

数字图书馆协同服务技术与实践 ——以国家科学图书馆为例

李春旺

(国家科学图书馆 北京 100190)

〔摘要〕 在介绍协同服务需求背景、发展特点基础上,以国家科学图书馆为例,重点介绍几种协同服务的关键技术与建设经验,包括面向用户工作流的协同服务、面向知识聚合的战略情报集成研讨服务、面向融汇服务环境的协同服务等。最后,分析数字图书馆协同服务技术发展趋势以及协同服务未来建设重点。

〔关键词〕 协同服务; 技术; 实践; 国家科学图书馆

Technology and Practice of Collaborative Services in Digital Libraries——taking National Science Library as an example LI

Chun - wang, National Science Library, Chinese Academy of Sciences, Beijing 100190, China

〔Abstract〕 Based on an analysis of background and situation of collaborative services in digital libraries, the paper introduces collaborative services in the National Science Digital Library, including user workflow oriented collaborative services, knowledge aggregation oriented collaborative services and mashup oriented collaborative services. Finally, the author analyzes trends of key technologies and directions of collaborative service development.

〔Keywords〕 Collaborative service; Technology; Practice; National Science Library

1 引言

随着科学问题的复杂化发展,用户对信息服务的需求也呈现跨学科、跨领域、跨组织机构的特点,相应地数字图书馆建设重点正在从个体产品服务转向跨机构、跨领域协同服务。国际图联 2008 年报告^[1]表明,建立区域、国家、洲际以及全球性协同服务的数字图书馆是当前数字图书馆建设的一个重要特征。为了推动数字图书馆协同服务建设,美国博物馆和图书馆协会(Institute of Museum and

Library Services, IMLS) 于 2008 年创立了“国际战略合作项目”,现已有 12 万家图书馆、1.7 万家博物馆成为联盟成员。IMLS 积极推动协同服务建设与研究工作,人权数字图书馆(CRDL, <http://www.civilrightslibrary.org/>) 是 IMLS 资助项目之一,它围绕人权研究这个主题,集成了多家图书馆、档案馆、博物馆收藏的相关文献资源,多家广播、电影、电视公司收藏的影音资料以及网络上可开放获取的故事、评论、人物介绍、多媒体信息等资源。同时,CRDL 建立了来自多家机构、不同领域的专业人员组成的协同服务团队,在相关信息技术支持下,向科研用户提供专业化服务。

与 IMLS 目标一致,国家科技图书文献中心(NSTL)也在积极推动国内数字图书馆协同服务建设。其中,作为 NSTL 成员馆之一,国家科学图书

〔收稿日期〕 2009 - 12 - 30

〔作者简介〕 李春旺,博士,信息系统部副主任,研究员,硕士研究生导师,发表论文 40 余篇。

馆(以下简称“国科图”)在协同服务技术研究与服务建设方面做了很多探索,本文将结合国家科学数字图书馆建设实践,介绍协同服务相关技术与方法。

2 面向用户工作流的协同服务

2.1 协同服务平台建设

通常,科研用户在使用数字图书馆服务时主要关心以下几个问题:如何发现自己所需要的资源、如何确定这些资源的收藏位置、如何获取到这些资源以及遇到问题时如何获得咨询解答服务等。为了回答以上问题,国家科学图书馆建立了由集成检索服务(Search)、情景敏感定位服务(Location)、协同支持的获取服务(Access)和咨询服务(Reference)等组成的集成服务平台,见图1,并实现各服务模块之间的无缝集成与协同工作,从而支持用户从发现信息到获取原文的平滑过渡,简化操作过程,降低操作难度,确保数字图书馆的可用、易用。

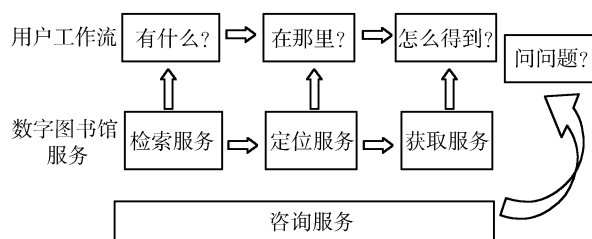


图1 面向用户工作流的协同服务

2.2 基于开放接口的集成检索服务

(1) 国家科学数字图书馆资源范围

用户信息需求很难由一个图书馆的馆藏资源给予满足。国家科学数字图书馆用户所需要的资源包括4类:第1类是国家科学图书馆收藏的、具有自主知识产权的资源,包括期刊集成目录、图书联合目录、中科院学位论文、中科院研究所产出的科学数据、ARP管理资源、机构知识库资源以及中科院出版的学术期刊资源等;第2类是国科图购买了使用权的商业数据库资源,包括电子期刊、电子图书以及基于协议而建立的本地长期保存资源;第3类是与国家科学图书馆建有协同服务关系的第3方机

构资源,如国防信息中心、标准研究院、计量研究院以及清华大学图书馆提供的国防科技信息、标准信息、标准规程信息、CASHL期刊等资源;第4类是支持开放获取的相关网络资源。为了支持用户从以上资源范围内发现个人所需信息,国家科学图书馆建立了一个基于SRU(Search and Retrieve via URL)^[2]技术的集成检索服务平台。

(2) 封装开放接口SRU

首先,国科图利用OCLC开源工具SRW/U^[3],为馆藏数字资源检索系统、面向商业数据资源的跨库检索系统、面向第3方协作资源的元数据仓储系统以及Web开放资源系统等封装了开放接口SRU,支持国科图应用系统以及第3方服务系统的开放调用。SRU是从Z39.50发展而来的Web信息检索协议,它建立在HTTP通信协议之上。一个SRU/GET请求是一个遵循RFC1738^[4]规范的URL,它由基本URL和检索属性两部分组成。基本检索属性元素包括:操作类型(Operation)、版本(Version)、查询式(Query)、开始记录号(StartRecord)、每页记录数(MaximumRecords)、结果信息表示方式(RecordPacking)、结果排序方式(SortKeys)等。下面是一个SRU查询请求实例:

```
http://crossdomain.las.ac.cn/SRW/services/DEIS?operation=searchRetrieve&query=dc.title="system"&version=1.1&startRecord=1&maximumRecords=10&recordSchema=info:srw/schema/1/dc-v1.1
```

应用系统只要构建并提交以上形式的查询请求,就可以实现与数字图书馆系统的集成,并返回以XML格式描述的相关文献信息。由于SRU具有开放性、易用性等特点,它已经被包括美国国会图书馆^[5]、OCLC^[6]、牛津大学、爱丁堡大学、南安普顿大学、多伦多大学^[7]、欧盟图书馆(The European Library)^[8]等众多国际上有影响的图书馆、信息服务机构广泛采用。

(3) 建立集成检索服务体系

在完成SRU接口封装之后,国科图重新构建网站主页系统,并新开发了用户桌面工具“e划通”、研究所信息服务平台等,建立了国科图集成检索服务体系,见图2。所有这些应用系统的顶层都采取

了用户需求驱动的设计思路, 底层通过 SRU 接口实现与资源服务系统的集成, 从而达到方便、快捷的协同服务目标。基于 SRU 接口, 国科图自己的应用系统、第 3 方系统都可以调用国科图的检索服务, 从而支持国科图与更广泛的组织机构之间的协同服务。例如, 中科院网络中心正在将国科图文献检索服务嵌入到科学数据检索系统中, 以便支持科研用户在进行科学数据检索的同时还可以查询相关的文献信息, 实现科学数据与科技文献信息的跨界协同服务。

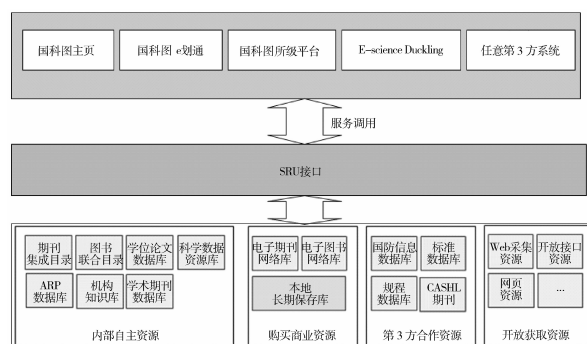


图 2 基于 SRU 接口的集成检索服务体系

2.3 基于情景敏感的动态定位服务

用户利用集成检索工具发现了自己需要的信息资源, 接下来想知道: 这些信息在哪里收藏、收藏机构的服务政策如何、如何请求这些资源等。针对这类问题, 国科图构建了情景敏感的动态定位服务。国科图数字图书馆资源分布在 1 个总馆、多个分馆、90 余家研究所图书馆以及多家第 3 方协作机构内, 不同机构的服务策略不同。为支持众多机构的协同服务, 国科图建立了一个知识库, 采集读者信息、研究所信息、资源信息、服务信息, 并建立读者——所属机构——服务代理关系、研究所——资源——访问控制——服务政策关系、数据库——期刊——论文目次——引文关系等。然后, 利用知识库对当前用户的当前检索结果信息进行分析, 根据用户 IP 确定其所属研究所, 再根据研究所订购资源情况判断用户对当前检索到的信息是否具有直接访问电子全文的权利, 如果有限则选择全文下载服务, 如果没有则为结果文献绑定原文传递、馆际

互借或参考咨询服务, 从而为每条检索结果信息动态选择最佳后续服务。

2.4 基于协同支持的原文获取服务

(1) 获取原文的多种形式

发现并定位信息后, 用户希望能尽快获取原文。获取原文有多种形式, 需要多个机构、多个服务系统以及多种角色馆员的协同工作。其中, 直接下载电子全文需要将请求提交给数据库商, 在通过其访问控制后从全文数据库返回文献全文; 原文传递及馆际互借服务需要将请求提交给资源收藏机构, 前者由收藏馆原文传递服务馆员扫描信息内容并通过电子邮件发送给请求者, 后者由收藏馆互借馆员作出反应; 咨询服务则要将用户的查询结果信息及问题提交给咨询系统, 由系统指定联合咨询馆员作出回答。

(2) 利用 OpenURL 实现基于协同支持的原文获取服务

为了支持用户请求在多个服务系统之间的传递与应答, 保证用户从信息检索到定位、获取全过程的顺畅进行, 国科图引入 OpenURL 技术。即利用 OpenURL 将检索系统生成的结果元数据由检索系统传递到定位系统, 增加定位信息后再由定位服务系统传递到全文数据库系统或原文传递与馆际互借服务系统、参考咨询服务系统等。利用 OpenURL, 不但可以实现用户检索结果与请求信息在多个服务系统之间的自动传递, 使得用户从检索、定位到获取过程变成一个整体, 而且, 也使得跨机构的服务单元被根据用户流程进行动态整合, 实现人员协同、机构协同的目标。例如, 用户检索到一条美军军事标准文献, 经定位分析得知国防信息中心图书馆有其印本收藏, 当用户点击“原文传递”请求后, 该文献元数据信息、用户联系信息等被自动传递到国防信息中心, 国防信息中心原文传递馆员接收到请求后, 启动进库取书、电子扫描、发送电子邮件等操作, 将相关信息发送到用户电子信箱。如果该用户尚不是原文传递系统注册用户, 其请求信息会先传给中间代理服务馆, 经代理服务馆确认后再将请求转发给文献收藏馆。国科图协同服务流程, 见图 3。

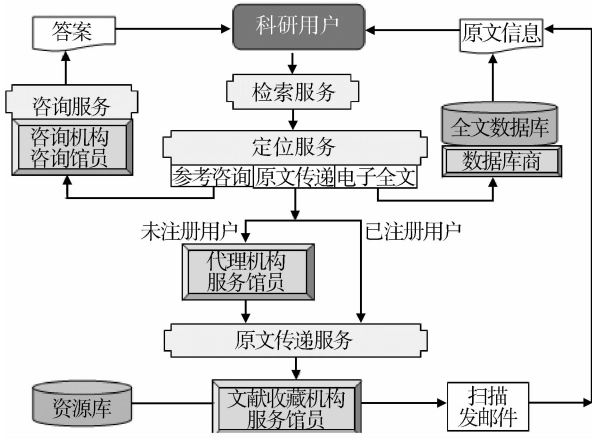


图 3 国科图协同服务流程

3 面向知识聚合的战略情报集成研讨服务

3.1 建立战略情报集成研讨厅

战略情报研究与服务一直是国科图重要任务之一。为有效提升战略情报研究能力，高效和快速地完成从 Data 到 Information 到 Intelligence 到 Solution 的升华，国科图建立了战略情报集成研讨厅。即在网络系统、视频会议系统基础上，通过构建虚拟协同工作空间，把专家、数据和工具融合构成一个智能系统^[9]，支持战略专家、情报专家、学科专家、技术专家针对特定的研究主题，利用共词、共引、关联、聚类等多种分析工具与方法，对多种相关的文献资源、情报研究数据等进行并行分析、多角度交互研讨、多渠道内容实时聚合、多专家即时合作写作，促进实现信息对照激发、多元交互激发、群体智慧捕获情形下的知识创造力聚合，共同创造生成新的战略情报研究产品。协同情报研究模型，见图 4。

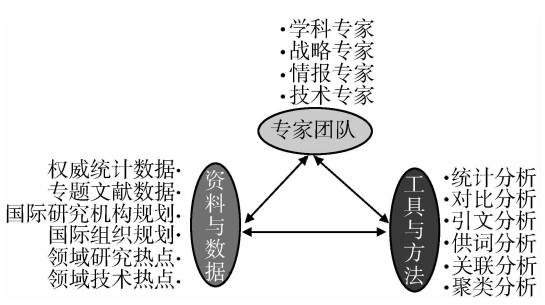


图 4 协同情报研究模型

3.2 战略情报集成研讨厅的技术实现

在技术实现上，战略情报集成研讨厅通过集成视频会议系统，支持本地及远程专家之间的协同交互，支持各类专家对分析过程的实时监控、中间结果判读、分析方向调整等，支持多路信息交互输出。同时，集成 Microsoft OneNote 支持多专家在同一 Word 空间里协同写作，并与多个本地、远程研讨参与者进行即时交流。并利用 SharePoint 建立异步联合工作空间，实现从同步写作、即时交流到异步写作与协同交流的集成。国科图战略情报集成研讨厅的建成，为协同情报研究提供了一个非常好的支撑平台，其应用效果受到广泛的好评。

4 面向集成融汇环境的协同服务

4.1 集成融汇技术的含义与应用

集成融汇是将来自多个来源的数据、服务合并到一个应用系统中，从而构建出新的应用形式。集成融汇的目标是充分利用已有的资源与服务，通过集成多方资源构建新的服务系统，支持图书馆服务向用户所在环境的嵌入，支持用户参与网络服务构建，使用户从单一的信息消费者（Read - only）转向消费者、创建者以及网络服务构建者的综合体（Read - write - and - participate）。作为 Web2.0 的重要技术之一，当前集成融汇技术在数字图书馆协同服务建设中得到广泛应用。例如，安大略剑桥公共图书馆通过融汇来自本馆的 OPAC 信息以及 Bowker Syndetics 公司的图书封面图片信息^[10]，从而构造出一个动态的新书展示服务^[11]；McMaster 大学图书馆将馆藏航空摄影照片资源与 Google 地图融汇，构建出一种新的信息揭示工具^[12]；纽约公共图书馆（NYPL）^[13]将基本服务封装成 Widget 组件，支持向用户桌面环境的嵌入。

4.2 国科图集成融汇环境下协同服务建设思路

(1) 封装融汇服务，支持其他系统发起的协同服务调用

集成融汇环境下协同服务建设的一个重要特点

是支持潜在的、甚至是未知的用户发起协同服务请求并迅速获得预期的服务结果,这也是实现长尾服务的一种重要形式。为了支持以上服务目标,国科图制定了集成融汇服务组件(Widget)封装计划,即在 SRU 检索接口基础上,将数字图书馆系统中的基本服务、专题服务封装成一个个具有特定功能、可以独立调用的服务组件,在图 2 SRU 接口层之上增加融汇服务组件层,支持目标应用系统的调用,包括:Facebook、Myspace 等用户桌面系统(Desk-Top),以及 iGoogle、Netvibes、Webwag 等用户起始页面(Start Page)。这些应用系统的一个共同特点是支持用户个人自我构建服务系统,不需要编程,不需要具备特殊的计算机技能,从而使用户参与服务构建成为可能。

(2) 创建融汇服务仓储,支持用户协同参与的融汇服务共建共用

创建融汇服务仓储系统,摄入、保存、管理融汇组件对象,支持开放的服务发现、调用。收藏对象既包括图书馆领域融汇组件,也接纳来自数字教学、数字科研、数字出版、数字生活等相关领域的服务组件。组件构建者既可以是国科图馆员,也可以是联盟馆馆员,以及其他机构人员和任意用户。在应用端,任何组织、机构和个人都可以根据个性需求从仓储系统中发现相关服务,并将其集成到个人应用系统中。同时,仓储系统支持服务组件的迁移,包括将 Google Gadget、Yahoo Widget、Netvibes Widget 等第 3 方融汇组件迁移到自己的仓储系统中,以及将自己仓储系统中的融汇组件迁移到第 3 方融汇平台上。融汇服务仓储系统架构,见图 5。

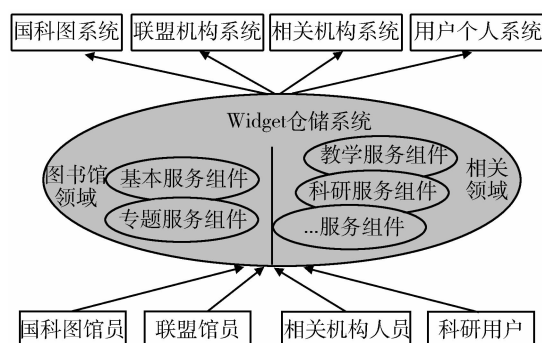


图 5 融汇服务仓储系统架构

(3) 开发桌面工具,支持用户构建 iLibrary

目前的图书馆网站都是由图书馆员设计的,在内容组织、流程设计上未必符合用户的个性需求。在融汇服务仓储之上,计划开发一个类似 iGoogle 的用户桌面工具,支持用户根据个性需求自我设计、构建个人图书馆页面。

5 结语

信息技术无疑是协同服务建设的重要支撑。当前协同服务技术呈现以下几方面特点:开放标准,支持跨机构服务协同,支持面向用户信息利用过程的服务协同,支持动态协商与双向融汇服务;开放关联数据(Linked Open Data),支持不同来源数据的协同计算,目前 W3C 已经发布 100 多个关联数据源;开放源码,实现协同设计开发, JISC 2008 报告认为,开放源码是 LMS 未来发展方向;开放服务,支持用户作为协作者参与服务构建, Netvibes 生态环境已经收集了来自 115 个国家、1.5 万用户贡献的 18.5 万次服务,每月提供第 3 方服务调用超过 5 亿次;新技术给协同服务建设提供新的机遇和挑战,如基于云计算的协同存储、分析与服务技术,动态服务发现、调用与融汇服务技术,以及基于情景感知的 U-服务技术等。国家科学图书馆在协同服务建设方面做了一些尝试,但在更大范围的跨领域协作、更深入的用户参与的协同服务建设等方面还有待做进一步探索,这也将是未来数字图书馆协同服务研究与建设的一个重要方面。

参考文献

- 1 Alexandra Y, Barbara C. Public Libraries, Archives and Museums: trends in collaboration and cooperation. [EB/OL]. [2009-12-24]. <http://archive. ifla. org/VII/s8/pub/Profrep108. pdf>.
- 2 SRU: search and retrieve via URL (standards, Library of Congress) [EB/OL]. [2009-12-24]. <http://www. loc. gov/standards/sru/index. html>.
- 3 OCLC SRW/U [EB/OL]. [2009-12-24]. <http://www. oclc. org/research/activities/srw/default. htm>.

(下转第 26 页)