

“CALIS 重点学科网络资源导航库”^{*} 标准与规范述评

□张惠君 张春红 萧德洪 林佳

摘要 “十五”期间,共54家高校图书馆参与“CALIS 重点学科网络资源导航库”项目的导航资源建设工作。为保证信息发现、资源描述等行为的一致性和连续性,项目组制订了相关标准与规范。该文对项目所涉及的标准和规范的建设情况进行了述评、探讨,包括资源选择标准、元数据规范、著录规则和资源类型设定、分类体系的选择等。

关键词 学科导航 标准 规范 CALIS

“CALIS 重点学科网络资源导航库”(以下简称“导航库”)是“十五”期间 CALIS 文献服务环境建设的重要组成部分。2003年10月,CALIS“十五”规划建设导航库子项目正式启动,西安交通大学图书馆作为项目牵头单位,负责项目的组织实施。自2003年10月至2005年底,54家高校图书馆共同完成了79个一级学科的导航资源建设工作,工程量大、参与单位多是该项目实施的一个特点。

根据“十五”CALIS 建设目标并针对导航库参建馆众多的现状,在项目启动伊始,项目组就制订了“统一平台、统一标准、合作建设、分头维护、集中服务、全国共享”的建设方针^[1]。其中的“统一标准”是保证数据质量的重要前提。学科导航库的开发与维护是一件非常费时费力的工作,要建设一个覆盖众多学科的导航系统,单靠某一个馆的力量是负担不了的,需联合多家单位跨地区协作完成。为保证信息发现、资源描述等行为的一致性和连续性,必须有一套标准、规范对其进行控制。CALIS 导航库标准与规范建设涉及资源选择、资源描述等多个方面。为此,项目组特别组织人力制订了 CALIS 重点学科网络资源导航库之《资源选择与评估标准》、《网络资源元数据规范》、《网络资源元数据描述著录规则》、《资源类型控制表》(分别简称为《资源选择标准》、《元数据规范》、《著录规则》、《资源类型表》)四个标准与规范,并为导航库

选择确定了合适的分类体系。

1 资源选择标准

学科门户的显著特点是其信息资源集合的质量受控特性^[3]。CALIS 导航库的目的是为高校教学和科研提供重要学术网站的导航和免费学术资源的导航,其资源的高质量意味着对读者需求的高满足率,它是导航库生存和发展的前提。

资源的高质量有赖于一套科学、客观、简便易行的资源选择标准。资源选择标准定义了与信息环境中选择资源的依据,它通常规定了资源的范围选择和质量选择策略。资源选择标准不仅可以保证分散在各地的众多参建馆在选择资源时遵循统一的标准,也可保持在今后漫长的数据库维护阶段与建库阶段工作的一致性。

1.1 基本内容^[2]

资源选择实际上是对资源进行评估的过程。CALIS 导航库《资源选择标准》主要从资源内容、资源形式、资源时效性和专家意见四个方面对资源进行评估,权重分别为50%、20%、10%、20%。从权重可以看出,CALIS 导航库特别强调资源内容及其专家意见的重要性。表1给出了导航库《资源选择标准》框架。

* 本文为“十五”期间“CALIS 重点学科网络资源导航库”项目研究成果之一。

表 1 导航库《资源选择标准》基本内容

评估项目	指标(说明)	权重
资源内容	用 7 个指标考查,分别为: 适用性 权威性 实用性 准确性 全面性 独特性 客观性	50%
资源形式	用 6 个指标考查,分别为: 内容形式 使用便利性 规范性 美观性 稳定性 影响力	20%
资源时效	主要考查反映学科发展的及时性、最后更新日期等。	10%
专家综合意见	由相关学科专家评分	20%

其中每个评估指标都给出了具体的考查内容,例如,用“可被多个外部的搜索引擎检索到”考查资源的影响力,用“资源发布者是有影响的团体或个人”及“服务器稳定、能提供正常服务”考查资源的稳定性。

1.2 使用尺度与评价

(1) 确切地说,资源选择标准是指导资源选择的策略,在使用时不可能像著录规则那样“钉是钉、铆是铆”地执行,它为资源发现者提供的仅是一个选择资源的依据。

(2) 根据《资源选择标准》,项目组推荐使用的是定性定量相结合的资源评估方法,不仅给出了每个评估项的权重,而且提供了一个详细的打分表。但是在实际使用时,出于建库效率等方面的考虑,较多采用定性的方式进行评估。对网络信息资源定性评估极易受主观因素的影响,要注意不能因评估者个人喜好而影响对资源价值的正确判断和对资源的选择,尽可能做到客观地评价和选择^[2]。

(3) 学科导航的目的并非搜集某学科的所有信息资源,它以质量控制为手段,目标是提供某学科领域可获得的、满足特定读者群需求的最优秀网络资源^[3]。因此,导航库项目特别强调在资源选择和评估过程中专业人员参与的重要性和必要性,并以此作为项目验收的一个指标。

(4) 资源发现是资源选择的前提,它是一个费时费力并需要专业知识支撑的工作(这正是导航库存在的理由)。导航库建库工具(本地系统)提供了两种发现资源的途径,一是基于搜索引擎搜索,另一个是利用爬行技术的智能搜索。无论哪种方式,都需要搜索者不断尝试,制定出科学、高效的检索策略。在建库实践过程中,应积极思考、积累经验,摸索出适合自己的资源发现方法。

(5) 每个学科的资源在因特网上的出现是不均衡的,有的非常多,有的却很少,这一现象的存在要求我们在使用标准时要灵活把握尺度。对于网络资源很丰富的学科,在评价、筛选资源时,应较为严格地遵照评估指标要求进行,即选择那些符合多项评估指标的资源。相反地,对于网络资源稀缺的学科,可以适当降低选择门槛,或许只满足一、二个评估指标就可入选。事实上,绝大多数资源不可能兼备《资源选择标准》中列举的那些特性,我们在进行资源选择时只能择优选取,把相对优秀的资源提供给读者。

2 元数据规范

学科导航库建立了第三方的元数据记录集合,即在数据库中用字段描述网络资源的不同属性,以帮助读者了解更多的网络资源。

在因特网上,有许多简单的学科导航门户,它们是由包含超链接列表的静态网页构成的,这种门户只可浏览、不可检索,并且只能通过手工修改网页的方法完成维护。利用数据库技术建立的导航库系统可以克服上述缺陷,但是需要使用一个元数据格式来建立和存储网络资源的描述信息。元数据格式就像一个表格或模板,编目人员只需将资源特性填入其中即可完成对资源的描述。与使用 HTML 列表建立学科门户相比,使用元数据具有明显的优越性^[3]: 元数据是结构化的,它是导航库实现高级检索功能的基础; 元数据可在需要的时候转换成其他格式,并且支持与其他不同检索协议的互操作; 维护一个资源描述数据库比维护一大堆 HTML 文件要容易得多。由此可见, CALIS 导航库制订和使用元数据规范是必须的。

2.1 元数据规范制订依据

使用什么元数据格式对导航库来说是一个重要选择,它不仅影响导航库系统的检索性能,而且关系到面向读者的资源描述的优劣。元数据资源是导航库最重要的财富,它所提供的独特的、详细的、标准化的描述信息为网络资源增值。但是,从经济和效率的角度考虑,并非字段越多越好。在导航库运行过程中,资源描述数据库的建立与维护是一项艰巨的任务,需要花费大量的人力和时间,因此,选择元数据标准必须考虑价值与费用的平衡。

选择或制订元数据格式通常需要考虑以下问题:

(1) 在网络资源的诸多属性中,选择哪些属性进

行描述就可基本反映其面貌,即设置哪些字段可较全面地反映一个资源的情况?

(2) 将来读者浏览和检索时会用到哪些字段?检索结果显示给读者时,应显示哪些字段?

(3) 管理人员进行导航库系统维护时需要哪些字段?例如,想跟踪资源的更新情况,系统可定期自动地给资源提供者发送电子邮件,这就需要某个字段提供相关的地址信息。

总之,导航库元数据既要能反映资源特征,又要便于使用和管理。

2.2 基本内容^[4]

2004年,科技部启动“我国数字图书馆标准规范建设”项目,并出台一系列标准与规范,其中的《网络资源描述元数据规范》(上海图书馆起草)专门针对网络资源描述而设计,CALIS 导航库《元数据规范》则是在此基础上制订的。

导航库《元数据规范》中元数据分为两种类型,即描述性元数据和管理性元数据。其中描述性元数据是根据导航库的实际需要从科技部《网络资源描述元数据规范》中选择了15个核心元素,并对元素的修饰词进行适当的扩展,比如题名增加了缩略题名和并列题名。对于其非核心元素导航库没有采用,而是根据自身特点设计了导航库核心元素和个别元素。管理性元数据是按照导航库的需要自行设计的。在设计规则方面,也部分遵循科技部标准的规定。

导航库元数据结构分为2个部分,即描述性元数据(18个元素)和管理性元数据(8个元素)。表2为导航库《元数据规范》框架^[4]。

表2 导航库《元数据规范》基本内容

类型	名称	属性
描述性元数据	题名 创建者 主题和关键词 其他责任者 出版者 描述 日期 格式 资源类型 资源标识符 来源 语种 相关资源 覆盖范围 权限管理	核心元素
	学科 版本	核心元素
	推荐级别	个别元素
管理性元数据	记录编目者 记录审核者 记录 编目日期 记录审核日期 记录提交 日期 记录更新日期 记录失效日期 使用方式	核心元素

3 著录规则

元数据规范规定了导航库使用的元数据格式,而著录规则的任务则是确认元数据格式的内容如何录入。一旦确定了元数据规范,接下来要做的是制订一套编目规则,用以规定如何使用元数据格式对资源进行描述。

学科导航系统的重要任务之一是建立网络资源的描述性元数据,它既是浏览和查询导航库的基础,又可帮助读者确认他们真正需要的资源,为读者节省大量宝贵的时间。对资源进行描述的行为通常被称为编目,通过编目建立的记录可帮助读者辨别、定位、访问和使用网络资源。与传统图书馆对图书的编目相似,对网络资源的描述同样需要遵循相关著录规则以及一些附加规则,如学科分类、资源类型、规范控制等。

与《元数据规范》相适应,导航库《著录规则》基于科技部立项、上海图书馆起草的《网络资源描述元数据规范(初稿)著录规则》而制订。

3.1 制订原则^[5]

针对网络资源本身的特征和CALIS 导航库建设的具体要求,在制定著录规则的过程中主要遵循了以下几大原则:

(1) 简单性和准确性原则。由于网络资源数量庞杂并处于经常变化之中,其著录规范应该简单易用,但追求简单性的前提是保证能对网络资源进行准确无误的描述与揭示。

(2) 通用性和容纳性原则。网络资源类型多样,可以区分为文本资源、图像资源、视频资源、声频资源等诸多类型,因此其著录规则应该既具有一定的通用性,又必须具备某种程度上的容纳性。

(3) 标准性和互操作性原则。为使CALIS 导航库建设成果能在更大范围内得到广泛利用,其著录规则应该与现有的著录规则保持基本的一致性。因此,在制定过程中充分参考了《中文文献编目规则》、《西文文献著录条例》以及ROADS和RDN的著录规则。

3.2 基本内容^[5]

CALIS 导航库建设《元数据规范》主要是在《网络资源元数据规范》(上海图书馆制定)和《管理性元数据规范》(CALIS 制定)的基础上并根据导航库项

目建设的某些特殊需求制定的,其元数据即著录项目包括 18 个描述性元数据及 8 个管理性元数据(见表 2)。描述性元数据中,必备元素有 6 个,分别为题名、主题和关键词、资源类型、资源标识符、语种、学科;而管理性元数据中,除记录更新日期外,其余均为必备元素。

著录信息源即为著录对象本身。对于网络资源,其主要信息源为网站的首页或资源所在的具体网页。编目语言用简体中文。描述性元数据中对题名项、责任说明项、版本项、出版发行项等资源本身所显示的内容进行著录时,著录文字必须按资源所用的文字转录;由著录人员给出的摘要、资源评价等信息,著录文字用简体中文^[5]。

4 关于分类体系与资源类型

如前所述,针对一个资源的著录包括描述性信息和管理性信息,管理性信息是系统自动生成的,不需要编目人员手工著录。在诸多描述性元素中,有些可根据资源本身提供的信息转录,有些则必须根据项目需求专门规定或制订。CALIS 导航库《著录规则》涉及的标准如下^[5]:

(1) 日期:建议采用 ISO 8601[W3CDTF] 规范进行著录,并使用 YYYY-MM-DD 格式。

(2) 格式:建议采用来自于受控词表中的值,使用[MIM E] 资源数据格式表,例如用 text/html 表示网页,用 image/gif 表示图形文件。

(3) 语种:建议采用“世界语种代码表”,例如,用 chi 表示汉语,用 jpn 表示日语。

(4) 学科:统一采用教育部正式颁布的《授予博士、硕士学位和培养研究生的学科、专业目录》(1997 年颁布)。

(5) 资源类型:采用 CALIS 导航库项目专门制订的《资源类型表》。

4.1 分类体系的选择

学科分类是通过资源的学科属性对其进行描述的一种方法。不同于标题、主题等学科索引系统,它试图建立一个相关学科资源的层次结构,通过使用符号或代码可方便地建立层次结构的学科树。

使用学科分类系统组织网络资源相比于早期用链接列表进行资源导航有着明显的优势,主要表现

在^[3]:易于浏览;缩小查询范围;为检索词提供上下位类,可解决同字异义词(或拼写相同不同义)带来的问题;可以将大的导航库按需要逻辑地分割成若干子库,方便数据库操作;通过数据库提供的学科浏览和高级检索,增强了导航库系统的互操作性。

选择或制定分类体系时应从分类体系本身和终端用户这两个角度进行考虑^[1]:

- 分类法应该是系统的、科学的;

- 分类法应该是成熟的、稳定的(经常变动的分类体系对用户使用不利);

- 所选择的分类体系应该是它的用户所熟悉的且经常使用的(有利于读者浏览);

- 类目级别适当,不宜过粗或过细。

基于上述考虑,CALIS 导航库项目决定使用教育部颁布的《授予博士、硕士学位和培养研究生的学科、专业目录》作为构建导航库分类体系的依据。该分类体系由 12 个学科门类构成,囊括社会科学和自然科学所有学科领域,全面、系统又不失简洁,在高校学科设置和重点学科建设中被广泛使用,并为广大师生所熟悉。

4.2 资源类型表

在建设学科导航系统的过程中,应该对哪些类型的资源进行收集与揭示是非常重要的问题,我们熟知的 CSDL、INFOMINE、SOSIG 等项目均采用了各自的资源类型标准。CALIS 导航库项目对国内外多个知名学科导航门户的资源类型标准进行了研究,分析其优点与缺陷,最终确定了自己的资源类型表。

资源类型的设定应从资源本身、用户需求、今后发展等角度进行全面考虑,遵循学术性、准确性、可操作性、通用性等原则^[6]。

CALIS 导航库的资源类型设定是一个二级体系,其中一级是属于数据库层次的,供系统内部进行逻辑组织使用,二级为用户端层次的,供标引和网页呈现使用^[6]。

任何一个标准都不是一成不变的。随着导航库资源建设的不断深入以及导航库网站的开通,应根据各个标准和规范的使用跟踪情况对其进行必要的修订。例如,依据读者对资源的访问情况,分析高点

表3 《资源类型表》基本内容⁷⁾

第一级	第二级
参考资源	资源导航 辞典与百科全书 文摘与索引 统计资料 其他参考资料
全文数据库	数据库 电子期刊 研究报告 政府出版物
教学资源	教学资料
多媒体资源	图像资料 音频资料 视频资料
黄页资源	协会/学会 大学院系 研究机构/中心 其他组织机构 专家学者
交互资源	邮件列表 论坛/讨论组 新闻组 搜索引擎
事件	会议 研究项目 学术动态
其他	网站 服务 软件 杂类

击率资源的特点,重新审查读者需求,对资源选择标准进行修订、完善。

参考文献

- 张惠君等. CALIS“重点学科网络资源导航库”“十五”建设思考. 大学图书馆学报, 2004(2): 34- 37
- 林佳, 李京华. “CALIS 重点学科网络资源导航系统资源选择与

- 评估标准”的建立与思考. 大学图书馆学报, 2005(1): 38- 41
- DESIRE Information Gateways Handbook. <http://www.desire.org>
 - 张春红等. CALIS 重点学科网络资源导航库之网络资源元数据规范. 2005. [2006- 01- 09]. <http://202.117.24.24/html/CALIS/daohangku/ysj.pdf>
 - 张春红等. CALIS 重点学科网络资源导航库之网络资源元数据描述著录规则. 2005. [2006- 01- 09]. <http://202.117.24.24/html/CALIS/daohangku/zhlgz.pdf>
 - 萧德洪等. 学术图书馆学科导航门户资源类型表的设定. 大学图书馆学报, 2004(5): 28- 33
 - 萧德洪等. CALIS 重点学科网络资源导航库之资源类型表. 2005 [2006- 01- 09]. <http://202.117.24.24/html/CALIS/daohangku/zylx.htm>

作者单位: 西安交通大学图书馆, 西安, 710049
 北京大学图书馆, 北京, 100871
 厦门大学图书馆, 厦门, 361005
 清华大学图书馆, 北京, 100084

收稿日期: 2006 年 1 月 9 日

Reviewing on Standards and Specifications about CALIS Subject Gateway

Zhang Huijun Zhang Chunhong Xiao Dehong Lin Jia

Abstract: During the 10th 5-year Plan period, fifty-four libraries take part in the project of CALIS Subject Gateway and creat the metadata records of Internet resources all together. In order to keep the consistency in selecting and describing resources, standards and specifications are made. In this paper discusses standards and specifications applied for CALIS Subject Gateway project such as selection criteria, metadata format, cataloguing rules, resource type and classification scheme.

Keywords: Subject Gateways; Standards; Specifications; CALIS Portal

(上接第 27 页)

The Evolution of the Information Commons in U. S. Academic Libraries and Its Revelation

Ren Shuhuai Sun Guichun

Abstract: The Information Commons (IC) is a new model of integrated library service in academic library. It has become an essential part of today's academic library environment, often the most popular service and space in U. S. academic libraries. This paper presents the concept and evolution of IC, its models and characteristics. It also provides a survey data of the IC development in U. S. academic libraries and then makes suggestions for implementing innovative user-centered services in China academic libraries as follows. The IC has become a most significant trend in the academic library. The library building design should reflect the new idea of IC. It is very necessary for the library services to establish a reasonable structure of information resources. The staff with a wide variety of literacy skill is one of the key elements to enhance the library service quantity. Library services evaluation will guarantee the improvement of the library service quantity.

Keywords: Information Commons; Open Access Environment; Academic Library; Service Facility; Information Literacy

2006年3期

大学图书馆学报