

地区性文献信息资源共享和联合服务的可持续发展研究

李苗苗, 张薇, 李鹏, 任佳妮, 周立秋
(陕西省科学技术信息研究所, 西安 710054)

摘要: 大数据时代的数字科研环境下, 文献信息资源共享和联合服务即为知识链接和知识组织下的知识内容服务, 公共图书馆、高校图书馆、情报(信息)机构以往建立的共享和服务平台渐显出不适应性。文章通过分析知识链接和文献信息资源共享及其服务之间的关系, 结合地区性文献信息资源共享及其联合服务在整合中具有优势的特点, 提出了其务实建设的“所建即有用, 所建即好用, 所建即可持续用”的“三用”理念, 并以陕西省为例, 尝试分析了省内已有的文献信息资源和共享服务实际工作可持续发展技术要点, 对再整合改造技术的选择、构建框架等方面进行了研究摸索。

关键词: 数字科研环境; 地区性文献信息资源; 联合服务; 知识链接; 知识组织; 知识服务
中图分类号: G252, G254
DOI: 10.3772/j.issn.1673—2286.2014.10.009

互联网迅猛发展和大数据时代效应下, 数字科研环境改变了以往的科研行为和学术、成果交流模式。知识资源的整合和开放共享, 知识链接和知识组织下的高度集成的知识元链接关联又使得文献信息资源服务走向知识内容服务模式。科研、教育面对不同格式类型的文献信息资源, 重叠交叉的各种文献信息共享服务平台, 层出不穷的新信息技术, 如何有效、全面、准确地获取知识集合是其需求所在。既存的共享服务平台陆续出现了对数据科研环境和用户需求变化的不适应性, 信息技术停滞加上服务模式观念陈旧造成的使用率低等问题比较突出, 亟待可持续性再发展。地区性文献信息资源共享和服务作为我国信息服务发展中的重要组成部分, 应响应国家当前对文献信息服务业的发展要求, 对系统进行改造, 推动文献信息资源再整合和联合服务的深层次集成化, 完成数字科研环境下的知识内容服务的模式转变。

1 地区性文献信息资源共享和联合服务概况

1.1 现状和问题

自20世纪90年代中期伊始, 以图书联盟为开端, 我

国开始了文献信息资源共享和服务建设。地区性文献信息资源共享和联合服务项目不断获得立项。随着计算机和信息通讯技术网络服务下的虚拟信息服务的产生, 数字数据存储、图形化、互联网链接、个人服务终端设备等关键点技术不断改变着世界信息交流服务模式, 加之大数据概念的提出, 科研环境也随之改变。如果文献信息资源服务不具备可持续发展性, 必然会造成前期建设资源的荒废、使用机能低等问题。随着持续发展意识的增强, 国家及地方开始对各类文献信息资源共享平台开始升级。其实质想要达到的效果就是统一机制下的同类知识机构的多个知识资源的动态无缝知识链接, 在相同标准协议的唯一服务平台上提供便捷而又行之有效的知识内容服务。目前大多数平台还处于散点分布、各自为营的状况, 资源和服务未得到深度整合, 管理未有统筹整合安排。

现存地区层面的文献信息资源共享和联合服务的问题主要有以下三方面:

(1) 数字科研环境的消化不良。数字科研环境带动起来的是密集型科研数据和开放的一站式集成服务模式, 单个的文献信息资源及服务是无法实现的, 而现存的地区层面的文献信息资源共享和服务仍处于分散和交叉重叠状况, 呈现对数字科研环境的消化不良病恙。

(2) “为项目而项目、为职称而项目、为业绩而项目”。个别虚空平台项目因技术停滞, 知识链接和知识服务构建浅, 跟不上数字科研环境下的新信息技术要求, 无法适应用户需求和行为的变化, 使用率低, 建立的数据资源荒废, 极大地浪费了社会资源。

(3) 区域圈化无统筹。系统的地区分布重复、小而繁密的, 并无太大作用和意义的小地区和个别机构项目合并入上一级建设, 也就是“范德华力”作用下的小分子合并成大分子。

针对以上, 推进地区联合服务作用于数字科研教育发展, 转变服务模式为一站式集成知识服务; 整合数据为可链接的知识元; 选择统一框架和标准的系统改造技术; 融入个人智能电子移动终端设备等新信息技术, 目标朝向知识内容服务和用户体验, 实实在在将已构建的共享知识资源价值体现和持续下去。将习近平主席的朴素、认真、严谨、不浪费、求实的发展精神融会, 按照“所建即有用, 好用, 可持续用”的“三用”理念朴实地提供知识服务, 避免虚大空项目, 为国家数字科研环境下长远的信息资源服务开放共享建设工作打好数据和技术基础, 做好“小分子”的角色。

1.2 可持续发展的优势

往往越简单的原理越能反映出事物间最本质的关联。通过对知识链接几个发展阶段的形态观察, 知识链接向下一阶段发展的效应力和物理学中的范德华吸引力有异曲同工之处。范德华力是存在于组成和结构相似的物质分子间的一种吸引力, 且一般随着相对分子质量的增大而增强, 相当于是实体物质间的整合, 进而通过整合的表现形态提供一种行为, 就好比是同一个玻璃平面上的水滴, 在范德华力作用下, 不断吸引汇融, 形成水泽, 进而可以折射出比水滴更广的外部景物影像。而知识链接的发展实质是虚拟物质间的关联体资源整合以及整合后更深层次的知识元链接表征, 不断发展出可以链接的更多关联节点, 以此提供知识集合。

文献信息资源和服务的整合从用户需求、学科类别、行业领域、地理因素等出发点, 归根到底都是遵循“相近、相似”原理。我国文献信息资源服务类型按区域划分有全国性、地区性两种, 随着互联网和通讯技术进步, 空间距离和地缘因素变得越来越模糊化, 但仍具有不可忽略的影响。地缘因素能够不同程度上强化其他的“相近、相似”特征。对于普遍的文献信息机构而

言, 地理因素结盟仍具有相当的吸引力, 无论从地区性合作寻求的优势性, 还是共同的上级部门和可能的资金来源, 或是相同的社会形态和文化背景来说, 其共性需求相近, 更容易达成共识和合作, 类似“范德华力”的效应愈是突出。同一地区内的文献信息资源的整合和联合服务具有其客观优势, 无论从地缘因素, 还是文化背景、用户需求、认知契合和使用度来说, 都更相近、相似, 则更容易吸引融合。这就是类似范德华力原理在文献信息服务的虚拟知识体之间的表现。另外同一地区内的文献信息资源的整合和联合服务在技术交流、合作沟通和管理培训等方面实效性强, 方便快捷, 项目再建设的可操作性和可实现性好。因此, 在综合考虑新信息技术与之匹配的数字信息资源建设情况, 利用已建立成熟的且具备一定量使用用户储备的地区内文献信息资源共享服务平台, 通过消除系统间、数字资源间的异构限制; 应用合乎成本效益、稳定且可以随着数据信息资源量变化而能后期灵活改动的系统技术整合文献信息资源; 统一检索引擎的集成检索等方法形成唯一门户的联合服务, 在根基上使资源成本得到统筹安排, 争取社会效益趋向最大化, 达到可持续发展的目的。

2 地区性文献信息资源共享和联合服务的可持续发展导向

地区性文献信息资源共享和联合服务是我国信息服务的组成部分, 因此和文献信息资源共享和联合服务整体所处的环境及可持续发展导向是一致的。

当下的数字科研环境具有基础数字数据信息密集、海量数据分析、开放式学术交流模式这三个特征。在这个大环境下, 科研的新亮点行为包括数据基础设施可持续发展和数据专员队伍完善。文献信息资源共享和联合服务的可持续发展则属于数据基础设施范畴, 通过数据基础设施科研的具体活动[可概括为数据抓取(Capture)、分类处理(Curation)和数据分析(Analysis)]最终将分散孤立的研究数据转变成相互关联的研究资源。而这一行为实际就是数字资源的知识组织和知识链接, 数字科研环境下的文献信息资源整合和联合服务于是可以理解为知识链接和知识组织(在这里, 本文所指的知识组织概念是狭义上的, 也就是指将知识加以描述、加工、形成系统的环节^[1])下的知识服务, 即为在传统收集、加工、检索、描述的基础

层面上对知识的一次或多次的提取和利用,并通过再组织关联点后寻求延伸文献信息资源,且有可能分析获取价值情报的过程。

可以看出,知识链接是文献信息资源整合和联合服务的本质核心。从辩证唯物主义方法来说,知识链接发展导向用户需求和行为的变化,反之用户的需求和行为变化又是知识链接发展的动因。铺开知识网后再到关注节点这样一个伴随用户需求和研究行为不断重复和探索的动作就是现代新信息技术行为下的文献信息资源服务,并且这种行为是具有可回溯性的,文献信息资源服务和知识链接便是这种动态相关联的发展关系。

表1为按知识链接发展阶段为顺序的代表性文献信息资源共享和服务: Dialog为代表的自体数据库查询检索的知识链接; Dialog为代表的自体数据库查询检索的知识链接; Springer link等数据库群的跨库知识链接; CALLS的同类型知识机构的知识链接; OpenURL/SFX的跨库、跨机构的知识链接。类似的检索服务系统还有有国家科技图书文献中心(NSTL)、ENCompass、MetaLib、NLM Gateway、CALIS统一检索系统、CNKI跨库检索系统、CSDL跨库集成检索系统等。

从表中不难看出知识关联的形态整体趋向于开放型的存储和共享模式,文献信息资源的整合和服务发展是基于实际环境中的知识链接特征,是建立在统一管理和技术规程的开放性的动态知识链接;服务导向于用户需求和体验,在便捷的检索和获取服务之后用户更需要能够解答实际问题的内容服务^[2]。这与大数据概念是为解决问题而提出的数据资源服务理念相吻合。

通过以上分析,依据数字科研对数据的聚集和分析需求,同时学习国外数据服务机构先进经验,相应地

区域性文献信息资源共享和联合服务的可持续发展的主要导向点可归纳为:用户需求和行为的变化;文献信息资源格式和种类的统一标准下整合。

另外要注意到的一点是知识链接的几个发展阶段是可以相互并行发展的,由此也说明了地区性文献信息资源共享和联合服务可持续发展具有兼容性、可行性。

2.1 用户需求和行为的变化

文献信息资源共享和联合服务中,用户的需求和行为即为人与知识体之间的知识链接,是一个主动信息查询行为获取知识和利用知识的过程。国外学者提出人类获取网络信息共有三种模式,即浏览模式(browsing)、搜索模式(searching)和询问模式(asking)^[1]。数字科研环境下,文献信息资源共享和联合服务可持续发展的目标便是可以达到最为复杂的询问模式,最终让用户可以根据自主提问的复杂度、主题位置、搜索栏和主题的明确性,附加自身知识来灵活选择浏览模式、搜索模式和询问模式中的一种,即为用户需求和行为变化的导向所在。

“所见即所得”是为用户提供的理想化的未来文献信息资源共享服务目标。

宏观角度来说信息服务中用户需求和行为变化的三个因素是:计算机信息技术发展、数据密集程度和知识链接形态,且用户需求和行为与这几个因素之间是相互交叉影响作用的。从图1中可以观察出现今用户的需求和行为变化特征大致有几点:

- (1) 文献信息资源和共享联合服务的物理空间和设施趋向一站式、自助式、多元化;
- (2) 更趋向于深层次的学科化、专业化、个性化需求;

表1 知识链接不同发展阶段和数字信息资源情况

阶段	知识链接范围	知识链接形式	跨库	跨机构	用户需考虑资源位置	用户需考虑资源归属	代表性信息资源
1	单个知识资源	自体数据库查询检索的知识链接			√	√	Dialog
2	单个知识机构内 多个知识资源	数据库群的跨库知识或知识元链接	√		√	√	Springer link; CNKI Elsevier Science Direct
3	联合服务系统内多个同类 知识机构的多个知识资源	相似知识机构共建共享服务系统 内的知识资源的知识链接	√	√	√	√	CALLS; NSTL; MetaLib; NSTL NLM Gateway; ENCompass
4	统一机制下的同类知识 机构的多个知识资源	多个机构的多个知识资源间统一且唯一的检索入口,实现用户知识需求的直接递交和对用户有用的知识链接的直接提供,用户不必再进行选库等动作,直接由服务系统完成	√	√			OpenURL/SFX

(3) 个人智能移动终端设备(平板电脑、智能手机、谷歌眼镜、智能手表、车载电脑等)被越来越多使

用, 移动数字化服务和网络化虚拟化服务的兼容。

结合图2、图3, 可以看到手机网民比例逐年递增, 而在电脑和手机使用频率最高的“了解感兴趣的信息”和“了解工作学习相关内容时”的两项行为统计中, 手机使用比例均高于电脑, 这是个人需求和行为变化的一个趋势表现。另外特别要提出智能手机和平板电脑在数字科研使用工具中所占比例增幅较高, 国家图书文献共享服务中心(NSTL)也关注到这点, 目前推出了安卓版手机客户端, 而我们大部分的地区性文献信息服务都还未完善这项无线虚拟服务。平板电脑、手机等个人智能移动终端设备越来越在生活中普及, 百度文库、豆丁书房、方正阿帕比电子书(Apabi Reader)等APP程序的应用, 均提供开放性的无线虚拟文献资源服务。数字网络化带来的是信息和数据爆炸, 遂导致了数字科研环境产生, 加大了虚拟化的文献信息中心和文献信息服务需求。虚拟技术依托于计算机、信息通讯技术, 而计算机、信息通讯技术是高速发展的产业, 新技术、新设备的更新换代迅速, 这些便迫使文献信息资源整合和服务的模式需要不断适应和发展, 不接轨便会导致平台的生命力枯竭。

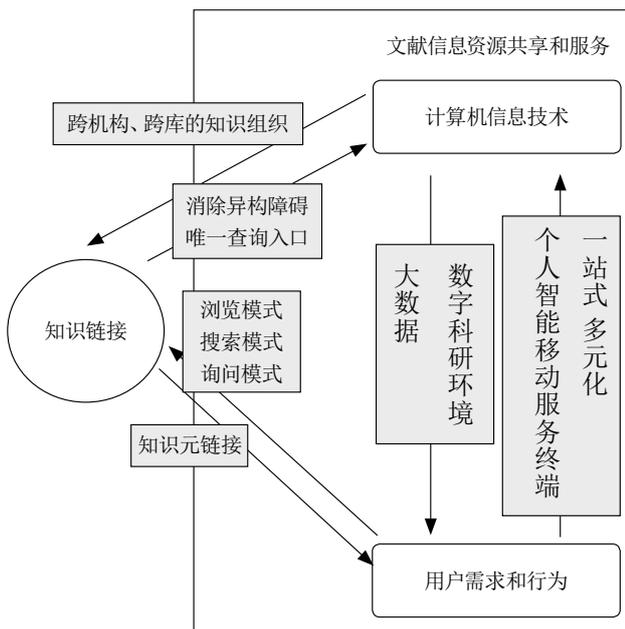


图1 用户需求及其行为和知识链接发展的相互作用关系

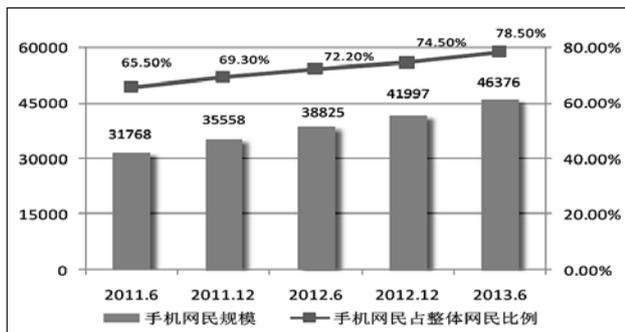


图2 中国手机网民占全部网民比例

(数据来源: 中国互联网络发展状况统计调查^[31])

另外, 包括下一代通信网络、物联网、三网融合、新型平板显示、高性能集成电路和以云计算为代表的高端软件的新一代信息技术中, 通信网络、平板、云计算、无线虚拟文献资源服务都是在分析用户需求和行为的变化中需要进一步关注的研究内容。

2.2 统一标准下文献信息资源格式和种类的调整

文献信息资源的联合服务实际上也可以认为是文献信息资源的深度整合及其服务, 亦涵盖了开放存取层面的发展研究。根据对文献信息资源整合的内涵和实现形式的理解, 深度整合资源大致包含了图书、期刊和信息存档。因此文献信息资源的整合模式可归纳为知识库整合、图书和期刊整合、搜索引擎整合、数字信息资源管理系统整合四种。概括而言, 即为文献数据和系统的整合。

从图4可以看到数据和文献之间的关联关系, 文献数据和原始数据分别位于顶部和底部, 中间是文献数据和原始数据的关联数据层。数字科研环境下, 这三部分既是研究内容的整体, 又是文献信息资源服务的内容整体。数据的结构、分类、存储等经过处理和整合转换, 通过跨数据基础设施的整合, 在服务平台上

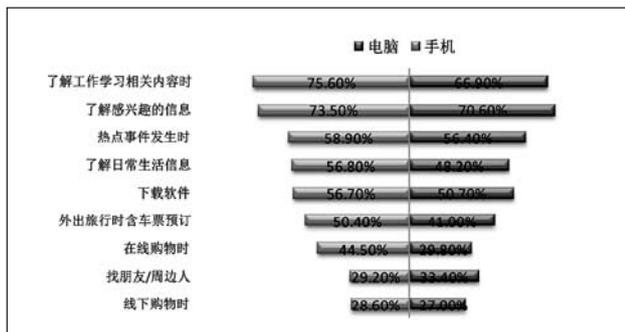


图3 2013年中国网民行为调查

(数据来源: 中国互联网络信息中心《2013年中国网民搜索行为研究报告》^[31])

搜索论文或科研数据,能得到原始数据,也能得到与之相关联的其他知识资源。例如用户通过检索到的论文,得到数据,进而还可以进行数据分析,再得到与其分析数据相关的文献。要实现这样的动态无缝知识链接,数据必须经过分类处理和整合转换才能够保存和进行分析利用,实现在统一协议下的处理。

文献信息资源整合的异构性消除必须有统一的架构和系统作保证,是具备可操作性和可发展性的关键点。

3 以陕西省为例的地区性文献信息资源的可持续发展方案研究

2004年陕西省基础条件平台开始起步建设,目前,陕西省科技图书文献共享服务平台已具有相当规模,除此外有省级层面的陕西省文献资源共享与服务平台、陕西省公共图书馆图书服务联盟等^[4,5]。另外,陕西省内的文献信息资源中,陕西省科技信息研究所作为省级情报信息机构,还拥有科技查新报告库、科技咨询报告库、科技成果库、科技档案库、科技报告库等特色科技数据资源,这便为陕西省文献信息资源的共享和联合服务提供了数据基础。

通过表2可以看出陕西省公共图书馆、高校图书馆、情报(信息)三大体系的文献信息资源的类型、分布、共享服务模式的概貌。

通过表2可以了解到陕西省文献信息资源共享和服务没有一致的标准和开发技术环境,新信息技术下的知识服务能力薄弱,仍处于知识链接的较早发展阶段,重复建设、资源重叠,缺乏整合和可持续发展性,和现今的数字

科研环境脱节,无法充分体现其信息数据资源的价值。

3.1 框架

文献信息服务的重点在知识的组织、分类与用户的行为趋势,单独的机构知识库服务已无法满足科研的需要,必须集成到整个数字科研环境中或某个特定区域内的信息系统才能充分发挥资源的作用。因此整体框架设计旨在通过以DSpace为核心基础开源技术推动图书文献机构资源和服务之间的互动和整合,跨越机构物理界限,构建统一标准下的联合体式虚拟图书文献服务平台,使陕西省内主要文献信息机构的资源和服务深层次整合,为用户提供“一站式”的信息服务和集成知识环境,即在开源软件DSpace的基础上设计和改造的知识服务平台。

整体设计沿用DSpace的存储层、业务逻辑层和应用层3层架构模式,通过分层封装和调用,满足二次开发时在相应层中进行扩展的需求。将对应机构知识库资源分类,通过Web Services封装将各分机构知识库嵌入整个文献信息资源联合服务平台中,对机构库的资源服务进行集成,利用DSpace和SRU接口共同完成检索服务,将检索嵌入集成一体化,形成了文献信息类型的数据描述模板。

3.2 技术选择

3.2.1 选择考虑因素

最高阶段的文献信息资源整合和共享服务形态为不

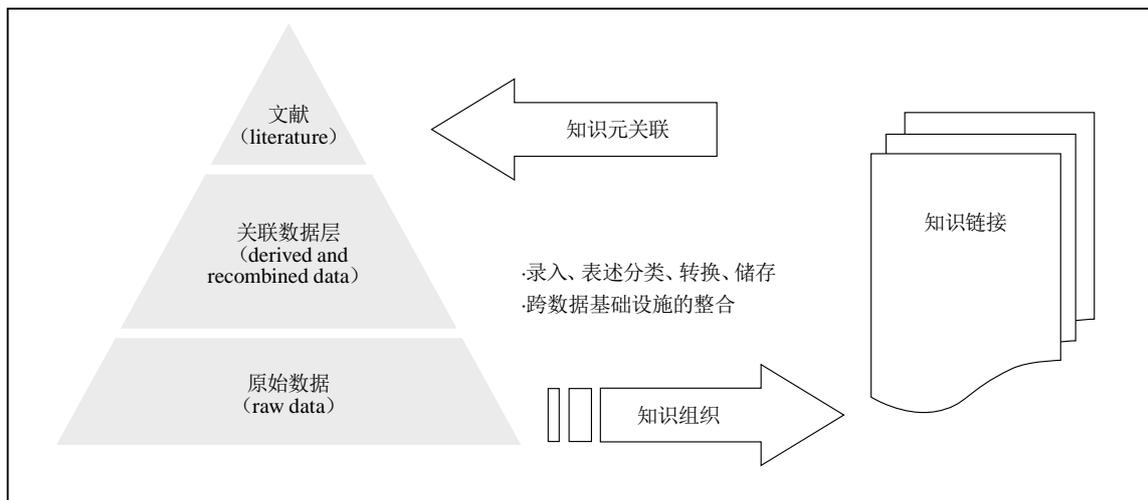


图4 数字科研环境下的文献信息资源服务内容

同类型、不同格式、分布式信息资源的无缝动态链接^[6], 结合陕西省文献信息服务开展的实际工作, 技术选择时考察的因素大致有: 能实现跨库、跨机构知识链接; 易于扩展; 支持多种数据类型; 便于索引; 统一标准;

可提供唯一查询入口(即统一的检索引擎)等, 因此具备开放性、对旧系统的易改造性、灵活的可再开发性等特点的开源技术便显示出了其优势, 并且需要此项技术服务的文献信息资源机构均可免费获取并自由使用。

表2 陕西省文献共享服务系统情况

名称	包含机构	文献信息资源类型	服务模式	初建时间
陕西省科技文献图书共享服务系统	211工程部分高校图书馆、情报(信息)机构	科技文献信息数据、电子期刊、部分高校图书馆馆藏文献	计算机文献书目数据; 原文传递、代查代借	2004
陕西省文献资源共享与服务平台	陕西省高校图书馆	陕西省各高校图书馆纸质图书、电子图书、电子期刊、学位论文、会议论文、报纸、标准、专利、视频、音频、自建数据库以及网页等	计算机文献书目数据; 原文传递、代查代借	2009
陕西省公共图书馆图书服务联盟	陕西省内公共图书馆	陕西省内公共图书馆文献资源	计算机文献书目数据; 原文传递、代查代借	2010

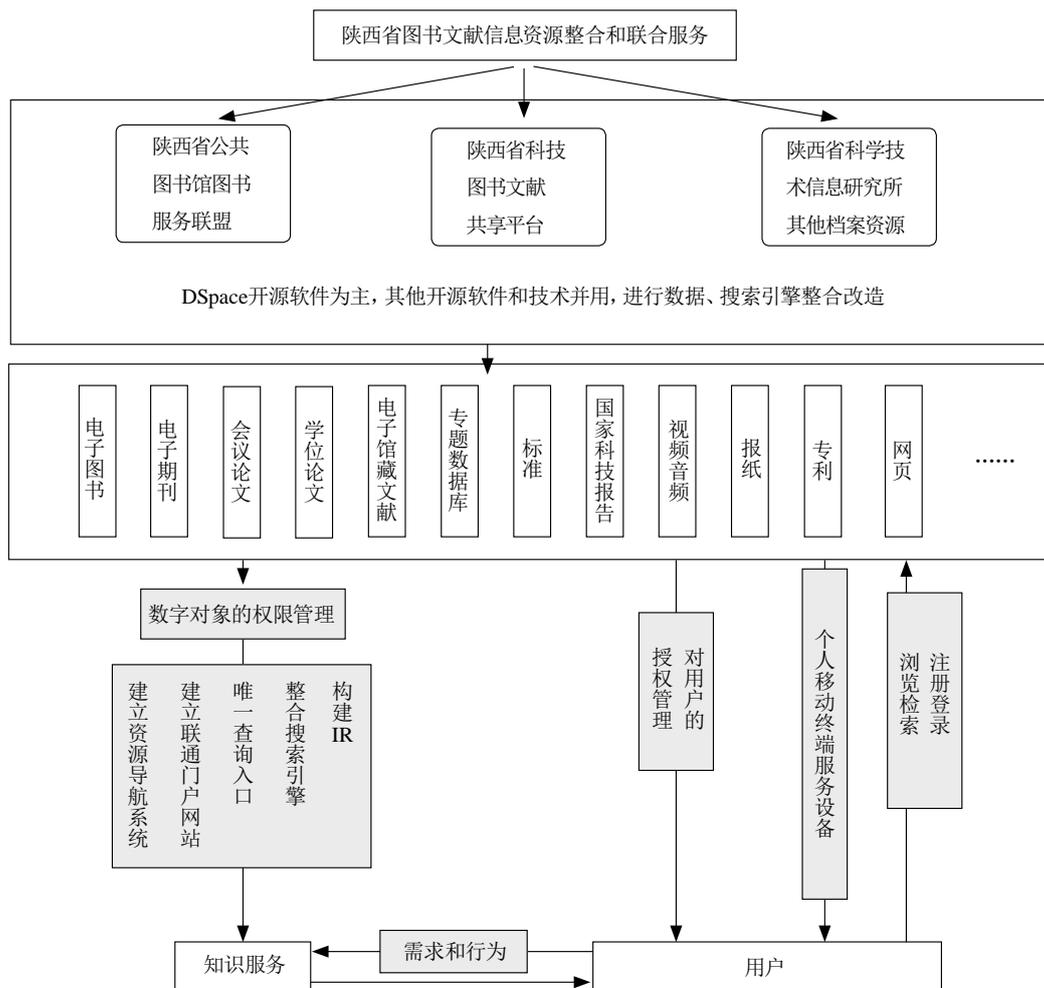


图5 陕西省文献信息资源共享和联合服务框架

目前开源软件应用研究较多集中在数字图书馆的数字集成建设和服务系统的二次开发方面, 我们可以看到, 开源技术的应用不仅于此, 还为同样服务本质的文献信息平台的整合及再开发打下了良好的基础, 提供了宝贵的实践经验。

3.2.2 开源技术筛选

中科院的科研人员将可以用于数字图书整合建设的开源软件大致分为门户类开源软件、机构仓储类开源软件、内容管理类开源软件和知识发现类开源软件等四类。表3列举了目前文献信息机构使用较多的开源技术软件。

面向服务结构 (Service-Oriented Architecture, SOA) 和随后大数据概念的提出, 表明数据和服务已经有机融合在了一起。图书文献的价值终究是要以某种信息服务来表现的, 反之信息服务的基础是大量的数据。在新型信息环境下, 图书文献机构传递信息的服务模式也在发生着改变, 开放数据和开放服务成为主流趋势。

从单个图书文献机构应用开源软件技术来说, 例如清华大学图书馆确定以开源软件Fedora为技术平台, 开发数字资源管理和长期保存系统; 浙江大学图书馆也已经开始尝试开发以DSpace为核心的数字图书馆系

统。因各自功能需求和业务模式的千差万别, 难以聚集大规模的用户群, 自身的可持续发展性受限; 从系统的角度看, 开源软件的关键是开放结构, 并不是单纯的开放代码。整个软件架构是开放的, 才能形成共享开发、联合应用服务的环境, 实现开源软件真正意义上的生命周期延续性。文献信息资源服务不单单是要追求开放代码, 根本上是要用开源软件的模式来开放数据和开放服务。

作为专门的数字资源管理系统 (数字空间) 的DSpace亦是知识内容管理平台类的研究型知识产品, 具有开放源代码社区的可持续服务保障; 数字科研环境下文献和学术资源具有长期互动存储等特点, 同时DSpace在我国基础性较好, 是国内目前各个文献机构使用较多的系统, 其数据模型更接近机构本身的特点, 易于在研究机构中使用。这里要注意的是Dspace还同时支持OAI协议和OpenURL协议, 而国外目前将近20个信息提供机构采用了OpenURL/SFX系统。基于国情和国际技术形势综合考虑, DSpace既是基于国内文献信息服务发展基础, 又满足了和国外的文献信息数据服务的接入性, 使整体数字科研环境下的数据也较为容易梳理、形成服务集合。另外, 陕西省文献信息资源共享和联合服务组建包括了多个现运行平台, DSpace系统则可以满足多平台运行改造的技术要求, 机构可通过直接改写

表3 目前使用较多的开源系统

名称	类型	技术特点及其应用机构
Koha	自动管理软件包	世界上第一个开放源码的图书馆自动化系统。Linux系统, 使用MySQL开源数据库, 根据不同类型和规模的图书文献机构, 提供分类检索、用户管理、收购和发行管理等功能。
DSpace	专门的数字资源管理系统 (数字空间)	由麻省理工学院图书馆和美国惠普公司实验室合作研究开发, 采用Java语言, 运行于UNIX系统, 可用于包括分布式唯一标识科研数据在内的不同格式的数据存储、处理、索引后的重新再发布。Dspace开源软件目前在我国使用较为普遍, 例如清华大学、中国西部环境及生态数据中心、中国国家图书馆、厦门大学图书馆、浙江大学图书馆、香港大学图书馆等相关文献机构均采用了DSpace系统进行二次开发。
Greenstone	内容管理类系统	基于Unicode编码多语言的开源软件, 采用JAVA技术, 可管理多个数字化的知识仓储。知识仓储结构和其配置文件对应, 可用于不同格式的文档、口述记录、歌曲等数据和内容处理。
Fedora	数字存储管理项目	由数字对象和仓储两大实体组成, 提供了集成数据和元数据, 对内容操作的框架、数据处理和对象链接接口, 可应用于文献信息机构数字仓储。
Scriblio	多融合系统	带有基于WordPress的面搜索和浏览功能, 检索界面简单, 单点获取数据信息, 可对结果集进行相关度排序和分面导航, 具有Web抓取内容功能; 可应用于下一代图书馆发现与获取界面搭建。成熟系统有Aquabrowser、Encore、Primo、Worldcat Local和Polaris Library System等。目前基于Scriblio架设OPAC系统的图书馆主要分布在国外, 有美国的库克记忆图书、Beyond Brown Paper、波斯顿大学神学院的传教历史资料库、奥地利维也纳医科大学图书馆等。

DSpace, 搭建具有新型陕西发展特色的文献信息资源类机构仓储系统, 并在各个成员机构间实现资源联网共享。因此, 基于总体平台建设经费和文献资源管理的特点, 可以优先考虑应用DSpace开源技术为主、其他技术为辅进行陕西省文献信息资源联合服务的可持续发展。

3.3 数据描述

陕西省文献信息资源共享和联合服务的可持续发展包括历史数据和新建元数据的规范化标引、网站风格统一化、网站栏目丰富化、新增专题数据库的集成与发布, 构建资源导航系统; 通过搜索引擎整合, 构建统一检索平台, 建立可联通的门户网站; 构建IR来实现对文献信息资源的全面管理等, 最终目标为实现公共图书馆、高校图书馆及省级科技情报机构这三大块文献信息资源的深度整合。而统一标准的数据存储、描述和处理是深度整合的重要基础。

DSpace包括Community (组织)、Collection (馆藏集合)、Item (信息对象)、Bundle (数据包)、Bitstream (位流数据) 这几种存贮数字资源类型。Community (组织) 为最大集合, 由Collection组成; 而每个Collection又是由多个Item集合而成, 在每个Item中包含多个Bundle, Bundle由多个Bitstream组成, 每个Bitstream中包含相关联的格式^[7]。这种数字存储形式的特点是分级灵活、适应性强, 可满足各个成员机构的

不同数据库共享改造。

该联合服务中的图书文献信息资源包括陕西省图书馆、西安图书馆等各级图书机构的计算机文献书目数据; 陕西省科技信息研究所、西安交通大学图书馆、西安电子科技大学图书馆等机构的中文期刊、干旱半干旱地区主要经济林、黄河水文、高校学位论文、中文图书书目、国外会议论文、外文期刊、国外科技报告、外文图书书目等电子数据。

在图书文献信息资源联合服务改造中, 各个成员机构所拥有的数据库等同于DSpace中的Community (组织), 在每个数据库下又分为科技期刊、学位论文、国内外专利、国内外标准、科技成果、会议论文等小类, 这些小类即为Collection (馆藏集合)。文献信息资源的再分类相当于Item, 而对文献信息资源进行标识的数据描述等同于Bundle, 文献信息的发布日期格式、标准页数等相当于Bitstream。

例如, 专利文献数据和DSpace类型对象的对应关系如下所示:

Dspace对象类型	知识仓储对应对象
·Community (组织)	装备制造数据库
·Collection (馆藏集合)	国内外专利
·Item (信息对象)	中国专利
·Bundle (数据包)	中国专利中的专利号、 公开日期等
·Bitstream (位流数据)	专利号格式、日期格式等

3.4 较之前文献信息服务功能增加的技术点内容

在文献信息资源整合方面: DSpace对文献信息资源机构各种数字信息资源的管理基本可以涵盖所有数字化格式, 图书、期刊文献、专利、标准等由DSpace管理后, 构成文献信息知识仓储的组建条目。而DSpace所包含的Handle (CNRI) 为文献信息、机构仓储、资源成员机构的唯一永久性标识符, 解决了文献信息资源共享和联合服务数据因内容或物理位置改变而丢失知识链接的问题。

文献信息资源联合服务方面: DSpace同时支持OAI协议和OpenURL协议, 满足了文献信息资源共享和联合服务中用户获取知识链接的检索需求, 可以实现浏览、搜索和询问。浏览 (浏览模式) 提供了按题名、作者、出版日期浏览整合数据 (DSpace存储库) 中全部文献信息 (Item) 的功能, 也提供了选择相应的文献信息资源机构 (Community) 或文献信息资源

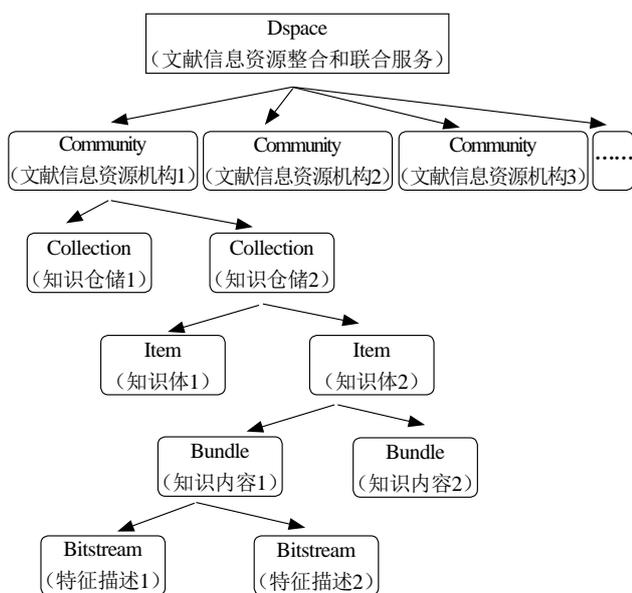


图6 DSpace的文献数据资源整合和组织模型

机构中某个知识仓储 (Collection) 下所有文献信息 (Item) 浏览的功能。具体文献信息可以本地下载或打开Web浏览器浏览。检索 (搜索模式和询问模式) 提供了API与免费搜索引擎Lucene^[8,9]检索、检索栏中输入检索词的简单检索和指定相应的知识仓储和DC元数据字段高级检索。

用户对象: DSpace系统中的用户可以是使用人,也可以是计算机,同时实现了知识链接的人和知识体、知识体和知识体之间的链接。

内容服务: DSpace具备日志型统计功能,可以提供文献信息的热点追踪、新技术预见等服务,通过兴趣订阅 (subscribe), 就可以得到专题的最新情报和文献资料。

3.5 待思考的问题

无论是DSpace系统还是兼容的OpenURL/SFX系统, 开源软件和无缝动态知识链接技术都发展了十年以上, 在国外也还仍处于摸索改进阶段, 其提供的文献服务质量问题逐渐明显, 基于我国文献信息资源基础数据建设和信息技术比国外薄弱的实际情况, 其动态链接不稳定、引文链接到获取全文过程存在相当失败率等问题会更加突出, 涉及相匹配的数字资源、管理制度、知识产权等方面的建设发展也有待后续进一步的研究。

4 结语

地区性文献信息资源联合服务可持续发展的总

路线思想可概括为数字化 (Digitalization)、开放化 (Opening)、结构化 (Structuralization) 和高度融合 (Integration), 我们的工作是为之实现而不断推进, 贴合数字科研环境下信息服务的发展需求。因此, 理论和实践还需更一进步的操作后进行研究。

参考文献

- [1] 周晓英. 知识链接和知识组织研究[C]. 第十届海峡两岸图书资讯学术研讨会论文集, 2010: 82-86.
- [2] 贺德方. 基于语义的数字图书馆服务创新[J]. 情报学报, 2012, 31(10): 1012-1017.
- [3] 中国互联网络信息中心. 2013年中国网民搜索行为研究报告[R].
- [4] 张薇, 孙滔. 网络环境下地方科技文献资源共享的实践: 陕西省科技文献资源共享服务系统分析[J]. 情报杂志, 2005, 24(6): 12-13, 11.
- [5] 李苗苗, 周立秋, 等. 陕西省区域一体化与信息资源管理研究[J]. 技术与创新管理, 2012, 33(6): 670-673.
- [6] CHANDLER A, WILEY G, LEBLANC J. Towards transparent and scalable OpenURL quality metrics [J]. D-Lib Magazine, 2011(17): 3-4.
- [7] SMITH M K, BASS M, et al. Dspace: an open source dynamic digital repository [J]. D-Lib Magazine, 2003, 9(1).
- [8] Apache Lucene [EB/OL]. [2014-07-11]. <http://lucene.apache.org/java/docs/index.html>.
- [9] GLASSER S. Broken links and failed access: how KBART, IOTA, and PIE-J can help [J]. Library Resources & Technical Services, 2012(1): 14-23.

作者简介

李苗苗, 女, 1981年生, 陕西省科学技术信息研究所助理研究员, E-mail: 30758215@qq.com.

Sustainable Development of Regional Literature Resources under the Digital Environment

LI MiaoMiao, ZHANG Wei, LI Peng, REN JiaNi, ZHOU LiQiu
(Shaanxi Provincial Science and Technology Information Institute, Xi'an 710054, China)

Abstract: In the digital environment, the sharing of the literature information resources and service are the service of knowledge and knowledge organization under the link. Public libraries, University Libraries, and information organizations are used to establish the sharing and service platform with the age of big data effect, which highlights the adaptability. Through the analysis of relationship between knowledge sharing and service links and information resources, this paper explains the necessity of sustainable construction and joint service in the area of category to integrate the advantages of the literature information resources sharing, and puts forward the literature information resources sharing, joint service and pragmatic construction of "built as useful, easy to use and sustainable." Taking Shaanxi Province as an example, combined with the existing literature information resources and sharing service of practical work, we discuss the points of sustainable construction technology of reintegration, transformation technology choice, construction framework, service mode, etc.

Keywords: Digital research environment; Regional literature information resource; Joint service; Knowledge link; Knowledge organization; Knowledge service

(收稿日期: 2014-10-15)