该按照规范形式和程序将自己的内容对象唯一标识符/描述元数据值对 (Value-pair) 提交公共的标识符登记解析系统，支持唯一标识符的可解析性和跨系统调度。

(7) 还可以建立数字对象的认证机制，例如数字签名或数字水印。认证机制可以在公共登记系统予以注册，可以支持对数字对象的公共验证。

4. 数字对象元数据标准规范建设的实施要求

4.1 元数据格式的基本建设框架

(1) 按照开放环境下数字资源服务的互操作需要，我国数字图书馆领域元数据建设将建立一个基本元数据格式 (Core Metadata Format)，作为所有数字对象和系统的公共的核心元数据集，支持对数字对象的检索、确认和调度；在基本元数据格式的基础上，可根据各个领域的特殊需要，按照标准方法进行扩展，形成各个领域的专门元数据格式 (Special Metadata Formats)；在基本元数据或专门元数据格式基础上，还可按照各个具体应用系统的个性化特点，按照标准方法再进行扩展，形成以应用协议表示的具体应用系统的应用元数据格式 (Applied Metadata Formats)。

(2) 对于资源性质、使用要求、管理与保存要求比较简单的数字图书馆系统，可以采用基本元数据建立自己系统的元数据，方便地与其他系统进行互操作。对于涉及专门资源类型或特殊使用与管理要求的系统，可以采用合适的专门元数

据格式，或者利用规范机制来定制、扩展基本元数据或专门元数据格式，形成自己的应用元数据格式。

(3) 由于基本元数据在语义和语法表示上一致，各个系统可以通过解析、识别和输出共同的基本元数据元素支持跨系统的数字对象检索、确认和调度；即使对方系统使用的是基于基本元数据进行扩展生成的专门元数据或应用元数据，也可以通过“Dumb down”方式舍去（横向扩展的）或归并（纵向扩展的）元素来保证最基本元素的传递和识别。

(4) 由于专门元数据格式、应用元数据格式都是在基本元数据基础上通过标准方法进行扩展，语义定义兼容（重叠和交叉情况较容易避免），语法表示一致，因此各个系统之间根据标准的扩展规则和语句语法规则，比较容易构建准确的元数据转换关系，比较容易构建元数据结构解析、语义解释和映射转换系统。

(5) 各个应用系统可以根据自己的需要和能力选择使用合适的元数据格式，并可比较方便地“向下”扩展或“向上”折叠元数据。而且，如果要求在建立专门元数据格式或应用元数据格式时复用各个领域已有元数据格式中的合适元素，也有利于充分利用现有元数据格式和元数据资源，有利于与现有系统（Legacy systems）的元数据进行互操作。所以，这种策略开发建设成本低，应用健壮性好，可伸缩和可扩展性好。

4.2 元数据格式标准规范建立与应用的实施要求

(1) 应该按照开放元数据机制来分析元数据设计需求，对所涉及的应用领域或应用过程的对象实体及其关系进行逻辑分析，划分针对不同对象和不同功能的元数据内容，建立面向应用的元数据逻辑框架。这个逻辑框架应该用 UML 描述。在有条件的应用领域，应该建立元数据词典，作为公共工具支持互操作。

(2) 应该采用已有的国际标准、国家标准、或者优势业务领域的行业标准或通用标准，避免从头创建自己的元数据格式。

(3) 在我国数字图书馆领域，采用简单Dublin Core^[27]作为基本元数据。对于一般数字对象或一般数字资源，应该采用Dublin Core作为描述元数据。

(4) 对于专门数字对象或特殊资源，应该建立基于基本元数据格式的专门元数据格式。我国数字图书馆标准与规范建设项目将根据上述原则建立相应的专门元数据格式（简称专门元数据格式），定义和描述有关的专门数字对象或特殊资源。这些专门元数据格式将建立与基本元数据格式的标准转换规则，将建立与本领域其他标准或通用的元数据格式的转换规则。各个数字图书馆系统应该采用这些专门元数据格式。

(5) 对于专门数字对象或特殊资源，如果数字图书馆系统需要根据特殊的本地化要求进行进一步描述，可以采用两种方式。一是以基本元数据格式为基础，

复用相关的专门元数据格式中的合适元素和相关领域已有通用格式的合适元素作为扩展元素，并可以自定义合适的元素作为本地扩展元素，构成符合本地需要的应用元数据格式。二是以相应的专门元数据格式为基础，复用相关领域已有通用格式的合适元素作为扩展元素，并可以自定义合适的元素作为本地扩展元素，构成符合本地需要的应用元数据格式。应用元数据格式必须建立与基本元数据格式和相应专门元数据格式的转换规则。

(6) 各个数字图书馆系统在描述专门数字对象或特殊资源时，可以采用该专门对象或特殊资源或特殊应用领域内的国际通用元数据格式，作为本地元数据格式。此时，必须建立本地元数据格式与基本元数据格式和相应专门元数据格式的转换规则。在确定与基本元数据的转换规则时，应该采用本地元数据格式所依据的国际格式制定单位所定义的与 Dublin Core 元数据的转换关系；在制定单位没有定义转换关系时，可以采用其他主要数字图书馆系统所定义的转换关系。在确定与专门元数据格式的转换规则时，应该遵循专门元数据格式在建立自己的相关转换规则时所阐明的原则和方法。

(7) 对于自己建立的各种元数据格式或应用规范，必须建立可靠的公开的版本管理制度，尽责维护相关格式定义、转换关系定义、转换模板定义和其他相关信息，有效跟踪所依据的标准或规则的版本变化。各种元数据格式，作为独立命名域或应用协议，应该在元数据规范登记系统中登记和发布。相应的数据元字典、内容描述体系、编制指南、编制模板等也在元数据规范登记系统中登记和发布。

4.3 代码与词汇体系应用的实施要求

代码与词汇体系（以下简称代码体系）指在具体描述元数据元素内容时，为保证内容语义的唯一性、明确性和一致性，所规定的代码体系（Encoding Schemes），包括（但不局限于）关于语种、国别、资源类型、文献类型、时间与日期、地理名称与位置、人名和机构名、主题概念、分类类目等方面的体系。

(1) 数字图书馆系统必须通过规范的元数据格式定义或编制指南，明确规定自己使用的所有元数据格式中的所有适用元素所采用的代码体系。这些规定应该是确定性规定，明确要求使用指定代码体系。在特殊情况下，也可以是选择性规定，提出根据情况选择使用代码体系的原则和方法。

(2) 有对应的国际和国家代码体系标准的元素，应该采用国际或国家标准；没有国际或国家标准时，应该采用优势业务领域或对应学科领域通用的代码体系；没有通用代码体系时，应该采用权威元数据编制机构或通用元数据格式所推荐的代码体系。

(3) 特殊情况下，如果需要编制自己的代码体系，必须明确定义和公开公

布自己的代码体系、明确描述代码及代码关系的语义定义、避免歧义，必须避免与国际、国家或行业标准规范中的代码相混淆。应该明确建立和公开公布这些代码体系与标准（或通用）代码体系的转换关系。

（4）采用我国基本元数据格式和专门元数据格式的系统，应该支持这些格式编制规范所规定的代码体系。

（5）元数据格式应该支持对代码体系的定制和扩展，在描述内容代码时应该通过三元组方式支持对代码体系名称、代码体系 URI、按照该代码体系所选择的代码的组合描述。

4.4 元数据格式标准规范描述和公开发布的实施要求

（1）在构建元数据格式时，必须按照ISO11179-3^[28]的要求定义元数据元素，必须按照规范格式建立元数据格式的文本定义文件，必须提供元数据元素的语义定义，尽量提供必要的规范化定义工具，例如所依据的数据元词典、语义网络和知识本体定义。必须建立针对元数据格式的编制指南。

（2）在建立专门元数据格式和应用元数据格式时，应该采用规范的定义机制、形式表示机制和应用协议机制（Application Profiles）。对于复用的元数据元素，必须明确说明其来源命名域、出现限制（例如是否必备、出现次数、限制使用条件等）、内容描述方式（包括代码或词汇体系、取值范围、标记语言等）、相互依赖关系等。有些情况下还可对复用元素的语义进行限定。应用协议必须采用严格的语义定义和使用规则，保障通过复用和扩展所形成的应用格式的描述清晰性、逻辑整体性、语义一致性和操作规范性。

（3）元数据格式必须建立基于XML Schema的定义文件、并经过XML语法验证。可以采用RDF^[29]、RDFS^[30]等开放内容描述语言进行描述。必须采用XML Namespace^[31]和URI等标准机制对被复用的命名域进行链接。

（4）所有转换规则必须通过明确的转换关系表予以定义和描述，应该采用基于XSLT^[32]的转换模板来描述转换关系。应该采用元数据格式制定单位所定义的转换模板，在制定单位没有定义转换模板时可以采用其他主要数字图书馆系统所定义的转换模板。采用第三方系统所定义的转换关系和转换模板时，必须提供可靠的链接和验证，保证转换的有效性和转换工具的可调用性。

（5）数字图书馆系统必须公开发布和可靠维护自己的元数据格式及其应用工具。应该建立自己的元数据应用规范，明确描述自己所建立或采用的元数据格式、代码体系、编制指南、编制范例、转换规则、转换模板定义等。如果在元数据应用规范中涉及所采用或引用的标准规范及其 URI，应尽责维护引用信息（包括 URI）的可靠性，保障它们的可查询和可获取。应该通过本系统公开和方便的渠道公布该规范，并随着所采用的标准规范及其应用工具的变化而及时更新。

(5) 如果自己建立专门的元数据格式或应用协议，必须在本系统公布相应的定义文件和相关应用工具的定义文件，最大限度地给出数据用例，并且尽责维护这些文件。在全国元数据规范登记系统建立起来后，必须按照规定元数据格式登记和发布自己的元数据规范信息，并随着自己资源与服务的变化而及时更新自己的登记信息。

5. 数字信息服务标准规范建设与应用的实施要求

数字图书馆系统必须保证从用户和第三方系统角度对自己信息服务的易用性和可互操作性，包括：基本网络服务的易用性，支持在复杂的网络环境下对所提供信息服务的方便获取、理解和利用；数据检索服务的可互操作性，支持在异构系统之间对信息内容进行检索或搜索；应用服务的可互操作性，支持在异构系统之间根据应用服务流程交换数据和控制；系统服务的可互操作性，支持在开放环境中对系统服务进行开放和公共的定义、登记、发现和调用。数字图书馆系统通过对基本技术要求、操作协议和管理机制的定义、描述和登记，支持各个层次的易用性和可互操作性，同时并不排斥任何系统在其本地系统机制内采用自己的方法支持自己的客户端系统。

5.1 网络服务易用性标准规范的实施要求

数字图书馆系统必须遵循W3C的Web Accessibility Guidelines^[33]中的易用性原则和分析要求，保障各种复杂条件下的用户都能方便地获取系统服务。

(1) 通过 Web 方式提供信息服务的系统，必须支持 WAI 内容易用性指南中优先级 1 和优先级 2 所规定的易用性标准，应该支持 WAI 内容易用性指南中优先级 3 所规定的易用性标准。应该支持 WAI 创作工具和用户代理易用性指南中规定的易用性标准。

(2) 必须支持基于 HTTP 协议的通用浏览器读取，如果某个系统需要使用其它通信协议，也必须提供与 HTTP 协议的接口。必须支持基于 HTML/XHTML/XML 的信息内容格式，如果某个系统需要使用其它信息内容格式，也必须提供与上述通用格式的转换接口。

(3) 应该通过通用浏览器提供全部服务，而且应该保证信息内容能够被所有的浏览器识别和呈现，能够在用户视觉和听力存在障碍时仍能被有效理解和利用，能够在不依赖特定硬件或脚本情况下仍能被有效理解和利用。

(4) 如果系统服务需要使用额外插件，应该尽量选择共享软件或开放源代码软件，而且必须作为系统实现和服务提供的一部分提供这些插件，或者提供能够方便获取相应插件的可靠链接或登记系统。服务提供者应该保证没有这些插件的用户得到必要的提示，并仍然能使用不直接受插件控制的其它服务。

(5) 如果所需要的插件是服务提供者的专用软件, 提供者必须按照规范的软件模块描述格式对插件进行描述, 并在公共服务登记系统上公开登记, 供第三方开放搜寻。提供者必须作为资源服务的一部分提供第三方安全复本, 供数字图书馆系统在提供者无法继续提供插件服务时向自己用户提供公共的插件服务。

(6) 应该采用开放标准来处理网络环境中的数据传递。元数据和数字对象在传输过程中应该支持基于HTML的内容封装方式, 应该支持基于XML的内容封装方式。在传输内容为元数据情况下, 应该采用相应元数据格式的标准或通用的封装方式, 例如DC-XML^[34]、MARCXML^[35]、MODS^[36]等; 如果自定义封装格式, 应该以XML Schema形式定义, 并予以公布。在传输内容为数字对象情况下, 应该采用标准或通用的封装方式, 例如METS^[37]、IMS Content Packaging^[38]、MPEG-21 DIDL^[39]; 如果自定义封装格式, 应该以XML Schema形式定义, 并予以公布。应该支持基于SOAP^[40]形式对上述元数据包或数字对象包进行传输封装的方式, 并与HTTP协议绑定。

(8) 对于特殊数据内容(例如GIS数据、计算数据、统计数据、虚拟现实数据等), 应该支持基于XML语言的定义、描述和封装方式, 例如SVG、SMIL、SSML、VRML/X3D^[41]等, 从而支持对这些数据的不依赖任何特殊硬件软件的应用。

5.2 检索服务标准规范应用的实施要求

数字图书馆系统的检索服务向用户提供最基本的数字对象查询和调用服务。检索服务一方面通过规范的检索界面支持用户的直接检索, 另一方面通过检索协议接口支持用户通过第三方系统的检索。

(1) 检索服务必须遵循网络服务易用性的实施要求。

(2) 检索服务必须支持基于HTTP的检索方式, 必须提供基于HTML的检索界面, 应该提供可被嵌入到第三方系统的基于HTML的检索接口, 支持用户和第三方系统将检索接口嵌入自己的系统之中。

(3) 检索服务应该在自己的系统信息中明确定义所采用的数据格式、标引语言、检索方法等, 定义文件应该置于本系统公知位置, 并在全中国信息服务登记系统公开登记。如果这些信息中涉及有关标准规范, 应该提供与标准规范文件的可靠链接。

(4) 对于面向元数据的检索服务, 应该提供基于Z39.50协议^[42]的检索接口; 在建立Z39.50协议客户端时, 必须确定所采用的Z39.50应用规范, 说明本系统在使用Z39.50协议时遵守的具体功能、参数和语义定义; Z39.50应用规范应该遵循已有的通用规范, 例如Bath Profile^[43]或One Profile^[44]; 应该逐步支持Z39.50/SRW/SRU^[45]协议。

(5) 对于面向元数据的搜索服务，应该采用OAI元数据采集协议^[47]来提供元数据开放搜寻服务；提供OAI数据服务的系统（数据提供者）必须提供Dublin Core形式的元数据记录，提供OAIMHP接口界面，支持第三方服务提供者对自己元数据的搜寻。如果数据提供者提供其他格式的元数据记录，应该按照OAI的规定链接和公布相应的元数据格式。

5.3 应用服务标准规范应用的实施要求

对于现阶段的数字图书馆应用服务，除检索服务外，还包括链接服务、信息推送服务、馆际互借与文献传递、参考咨询、个性化定制、集成定制等面向用户或面向第三方系统的服务。这些服务一般会依据一定协议、通过一定界面形式、有些还借助一定的客户端向用户或第三方系统提供。

(1) 各类应用服务必须遵循网络服务易用性的要求。

(2) 各类应用服务必须支持基于 HTTP 的调用，必须提供基于 HTML 的使用界面，应该提供可被嵌入到第三方系统的基于 HTML 的接口，应该采用本应用服务领域的国际标准或通用标准。

(3) 各种应用服务应该在自己的系统信息中明确定义所采用的数据格式、流程规范、应用接口等，定义文件应该置于本系统公知位置，并在全国信息服务登记系统进行公开登记。如果这些信息中涉及有关标准规范，应该提供与标准规范文件的可靠链接。

(4) 对于面向分布资源的链接服务，应该遵循OpenURL模式建立开放链接机制。当前，应该遵循OpenURL 0.1 版^[48]；在ANSI/NISO Z39.88^[49]成为正式标准后，应该遵循该标准。作为链接源和链接目标时，应该详细说明支持第三方链接服务器的技术规则，并予以公布；在提供链接服务器时，应该详细说明对链接源系统和链接目标系统的要求，并予以公布。

(5) 对于面向分布接收者的信息报道和推送服务，应该遵循RSS协议^[50]，建立规范的RSS信息聚合接口，支持第三方建立信息聚合服务。

(6) 对馆际互借服务，可以采用ISO的开放系统互连环境下馆际互借协议（ISO 10160 和ISO 10161）^[51]，支持第三方系统的馆际互借功能相互交换和处理馆际互借信息。

(7) 必须密切跟踪有关应用服务领域的标准规范发展，及时采用合适的标准来促进应用服务本身的可互操作性。

5.4 系统服务描述标准规范应用的实施要求

数字图书馆系统作为特殊的服务对象，应该利用开放语言进行规范描述，以保证用户和第三方系统能够搜寻、分析和调用所需要的系统服务。

(1) 应该按照Web Services框架^[52]来建立系统服务的描述、发布、登记、发现和调用机制，支持对系统服务的公开发布、公共查询和开放调用。

(2) 应该建立基于WSDL^[53]的服务系统格式，作为系统描述元数据，对服务系统的操作类型、输入输出数据流、数据格式、操作协议、与传输协议的捆绑机制、与安全控制机制的捆绑方式、网络位置等进行描述，成为系统服务的标准界面。

(3) 应该采用开放元数据机制来建立系统服务的描述元数据，通过复用、嵌套或链接方式描述服务系统所采用的有关界面、格式、组织语言、操作协议、控制机制等。

(4) 应该保证系统描述元数据与相应的资源集合元数据、资源集合内容结构元数据、资源集合知识组织系统元数据、资源集合管理机制元数据之间的有机关联，应该保证这些元数据之间的连接、调用和验证。

(5) 上述服务系统描述元数据应该置于服务系统公知位置，同时应该在全国信息服务登记系统进行公开登记，供第三方系统开放搜寻、公共查询和调用。

6. 组织管理机制的标准规范建设与应用的实施要求

数字图书馆系统往往构成按照一定的主题、资源类型、用户范围、使用管理机制等因素组织起来的资源集合，对这些集合本身、集合的对象组织结构、集合的知识组织体系的描述，有助于发现、解构、浏览、集成相关的资源集合。传统地，这些描述往往通过本地系统的内在程序机制来即时呈现。当对资源及其结构的描述以元数据形式描述时，既可以通过将这些元数据与具体系统绑定来形成实际的呈现机制，又可以通过对这些元数据的修改来方便地实现呈现机制的定制和修改，还可支持智能代理对资源集合的自动搜寻和匹配，支持智能代理对资源结构的解析以及在此基础上对多个资源集合的集成。

数字图书馆系统通过一定的机制来实现对所提供服务的管理。这些管理机制的核心实质是相应的管理政策和规则，这些政策和规则建立在特定管理模式基础上，通过一组描述该模式中各个实体及其相互关系的管理词汇来具体描述。当管理政策和规则以元数据形式（管理元数据）描述时，既可以通过将管理元数据与具体系统绑定来形成实际的程序运行机制，又可以通过对管理元数据的修改来方便地实现管理机制的定制和修改，还可支持智能代理对管理政策和规则进行搜寻、解析、并根据它们进行相应自动处理。

为了支持对数字资源与服务的组织管理机制的规范描述，许多领域正在探索相应的标准规范。数字图书馆系统应该紧密跟踪这方面的发展，及时引入相应标准规范，支持对数字图书馆系统的广泛使用和有机集成。

6.1 资源集合描述标准规范应用的实施要求

资源集合本身的描述构成资源集合元数据，支持人们对资源集合的规范描述、公共登记和开放检索。

(1) 数字图书馆系统都应该对自己的资源集合进行规范描述，建立相应的元数据记录。数字图书馆系统应该采用《数字图书馆标准与规范建设》项目组制定的我国数字图书馆资源集合元数据格式，对资源内容范围、资源建设过程、资源位置、资源责任者、资源使用与管理条件、与其它资源集合关系等内容进行描述，支持对资源集合的发现。数字图书馆系统也可以采用RSLP CDS格式^[54]和NSDL资源集合格式^[55]。

(2) 资源集合元数据的描述对象可以是多个层次，主要是门户级（系统级）、频道级和数据库级，并通过资源集合元数据内在机制将相互关联的多个层次连接起来。

(3) 数字图书馆系统可以在上述资源集合元数据的基础上，采取开放元数据机制进行扩展，描述资源集合中所涉及的数字对象标识规则、数字对象元数据格式、资源集合结构组织机制、资源集合知识组织体系等进行描述，支持对资源集合的智能利用。这些被描述的内容可以直接嵌入资源集合元数据中，也可以作为外部元数据被链接。

(5) 数字图书馆系统可以在上述资源集合元数据的基础上，采取开放元数据机制进一步扩展，描述资源集合的资源选择规则、元数据编制规则、资源使用政策、知识产权管理政策、隐私保护政策、资源长期保存政策等，支持对资源集合的分析选择以及管理。这些被描述的内容可以直接嵌入资源集合元数据中，也可以作为外部元数据被链接。

(6) 资源集合元数据应该采用 XML Schema 方式进行定义，其内容代码体系和语义定义应该参照数字对象元数据标准规范实施要求中的有关规定。应该建立资源集合元数据与基本元数据和 RSLP CDS 的标准转换关系，并建立基于 XSLT 的转换模板。

(7) 资源集合元数据的扩展内容构成新的资源集合元数据，应该采用基于 XML 的格式定义，应该采用相关领域已有的标准或通用格式，如果自己定义格式时应该采用相关领域已有的定义语言。

(8) 数字图书馆系统应该在自己系统公知位置上发布资源集合元数据文件，并在全国数字图书馆资源登记系统登记自己的资源集合元数据。

6.2 资源集合组织机制标准规范应用的实施要求

资源集合的组织结构机制包括资源集合组织结构和资源集合知识组织体系；

相应的标准规范，包括这些结构或体系本身（例如 Information Architecture、分类法、叙词表等），也包括对这些结构或体系及其与资源内容的联结方式进行定义和描述的标准规范。本指南主要涉及后一类标准规范。

（1）应该积极建立对资源集合的组织结构或知识组织体系进行定义和描述的规范语言，这些语言应该支持对复杂的集合结构或知识关系的描述。利用这些语言对具体资源集合进行描述后生成的结果，是资源集合组织结构元数据或知识组织体系元数据。

（2）可以建立基于XTM语言^[56]的资源集合组织结构描述语言，可以采用ISO 13250/XTM标准^[57]，可以选择在一个XTM文件中描述本系统各层集合的细节，也可选择采用嵌套的多个XTM文件来分别描述。

（3）可以建立基于XTM的知识组织体系描述语言，形成XTM知识组织元数据。可以利用基于XML的词表描述语言建立知识组织系统描述语言，形成开放词表系统。可以基于知识本体语言（例如OWL^[58]）建立知识组织体系描述语言，形成可推理的知识组织系统。

（4）可以采用XTM将组织结构元数据与知识组织体系元数据结合在一起，但由于知识组织体系侧重描述知识概念及其关系，组织结构侧重描述资源集合的内容单元及其关系，从模块化、可解析性和可维护性考虑，各自作为相对独立又互相联系的元数据更为合适。

（5）在建立描述语言时，应该严格基于XML语言，应该积极采用优势业务领域通用的描述语言，以保证组织结构元数据和知识组织体系元数据的开放性、可解析性和可扩展性。

（6）数字图书馆系统应该在自己系统公知位置上发布自己采用的资源集合组织结构和知识组织体系描述语言，发布基于这些语言所生成的组织结构元数据和知识组织体系元数据，并在全国数字图书馆资源集合登记系统上公开登记这些元数据，支持第三方系统自动发现和调用这些元数据。

6.3 资源集合管理机制标准规范建设的实施要求

资源集合管理机制涉及多个对象和若干层次，例如资源选择规则、使用控制政策、知识产权管理政策、隐私保护政策、长期保存政策、个性化政策等。由于管理机制实际上是管理政策的表现形式，通过对管理政策的规范描述，可以支持对管理机制的发现、解析、链接和整合，从而支持对资源和服务的分析、选择、组合和控制。管理机制的标准规范建设主要涉及对资源集合管理政策的规范描述。

（1）应该明确分析自己的管理需求，建立明确的管理政策，以文本形式在本系统主页予以公布，并提供明确的版本信息。

(2) 应该跟踪技术支撑领域和密切相关领域（作为业务优势领域）在管理机制标准规范方面的发展，积极采用这些领域的通用或主流标准规范。

(3) 应该在若干互操作迫切、管理政策已形成初步基础、已有可参照元数据的管理领域（例如使用管理、权益管理、开放链接、定制与推送等），积极推动建立管理模式框架，描述相应的实体、实体关系、管理过程、管理规则，并建立公共的管理词汇集。

(4) 应该积极利用管理模式框架和公共管理词汇集，积极选择开放的和具有良好推广前景的管理规则描述语言，建立各个业务管理的政策描述元数据格式，支持各个数字图书馆系统依照这些格式规范化和开放地描述和标记自己的管理政策。

(5) 数字图书馆系统应该在资源集合描述机制中建立开放接口，使其可以嵌入或链接相应的管理机制元数据，从而可以通过资源集合元数据对管理机制元数据进行调用。

(6) 管理机制元数据格式应该严格基于 XML 语言，应该采用通用的应用描述语言，保证管理机制元数据的开放性和可解析型。而且，内容代码体系和语义定义应该参照数字对象元数据标准规范建设实施要求中有关规定。

(7) 应该建立管理机制元数据的编制与登记发布规则，在各自系统公知位置上发布相应元数据文件，并在我国数字图书馆公共服务登记系统公开登记。

(8) 应该保证一个数字图书馆系统的资源集合元数据、资源集合组织结构元数据、资源集合知识组织体系元数据、资源集合管理机制元数据之间的有机关联，应该保证这些关于资源集合的元数据与资源集合中的数字对象元数据格式和数字对象唯一标识符规范的有机关联，从而保证这些元数据之间的连接和调用。

7. 数字资源系统长期保护标准规范建设的实施要求

为了保证数字信息资源的长期可获得性，数字图书馆系统、尤其是自身生产原生数字资源、并作为主要的资源提供者的数字图书馆系统，应该建立相应的长期保存机制，并通过规范的管理机制和技术机制来保证长期保存过程的可靠性。

(1) 应该从全面保存内容比特流、数字格式与处理信息、数字信息呈现与服务环境的角度，系统考虑保存对象范围、保存责任体系、保存对象选择鉴别机制、数据捕获转移机制、元数据描述、技术保存机制、知识产权管理机制、经济支撑机制、长期保存资源的利用机制等，建立长期保存的决策分析框架。

(2) 应该建立明确的长期保存战略，系统考虑保存对象范围界定、责任分配、对象选择鉴别规范与程序、对象搜集与捕获的规范与程序、长期保存资源的描述与技术处理规范、知识产权管理规范与程序、长期保存资源的利用管理等，建立长期保存的政策管理框架。

(3) 应该遵循OAIS参考模型 (ISO 14721:2002) [59], 建立长期保存系统的功能框架和信息框架, 包括摄取模块 (Ingest)、长期存储模块 (Archival Storage)、数据管理模块 (Data Management)、系统管理模块 (Administration)、保存计划模块 (Preservation Planning) 和检索传递模块 (Access)。在此基础上, 参考 RLG/OCLC 的可信赖数字存储库属性要求 [60], 界定数字信息长期保存系统具备的基本条件和责任体系。

(4) 《数字图书馆标准与规范建设》项目将制定我国数字资源长期保存元数据格式。在此之前, 应该按照OAIS信息模型选择采用通用的长期保存元数据格式 (例如CEDARS [61]、PANDORA/NLA [62]、NEDLIB [63]、RLG/OCLC格式 [64]), 或在自己的元数据格式中包含长期保存所需要的内容信息 (数据内容本身和数据内容表现)、保护描述信息 (确认信息、起源信息、环境信息、固化信息)、封装信息和内容描述信息, 其中内容描述信息应该采用基本元数据格式。我国数字资源长期保存元数据格式制定后, 应该采用该格式。

(5) 应该建立规范的数字对象技术保护规范 (例如数据更新、技术仿真、数据迁移) 和数字保存系统管理规范。应该建立长期保存政策的管理和发布规则, 在各自系统的公知位置上发布, 并作为系统规范的一部分在全国数字图书馆公共服务登记系统公开登记。

参考文献

- [1] 张晓林、王惠临、牛振东等.我国数字图书馆标准规范的发展战略。2004.4
- [2] 张晓林、牛振东等.我国数字图书馆标准规范的总体框架。2004.4
- [3] Extensible Markup Language (XML) 1.0 (Third Edition), W3C Recommendation 04 February 2004. <http://www.w3.org/TR/REC-xml> (2006/01/04 检索)
- [4] World Wide Web Consortia. <http://www.w3.org/> (2006/01/04 检索)
- [5] Organization for the Advancement of Structured Information Standards (OASIS) <http://www.oasis-open.org/home/index.php> (2006/01/04 检索)
- [6] Scalable Vector Graphics (SVG) 1.1 Specification. W3C Recommendation 14 January 2003 <http://www.w3.org/TR/SVG11/> (2006/01/04 检索)
- [7] Synchronized Multimedia Integration Language (SMIL 2.0) . W3C Recommendation 07 August 2001 <http://www.w3.org/TR/smil20> (2006/01/04 检索)
- [8] Speech Synthesis Markup Language Specification. W3C Working Draft 2. December. 2002 <http://www.w3.org/TR/speech-synthesis/> (2006/01/04 检索)
- [9] Mathematical Markup Language (MathML) Version 2.0. W3C Recommendation 21 February 2001 <http://www.w3.org/TR/MathML2> (2006/01/04 检索)
- [10] Geography Markup Language (GML) 2.0, OpenGIS® Implementation Specification, 20 February 2001 <http://www.opengeospatial.org/docs/01-029.pdf> (2006/01/04 检索)

- [11] Chemical Markup Language. 15.July. 2003 <http://www.xml-cml.org/> (2006/01/04 检索)
- [12] Extensible Markup Language (XML) 1.0 (Second Edition) W3C Recommendation 6 October 2000
<http://www.w3.org/TR/REC-xml>. (2006/01/04 检索)
- [13] XML Schema Part 0: Primer W3C Recommendation, 2 May 2001.
<http://www.w3.org/TR/xmlschema-0/> (2006/01/04 检索)
- [14] TIFF Rev 6.0,3.June. 1992. <http://partners.adobe.com/asn/developer/pdfs/tn/TIFF6.pdf>
(2006/01/04 检索)
- [15] JPEG Specification. Sept. 1992 <http://www.w3.org/Graphics/JPEG/itu-t81.pdf> (2006/01/04 检索)
- [16] PDF Reference, Third Edition, Version 1.4. Addison-Wesley, 2001 (ISBN 0-201-75839-3)
- [17] Graphics Interchange Format, Version 89a. July 31, 1990
<http://astronomy.swin.edu.au/~pbourke/dataformats/gif/> (2006/01/04 检索)
- [18] Portable Network Graphics. 01.October.1996 <http://www.w3.org/TR/REC-png.html>
(2006/01/04 检索)
- [19] MPEG. <http://www.mpeg.org/MPEG/index.html> (2006/01/04 检索)
- [20] Apple Quicktime Developer Resources <http://developer.apple.com/quicktime/> (2006/01/04 检索)
- [21]Microsoft Windows Media Technologies.
<http://www.microsoft.com/windows/windowsmedia/default.asp> (2006/01/04 检索)
- [22]MP3 Moving Picture Experts Group, Audio Layer III
- [23] Microsoft Windows Media Technologies.
<http://www.microsoft.com/windows/windowsmedia/default.asp>
(2005/07/12 检索)
- [24] Apple Quicktime Developer Resources <http://developer.apple.com/quicktime/> (2006/01/04 检索)
- [25] Waveform Audio File Format. <http://www.borg.com/~jglatt/tech/wave.htm> (2006/01/04 检索)
- [26] Internet Engineering Task force. <http://www.ietf.org/> (2006/01/04 检索)
- [27] Van de Sompel, H., Hammond, T., Neylon, E., and Weibel S.L. "The "info" URI Scheme for Information Assets with Identifiers in Public Namespaces",December 2003.
<http://www.ietf.org/internet-drafts/draft-vandesompel-info-uri-01.txt> (2006/01/04 检索)
- [28] ISO 11179: Specification and Standardization of Data Elements.
http://metadata-standards.org/11179-2/ISO-IEC_11179-2_2000_IS_E.pdf (2006/01/04 检索)
- [29] Resource Description Framework (RDF) <http://www.w3.org/RDF/> (2006/01/04 检索)
- [30] Dan Brickley, University of Bristol R.V. Guha, Epinions Resource Description Framework (RDF) Schema Specification 1.0 W3C Candidate Recommendation 27 March 2000
<http://www.w3.org/TR/rdf-schema> (2006/01/04 检索)
- [31] Leigh Dodds XML Namespaces 1.1 April 10.2002
<http://www.xml.com/pub/a/2002/04/10/namespaces-1.1.html> (2006/01/04 检索)
- [32] James Clark XSL Transformations (XSLT) Version 1.0 W3C Recommendation 16 November 1999 <http://www.w3.org/TR/xslt> (2006/01/04 检索)
- [33]Web Content Accessibility Guidelines 1.0. W3C Recommendation 5-May-1999
<http://www.w3.org/TR/WAI-WEBCONTENT> (2006/01/04 检索)

- [34] Andy Powell, UKOLN, Pete Johnston. 2003-04-02. Guidelines for implementing Dublin Core in XML.
<http://dublincore.org/documents/2003/04/02/dc-xml-guidelines/> (2006/01/04 检索)
- [35] MARCXML10.June.2003 <http://www.loc.gov/standards/marcxml/> (2006/01/04 检索)
- [36] Metadata Object Description Schema. <http://www.loc.gov/standards/mods/> (2006/01/04 检索)
- [37] Metadata Encoding and Transmission Standard <http://www.loc.gov/standards/mets/> (2006/01/04 检索)
- [38] IMS Content Packaging. <http://www.imsglobal.org/content/packaging/index.cfm> (2006/01/04 检索)
- [39] ISO/IEC WD 21000-2:2001. Information Technology—Multimedia Framework-Part 2: Digital Item Declaration, March 2001
<http://xml.coverpages.org/MPEG21-WG-11-N3971-200103.pdf> (2006/01/04 检索)
- [40] Simple Object Access Protocol. <http://www.w3.org/TR/soap/> (2006/01/04 检索)
- [41] Creating Open Standards for Communicating 3D. <http://www.web3d.org/> (2006/01/04 检索)
- [42] Z39.50 Standards. <http://www.loc.gov/z3950/agency/> (2006/01/04 检索)
- [43] The Bath Profile: An International Z39.50 Specification for Library Applications and Resource Discovery
 Feb/2001 <http://www.nlc-bnc.ca/bath/bp-current.htm> (2006/01/04 检索)
- [44] ONE-2 Profile. <http://www.portia.dk/pubs/one2/Profiles/ProfileV3R2/ONE-2ProfileV3R2.pdf> (2006/01/04 检索)
- [45] Z39.50 Search/Retrieval Web Service.
<http://lcweb.loc.gov/z3950/agency/zing/srw/version.html> (2006/01/04 检索)
- [46] Lightweight Directory Access Protocol (Version 3) Specifications Dec.30.1998
<http://www.innosoft.com/ldapworld/ldapv3.html> (2005/07/12 检索)
- [47] Open Archive Initiative. <http://www.openarchives.org/> (2006/01/04 检索)
- [48] Van de Sompel, H., Hochstenbach, P. and Beit-Arie, O. "OpenURL Syntax Description," 2000. <http://www.openurl.info/registry/docs/pdf/openurl-01.pdf> (2006/01/04 检索)
- [49] OpenURL Framework Standard. ANSI/NISO Z39.88.
<http://library.caltech.edu/openurl/Standard.htm> (2006/01/04 检索)
- [50] Rich Site Summary (RSS) <http://www.oasis-open.org/cover/rss.html> (2006/01/04 检索)
- [51] Documentation - Interlibrary Loan Application Protocol Specification -
 Amendment 1: ILL Support for Electronic Document Delivery ISO TC46/SC4/WG4 21
 November 1995
 Status: Final IS text
- [52] Kreger, H. IBM Web Services Conceptual Architecture 1.0, May 2001
<http://www-4.ibm.com/software/solutions/webservices/pdf/WSCA.pdf> (2006/01/04 检索)
- [53] Web Services Description Language (WSDL) Version 1.2. W3C Working Draft 9 July 2002
<http://www.w3.org/TR/2002/WD-wsd112-20020709/> (2006/01/04 检索)
- [54] RSLP Collection Description. <http://www.dlib.org/dlib/september00/powell/09powell.html> (2006/01/04 检索)
- [55] NSDL Metadata Primer. <http://metamanagement.comm.nslib.org/outline.html> (2006/01/04 检索)

- [56] XML Topic Maps (XTM) 1.0, TopicMaps.Org Specification.
<http://www.topicmaps.org/xtm/1.0/> (2006/01/04 检索)
- [57] ISO/IEC 13250 Topic Maps. Information Technology Document Description and Processing Languages
Technologies de l'information -- Cartes topiques
http://www.y12.doe.gov/sgml/sc34/document/0322_files/iso13250-2nd-ed-v2.pdf (2006/01/04 检索)
- [58] OWL Web Ontology Language Guide, W3C Recommendation 10 February 2004
<http://www.w3.org/TR/owl-guide/> (2006/01/04 检索)
- [59] http://www.gea.nu/pdf/tillit/Arkivering_GEA.pdf (2006/01/04 检索)
- [60] Attributes of a Trusted Digital Repository: Meeting the Needs of Research Resources
An RLG-OCLC Report DRAFT FOR PUBLIC COMMENT RLG Mountain View, CA August 2001
<http://www.rlg.org/longterm/attributes01.pdf> (2006/01/04 检索)
- [61] Metadata for Digital Preservation : the Cedars Project Outline Specification.
<http://www.leeds.ac.uk/cedars/colman/metadata/metadataspec.html> (2006/01/04 检索)
- [62] National Library of Australia Preservation Metadata for Digital Collections. Exposure Draft
<http://www.nla.gov.au/preserve/pmeta.html> (15 October 1999) (2006/01/04 检索)
- [63] Metadata for long term-preservation, July 2000.
<http://www.kb.nl/coop/nedlib/results/D4.2/D4.2.htm> (2006/01/04 检索)
- [64] <http://www.rlg.org/legacy/longterm/repositories.pdf> (2006/01/04 检索)