

●王芬林

关于全国文化信息资源共享工程数字 资源建设的思考和对策

摘 要 文化信息资源是一个范围很广泛的概念。就力所能及而言,全国文化信息资源共享工程的资源建设应以文化部可控制资源和实施方案中既定的资源建设目标为内容。工程应着手建设数字化图书馆,可为实现其数字图书馆服务的终极目标奠定基础。在资源共建模式和选题策划上,也都应有科学的规划。今后的全国文化信息资源共享工程建设应在总体思路、建设目标、技术支持、知识产权配合、组织实施、用户服务、工作流程等方面进行调整和改进。

关键词 文化信息资源 全国共享工程 资源范围 共建模式 发展对策

分类号 G253

ABSTRACT Cultural information resource is a broad concept. Practically speaking, the development of digital resources for the National Cultural Information Resources Sharing Project should have the contents of resources controllable by the Ministry of Culture and resources clearly specified in the plan. The project should be focused on the development of digital libraries and to realize its final objects of services via digital libraries. In this paper, the author proposes some recommendations concerning the overall plan, objectives, technical supports, intellectual property, implementation, customer services, work flows, etc.

KEY WORDS Cultural information resources. National Cultural Information Resources Sharing Project. Scope of resources. Pattern of cooperation development. Development strategies.

CLASS NUMBER G253

数字资源建设是全国文化信息资源共享工程(以下简称共享工程)的建设核心,是工程的命脉,这已是共识,无需再来论证。笔者从2001年10月开始共享工程试验系统研发至今一直负责资源建设工

我国机构特点,采取纵向——行业内、横向——地区内双管齐下的合作,建立分布式存储系统。四是形成上下合作链,即个人与机构进行合作、基层机构与省级机构合作、省级机构与国家机构合作、国家机构与国际机构合作等。当然这不是固定的模型,基层机构如一般大学图书馆在有条件时也可直接参与国际合作,但从总体上形成这样的合作链有利于层层带动,也有利于为将来的国家长期保存体系结构的构建打下基础。

(4)积极推进学术性资源永久使用权的建设。我国各类图书馆近几年购置了大量网络数据库。如CALIS目前已组团引进了近20个厂商的90多个数据库(含电子刊),参加的高校遍布全国各地;中国国家科学数字图书馆项目管理中心在中国科学院院内组织引进了10多个网络数据库。对于网络数据库,使用权与拥有权通常是分离的,一旦终止订购,或受社会不可抗力因素影响,订购方都有可能瞬间就失去了所有资源。在目前法律机制下,订购方一定

要积极向商家争取所购置数字资源的永久使用权。永久使用权的获取是国家长期保存战略机制建立的基础、实施的重要组成部分。

另外还包括建立引进、借鉴和研究有关标准规范的机制,建立责任分配与认证机制,加强相关法律、经济运作模式、管理模式等问题的研究,积极开展自建数字资源的长期保存活动等。

参考文献

- 1 A Metadata Framework to Support the Preservation of Digital Objects Final Report (June 2002). http://www.oclc.org/research/pnwg/pm_framework.pdf
- 2 Task Force on Digital Repository Certification. <http://www.rlg.org/longterm/certification.html>.

宛玲 河北大学管理学院副教授,中国科学院文献情报中心在读博士生。通讯地址:北京中关村北四环西路33号。邮编100080。(来稿时间:2003-08-22)

作,始终被资源建设定位、资源目标用户、资源共建的模式、网站怎样策划设计才能吸引读者,以及“数字化图书馆和数字图书馆的区别、共享工程和普通网站的区别”等诸多问题所困扰。笔者认为,很有必要对一些似乎已成定论的问题进行进一步的讨论,以提出行之有效的措施,使有限的资金投入发挥更大的作用,加快共享工程的资源建设。

1 以往资源建设中需要再斟酌的几个问题

1.1 文化信息资源范围的界定

无论是在最初的共享工程资源建设实施方案和之后的后续建设计划中,还是在共享工程专家咨询委员会对资源建设的数次论证中,都把注意力集中到了“文化信息资源所应涵盖的范围”上。尽管一再强调工程的文化概念是狭义概念而非广义概念,但即使是狭义,也依然是无所不包。因此,为了吻合每个人心中的文化概念,就使得其含义范围在不断扩大,资源专题在不断增多,后续发展的困难也越来越大。如此,资源建设的范围之广足以让人望而却步,也已经背离既定目标越来越远。

其实,无论文化的范围可以有多么广的外延,国务院界定的文化部职能却是明确的,文化部直属单位所涉及的文化艺术类别也是明确的。就目前力量之所能及,共享工程应紧紧围绕“文化部可控制资源”以及“实施方案中的资源建设目标”,结合启动以来的实际经验教训,重新修订资源建设方案,投资并整合开发文化部自有资源,不要为更广大的用户需求所干扰。如果需求不在文化部所控制资源范围之内,可以不予考虑,因为我们目前的力量达不到。

1.2 数字图书馆技术的应用

共享工程以“试验型数字化图书馆”和“中国数字图书馆有限责任公司”在数字图书馆技术、资源建设方面的建设成果为基础,强调技术的与众不同,在资源建设和服务中,过分注重用资源去诠释数字图书馆的概念,对目前大众普遍理解并接受的数字图书馆(另将其定义为数字化图书馆)未给予足够的重视。在笔者现在看来,我们定义的“数字化图书馆”应该对应的是“传统图书馆”,其工作目标和落脚点是对各种原始对象数据的收藏、加工、整理、揭示并提供检索查询服务,这是图书馆界的看家本领;而我们一再宣称的“数字图书馆”对应的却是传统意义上的“情报研究中心”,其目标和工作核心是对资源内容的深度挖掘和个性化服务,需要有相应的专业基

础,这恰恰是图书馆界的弱势。以工程现有资源建设人员的知识水平和知识结构去完成数十个专业的信息增值服务,绝对是天方夜谭。

其实,无论是哪种定义,“数字化图书馆”都是不可逾越的发展阶段,也是最能满足现阶段用户需求的一种资源服务形式。全国文化资源共享工程的终极理想和目标可以是“数字图书馆”,但也必须经过“数字化图书馆”这一原始积累过程,正如同社会主义是共产主义的发展阶段一样。用户的需求并不难满足,也不管你采用了什么技术,问题是他期望中的功能需求你是否能最大限度地给予满足,用户关心的是前台的服务是否便捷、周到,并不关心后台如何实现。“不选最好的,只选最合适的”,yahoo在这方面已有深刻的教训。比如财经栏目,yahoo采用美国的组织方式,但完全不能适应中国股民的查询浏览习惯,最后导致本栏目访问量锐减。与此同时,新浪网、和讯网等网站开设的财经栏目却越来越成气候,yahoo想收回领地,非常之难。如果普遍采用尚未开始大规模推广应用的技术来组织资源,并提供服务,前期需投入大量人力、物力、财力去培育市场,甚至改造市场,其结果可能是市场被你控制,你牵着市场走,但另一个可能是市场根本不给你机会。

笔者认为,文化资源共享工程现在可以发挥自身优势,着手建设数字化图书馆,为实现终极目标奠定扎实的基础。

1.3 资源共建模式

尽管国家中心已与31个省省级分中心、14个文化部直属单位签署了试点实施协议,但几乎所有签约单位都未完全履行协议条款——尤其是在资源共建方面更是令人失望。共享工程的资源建设工作基本上是国家中心一家在独撑局面。究其原因,除各地领导重视不一、认识各异、资金有多有少等老生常谈的问题之外,还有一个原因是资源软件的使用和库结构的设计问题。资源共建的第一步可以是动员各级中心搜集、整理各地具有知识产权的资源素材,或仅仅获得使用授权,而不是过分强调按照我们的“模具”去装资源,挫伤了很多单位的积极性。国家中心应参照四川省中心调动基层中心提供素材进行简单粗加工的做法,打通原始资源素材采集的渠道,加快资源整合的步伐。

1.4 资源选题策划的切入

试验系统研制之初,为了争取各方的认同,并能形象直观地表现共享工程的建设思路、规划和目标,选题

的策划是从用户端切入,先确定栏目,再围绕栏目组织资源。这样做的优点是容易出亮点,短期内效果显著;缺点是积累的素材比较零散,容易偏离方向。为了不断出彩,必须不断策划新的栏目,致使范围越来越宽,摊子越铺越大。此外,从栏目切入很难统计数据量,其单位或者是多少条数据,或者是多少 MB、GB,而不是老百姓十分熟悉的多少册书、多少部作品、多少场讲座,而且,其他工作比如知识产权的解决等也难以配合。

共享工程实施一年多之后,应该把选题的切入点调整为策划底层的对象资源。结合文化部和财政部联合下发的文件要求,围绕百万册件文献、地方剧目、音乐作品、美术作品、珍贵文物做文章,根据实施的具体情况和用户反馈,可以增加电影、讲座的资源。由于文物局未参加,可以将文物规划适当压缩。之后,根据实际完成的对象资源情况策划用户界面的栏目和特色资源推荐。

2 今后文化信息资源建设的思路调整和操作方案

2.1 总体思路调整

进一步明确共享工程的宗旨,为资源建设指明方向,为素材选题提供依据;资源建设应紧紧围绕文化部和财政部联合下发文件中规定的资源建设目标来开展;目前资源建设的方式相对于资源深加工来说,更应注重素材的积累;资源应突出文化部管理范围内的资源特色,资源建设应以文化部直属单位的特色收藏为主要依托,从各单位文化信息资源的联合目录入手,有计划地组织对象资源的加工整合。

2.2 建设目标调整

两部委联合文件中规定的资源建设目标为:

(1)完成全国图书馆、博物馆、美术馆、艺术研究等机构的文化信息资源联合目录。

(2)完成以“百万册(件)文献共建”与“四个一优秀作品”为核心的数字资源建设,即完成 100 万册(件)文献、1000 台优秀地方剧目、1000 部优秀音乐作品、1000 部优秀美术作品、1000 件珍贵文物的数字化,并提供网上服务。

(3)整合现代贴近大众生活的社会文化信息资源。

(4)建设支持文化信息资源共建的基础信息资源。

对其中第 1 个目标,应先从文化部直属单位的联合目录做起,然后再扩展到北京市,同时以各省分中心为依托,汇总各省相关信息的联合目录,最后完成全国图书馆、美术馆、艺术研究等机构的联合目录。

对第 2 个目标,目前共享工程的许多栏目都是围绕这一目标建设的,例如“文化源流”、“哲学与宗教”、“中国历史”、“中国文学”、“地方戏曲”、“新中国美术文献”等。根据资源建设的实际情况,我们可增加 1000 部优秀电影(电视剧)和 1000 场优秀讲座的资源建设目标。但是百万册(件)的数量太过庞大,应征求专家意见,确定一个比较实际的目标;为了加强对资源建设的指导,应对各具体建设目标成立专家组,由专家对什么是文献、什么样的作品才是优秀作品及其数量等进行界定,并提出具体的素材目录。资源建设工作还应充分借助“中华再造善本工程”以及“民族民间文化保护工程”的工作成果。

对第 3 个目标,共享工程已建设了许多相关资源,如中华食文化、烟文化、酒文化、茶文化、文化旅游等,从资源数量以及资源建设经费出发,这一部分可以维持现状,不再增加新的内容。

对第 4 个目标,共享工程已建设了相当数量的资源。可请专家组再进行界定与把关。

2.3 技术支持调整

资源建设的最终目的是用户服务,因此,技术支持的重点放在所建资源的有效揭示、全方位检索与查询功能上,一切以用户方便、易用为出发点,具体表现在:

(1)现行资源加工与发布系统软件无自动标引、关联功能,在保证资源库元数据准确、完整的情况下,减少跨库关联(即减少库结构中相关字段、多行置标),从而降低资源库拆分困难与手工置标的错误率。

(2)资源库后台管理采用 XML 技术,前台因特网上的服务采用浏览器普遍支持的 Html 网页,以支持网站的信息动态发布、兼顾最大范围的用户群与先进技术的应用。

(3)构建支持多种系统与资源库的发布平台,提供具有跨库检索和站内网页内容查询功能的搜索引擎或检索框。

(4)在资源库发布模板上增加二次检索、页码跳转功能。

(5)增加资源加工界面上数据批查重、导入与输出功能,增加服务界面(国家数字文化网)上的用户访问路径显示功能。

(6)根据不同资源访问权限应用共享工程数字版权保护系统进行控制。

2.4 知识产权配合

共享资源要突出主题,围绕一个主干,即以“七

个“优秀作品”的建设为核心,即收入100万册文献、1000台优秀地方剧目、1000部优秀音乐作品、1000部优秀美术作品、1000件珍贵文物,以及1000场优秀讲座和1000部优秀电影,按此内容在专家组制定的框架下进行资源的采集,积累能经得住时间考验的精品,避免在选择素材时因为缺少框范而过于随机而导致的鱼龙混杂。其中社会方面的资源买进作为辅助的内容建设手段。

由于版权谈判中存在的困难,我们所选资源素材不可能全部谈成,可以确立四个类别的资源素材,即“必选资源”、“首选资源”、“次选资源”及“替代资源”等,进行选择。

由于资源建设目标中所确定的对象资源种类明确,因此,采集资源时针对性很强。这在资源建设资金非常紧张的情况下,能免受社会上庞杂的资源素材的干扰,突出主题,打造共享工程的特色资源库。

2.5 组织实施调整

(1) 专家组及其职责

根据调整后的资源建设目标,针对各项具体的资源建设目标,成立专家小组或与相应的专业机构合作。专家组可以包括:联合目录建设专家组;文献资源专家组;地方剧目专家组;音乐作品专家组;美术作品专家组;讲座策划专家组;电影资源专家组;用户服务策划专家组等等。这些专家组的主要职责是:制定标准、规则;框定某类型资源素材采选的原则,尽可能细化到素材目录;审定元数据结构;审定相关知识的内容;审定特定资源服务的功能设计。

(2) 共建模式

资源建设的共建模式对于省级中心来说,可根据实际情况分解资源建设目标。以讲座为例,假设国家中心和直属单位可制作640场讲座的资源,30个省级分中心共制作360场,如果在3年内完成,则每个省级分中心每年需制作4场。签订试点协议时,可明确具体上交的资源内容包括4场讲座、N部地方剧目、N册文献或地方文学,等等。在完成应上交的资源任务之后,可签订经济合同,按双方认可的价格购买省级分中心制作的符合资源建设目标的资源。

对于文化部直属单位,可以根据每个单位的主要业务主题有不同的侧重点——如国家图书馆以文献为主、中国艺术研究院以专业指导和艺术研究文献为主、中国京剧院以京剧为主、中国交响乐团以音乐作品为主、中国美术馆以美术作品为主等等。分中心协议签订时规定应上交的资源基数,之后,按照

双方议定的价格签订经济合同,购买符合资源建设目标的资源。

(3) 成果形式

共建单位提供的资源可以以三种成果形式上交:

①上交获得使用权的原始素材,如图书、光盘等;

②上交已获得使用权的素材,并进行数字化加工;

③如有资源加工基础,可按照统一的加工标准进行整库的加工。

2.6 用户服务调整

(1) 国家中心在因特网上的资源服务

资源服务界面采用两种方式:一种是资源类别界面,以原始信息类型为分类标准,如图书、期刊、图片、讲座、电影等,展示全部资源建设成果。另一种是主题资源库界面,根据新建资源特点、不同时期宣传重点,有选择地围绕主题进行资源组织与发布,形式上同于现有国家数字文化网。

两种服务界面均提供全文检索、网站地图、查询路径显示功能,保证资源库中对象数据的稳定性。

两种服务界面中的分类要科学、清晰,每一类名具有说明文字。

所有网页进行优化,加快用户的访问速度。

(2) 资源产品服务

整理不同角度分类的资源清单。

根据目标用户群特点,设计相应的产品界面风格。

2.7 加工流程调整

(1)各专家组确定总体的建设框架和标准、确定素材采选原则、细化素材目录,并审核通过元数据机构、服务功能设计;

(2)知识产权部负责按照资源素材目录采集资源素材,通过合法的途径获得资源加工和资源服务所需的各项权益;

(3)资源建设部负责组织资源的数字化加工、制作元数据、按照资源类别发布信息并提供检索、查询服务——建设维护共享工程资源类别的服务界面。

(4)产品设计部负责根据加工的对象资源,结合用户需求、社会热点、重大事件等策划不同的产品系列,归并不同的栏目,提供用户端的资源服务——建设维护主题资源库的服务界面,以及根据用户需求抽取数据形成产品。

王芬林 国家图书馆副译审。通讯地址:北京中关村南大街33号。邮编100081。(来稿时间:2003-10-24)