

# 联合目录中电子资源的集成与揭示

路纳新

(中国科学院国家科学图书馆 北京 100190)

**【摘要】** 分析国内外电子资源组织的发展现状及面临的问题,结合用户的需求,提出在中国科学院联合目录数据库中建立电子资源知识库。该文介绍了中国科学院基于联合目录的电子资源知识库的建设目标和思路,详细描述了电子资源知识库的数据结构、集成组织方式及维护更新机制等,及支持情景敏感的集成发布。最后总结了电子资源知识库的不足及发展方向。

**【关键词】** 电子资源知识库 联合目录 集成揭示 情景敏感

**【分类号】** G250.74

## 1 引言

中国科学院联合目录数据库是一个多学科的大型数据库,学科范围覆盖自然科学、农业科学、医药卫生、工业技术、社会科学等,资源类型包括各文种印刷版图书、连续出版物,电子版中西文图书、连续出版物以及数据库信息。经过多年的发展,联合目录数据库已成为多类型、多载体的集成揭示目录,用户可以通过联合服务系统,方便地检索到各种类型的文献。

随着网络技术的发展,读者对网络电子资源的需求的不断增加,电子资源在图书馆资源建设中正呈异构性,多样性和动态性不断增加。尤其是电子期刊,由于其具有很强的实效性,是科研人员在科研活动中非常关注且使用广泛的文献资源。随着各成员单位电子文献资源比例的不断增加,将成员单位的电子资源在联合目录中进行组织和揭示已成为各成员单位的强烈需求。

## 2 国内外电子资源组织揭示现状

国外许多电子期刊集成系统中,都有类似“知识库”的组件。它们主要有两种用途:(1)作为电子期刊列表的底层数据,支持开放链接;(2)作为链接数据库。比如 SFX 系统的 Global KnowledgeBase,在支持开放链接机制的基础上,内容包括了从集成商或出版社获取的刊名列表,覆盖日期,以及要求的链接语法等。Icate 的知识库是 Openly 公司花六年时间建设起来的,该知识库存储着来自 1,200 多家出版社的 75,000 多种电子期刊的描述信息和链接信息。此外,系统还包括了 600 多个款目表,以通用款目格式(UHF, Universal Holdings Format)存放,范围基本上涵盖了目前顶尖的电子期刊集成商(ACM、ALPSP、NASA Astrophysics Data System、AIP、Blackwell-Synergy、EBSCO Host、Elsevier Science Direct 等)。还有 Serials Solutions 的 E-Journal A. M. S 也有相应的电子资源知识库做为底层数据支持开放链接。

而国内图书馆对电子资源的组织和揭示大部分采用浅层组织的方式,如按电子资源的名称或数据库商的名称在图书馆的主页上进行罗列,或按电子资源学科进行分类罗列,或按文

献类型进行区别,或综合以上方式进行揭示。也有图书馆建立了电子资源导航系统,将各种电子资源集成检索和显示,如清华大学的电子资源导航系统;部分图书馆或机构对电子资源进行了编目揭示,如 CALIS 联合目录及 NSTL 联合目录对电子资源(数据库及科技期刊)的揭示主要继承印本资源的组织揭示方式,以 MARC 方式进行书目级的描述著录,在 856 字段对电子资源进行检索定位。

虽然国内图书馆对电子资源组织揭示的研究是近年来的热点问题,而且很多图书馆已着手电子资源系统和数据库的建设,但基本都是基于单馆模式下的电子资源集成与提示,读者只能通过本馆主页查询和使用本馆订购的电子资源,无法获知电子资源在其它图书馆的收藏情况。目前,在国内尚无多馆联盟方式下的电子资源知识库。

### 3 中国科学院电子资源知识库建设

中国科学院联合目录拥有包括院内研究所在内的 400 余家成员馆,各单位学术领域不同,所订购的印本及电子资源也不尽相同,从而产生了不同用户对电子资源访问权限多样性的特点。中国科学院联合目录通过对用户需求的调查,在分析了自身情况,跟踪国内外相关系统的发展后,于 2009 年对联合目录的系统功能进行改造优化,建立了基于联合目录的电子资源知识库。

#### 3.1 建设目标及思路

中国科学院电子资源知识库的建设目标是:建立针对联盟图书馆或多分馆系统的电子资源知识库,支持成员单位电子资源的集成、动态更新和动态发布,实现印本与电子资源的有效集成与揭示,实现印本与电子资源的同步更新,实现电子资源信息的自动更新与批量更新,支持情景敏感的集成服务。

基于以上目标,中国科学院联合目录扩展了系统功能,对联合目录数据库的数据结构,以及电子资源的集成组织和揭示方式进行改造和优化,建立了电子资源知识库。解决了电子资源的更深层的揭示,情景敏感以及电子资源维护更新等问题。围绕着电子期刊资源和服务的相关信息,建立起成员单位、订户、印本资源、电子资源及服务之间的多种层次的复杂关系,制定了多层次的资源和服务策略与规则,形成支持情景敏感服务的知识库。联合目录集成服务系统通过印本资源、电子资源的集成揭示,与用户 IP 结合,形成独特的情景敏感链接,可以使用户方便地获取有使用权限的电子资源全文。通过该系统,用户可获得电子资源在联合目录成员单位的分布情况,即使没有权限直接获得全文,也可通过原文请求方便获取所需资源。

#### 3.2 系统的整体架构

电子资源知识库通过电子资源信息、订户信息、收藏机构信息,关联信息,几部分之间复杂多层次的关系,实现根据读者特点动态为其呈现恰当的资源 and 所能获取的服务,构建情景敏感的服务系统环境。

(1) 电子资源信息:包括数据库和电子期刊的基本描述信息以及开通信息等。

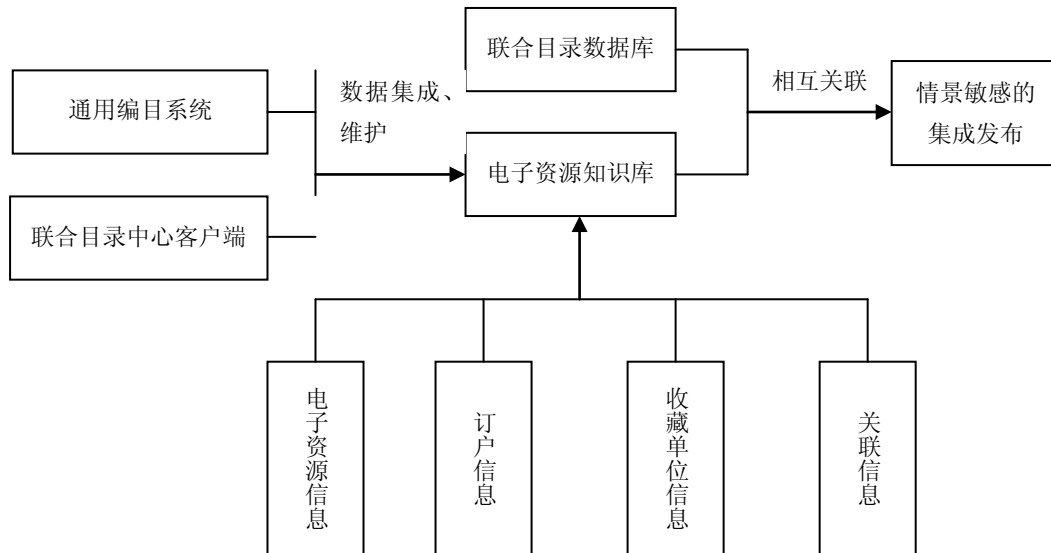
(2) 订户信息:包括资源订购的集团信息,如 CSDL、NSTL。电子资源知识的馆藏结构分为多层,订户可视为第一层的馆藏机构。订户可能是包含多个单位的集团,也可能是一个单位。

(3) 收藏单位信息:包括联合目录的成员馆,电子资源的收藏单位,是电子资源知识库的第二层馆藏机构。包括单位代码、单位名称、联系人、联系方式,IP 地址等信息。IP

地址是电子资源知识库最终实现情景敏感的重要信息。

(4) 关联信息：包括电子资源信息与订户信息，电子资源信息与印本资源信息之间的关联关系。通过各种关联信息，最终形成复杂的多层次关系的电子资源知识库，支持情景敏感的集成揭示。

中国科学院电子资源知识库的整体架构如下图：



### 3.3 元数据及数据结构

中国科学院电子资源知识库采用多元化的元数据组织（MARC、电子资源元数据）。电子资源信息改变了以往在书目 MARC 中添加 856 字段检索定位的组织方式，而是用电子资源元数据对电子资源信息进行组织描述，同时与印本资源的 MARC 数据进行整合，形成即分散又关联的电子资源知识库。通过与印本资源的关联，将同一资源不同载体的资源进行了统一的描述揭示，最终呈现给用户最为完整和准确的资源描述信息、电子资源定位信息以及馆藏信息。

数据库元数据在知识库中的显示：

序号	数据库id	数据库名称	数据库名称缩写	数据库url	数据库类型	数据库内容类型	数据库主题词	揭示深度
1	2010a002	OCLC FirstSearch-f	OCLC FirstSearc	http://firstsearch.oclc.org/	文献	索引		文摘
2	2010a002	CSA-Aqualine	CSA-Aqualine	http://csa.tsinghua.edu.cn/	文献	索引		文摘
3	2010a002	Cambridge Journals	Cambridge	http://journals.cambridge	文献	期刊		全文
4	2010a002	OCLC FirstSearch-i	OCLC FirstSearc	http://firstsearch.oclc.org/	文献	索引		索引
5	2010a002	ProQuest-EIU View	ProQuest-EIU	http://proquest.umi.com/	文献	商业信息		全文
6	2010a002	Dialog	Dialog	http://www.dialog.com/	文献	检索平台		索引
7	2010a002	EBSCO host-Maste	EBSCO host-MFF	http://search.ebscohost.c	文献	多出版类型		全文
8	2010a002	CSA-Library & Infor	CSA-LISA	http://csa.tsinghua.edu.cn/	文献	索引		文摘
9	2010a002	EBSCO host-Regioi	EBSCO host-RB	http://search.ebscohost.c	文献	商业信息		全文
10	2010a002	PsycInfo	PsycInfo	http://csaweb111v.csa.c	文献	多出版类型		文摘
11	2010a002	EBSCO host-Acade	EBSCO host-ASF	http://search.ebscohost.c	文献	期刊		全文
12	2010a002	CSA-Physical Educ	CSA-PEI	http://csa.tsinghua.edu.cn/	文献	索引		文摘
13	2010a002	IWA	IWA	http://www.swetswise.co	文献	期刊		全文
14	2010a002	SPIE Digital Library	SPIE	http://scitation.aip.org/	文献	期刊、会议论		全文
15	2010a002	OCLC FirstSearch-l	OCLC FirstSearc	http://firstsearch.oclc.org/	文献	期刊		全文
16	2010a002	ProQuest-ProQuest	ProQuest-Newsp	http://proquest.umi.com/	文献	报纸		全文
17	2010a002	SciFinder-CHEMLIS	SciFinder-CHEM	http://lib.utsz.edu.cn/view	文献	化合物性质		文摘

电子期刊元数据在知识库中的显示：

序号	期刊记录控制号	刊名	P-Issn	E-Issn	所属数据库名称	出版商	起始年	截止年	学科分类	主题关键词	获得来源
1	2009a0020011	4OR	1619-4500	1614-2411	Springer(Calis)						CALIS
2	2009a0020002	4OR: A Quart	1619-4500	1614-2411	SpringerLink期刊	Springer Be	1997				CSDL
3	2009a0020001	sterreichisch	0945-358X	1613-7566	SpringerLink期刊	Springer Wi	1997				CSDL
4	2009a00200017	sterreichisch	1011-0070	1862-2585	SpringerLink期刊	VS Verlag fi	1997				CSDL
5	2009a0020008	AAPS Pharm:		1530-9932	SpringerLink期刊	Springer Ne	1997				CSDL
6	2009a0020000	ACS Chemica	1554-8929	1554-8937	ACS	American C	2006				CSDL
7	2009a0020005	ACS Nano	1936-0851	1936-086X	ACS	American C	2007				CSDL
8	2009a0020005	ADVANCED	0921-8831	1568-5527	网络版电子期刊	Brill Acaden	2007				CSDL
9	2009a0020012	AGE	0161-9152	1574-4647	Springer(Calis)						CALIS
10	2009a0020001	AGE	0161-9152	1574-4647	SpringerLink期刊	Springer Ne	1997				CSDL
11	2009a0020004	AGE	0161-9152	1574-4647	SpringerLink回溯	Springer	1978	1996			NSTL
12	2009a0020011	AI & SOCIETY	0951-5666	1435-5655	Springer(Calis)						CALIS
13	2009a0020001	AI & Society	0951-5666	1435-5655	SpringerLink期刊	Springer Lo	1997				CSDL
14	2009a0020003	AI & Society	0951-5666	1435-5655	SpringerLink回溯	Springer	1987	1996			NSTL
15	2009a0020009	AI Communic	0921-7126	1875-8452	IOS Press Journa	IOS Press	1997			Computer S	NSTL
16	2009a0020000	AI EDAM	0890-0604	1469-1760	Cambridge	Cambridge	1998				CSDL
17	2009a0020004	AIChE Journa	0001-1541	1547-5905	Wiley回溯	wiley	1955	1997			CSDL

### 3.4 数据集成方式

资源集成是中国科学院电子资源知识库建设最重要的工作。中国科学院电子资源知识库的集成提供批量转入和远程在线提交两种方式。对于中国科学院集团订购的电子资源，采用批量转入的方式集成，对于成员单位自主订购的电子期刊采用远程提交的方式。

(1) 批量转入。中国科学院联合目录系统在原来只提供 MARC 数据转入功能基础上进行了扩展，系统可以提供 excel 表的转入接口，改变了以往必须以 MARC 格式转入的方式。数据转入过程中，系统根据设置的查重点可进行自动查重入库或归并，同时在数据转入时根据特殊的信息点建立各种关联关系。

(2) 远程上载。中国科学院联合目录在进行系统功能扩展的同时，对“通用编目系统”进行了升级改造，增加了电子期刊集成模块。通过该模块，实现对联合目录成员单位自主订购电子期刊的远程提交和电子期刊信息的维护更新。

### 3.5 维护更新机制

电子资源维护更新主要涉及以下方面：电子资源品种变化、电子资源馆藏范围变化、电子资源开通范围变化、电子资源 URL 地址变化等。中国科学院电子资源知识库的维护更新根据不同的情况，采用批量处理或远程维护的方式，以保证数据在第一时间得到完全准确的维护更新，保证读者随时获得准确方便的服务。对于中国科学院集团订购的电子资源，我们可随时掌握电子资源的变化及参团单位馆藏的变化情况，采用批量集中处理的方式进行维护，系统不提供远程维护的权限。成员单位自主订购的电子资源如果发生变化，则需成员单位通过“通用编目系统”的电子期刊集成模块进行处理。电子资源知识库对各种类型的电子资源数据均提供人工修改处理的功能。

## 4 支持情景敏感的集成发布

通过中国科学院电子资源知识库的建设，以及知识库各系统模块与元数据间的关联关系，实现了电子资源与印本资源情景敏感的集成发布，实现了电子资源更深层次内容的揭示。

(1) 支持系统判断读者 IP 范围以确定读者对电子资源的访问权限，不同的权限有不同

的显示方式。如果有访问权限的资源,则在显示检索到的资源列表时,显示“+e”的标记。无访问权限的资源,则没有“+e”的标记。

(2) 为读者提供情景敏感的电子期刊开通范围。除可判断读者是否有读取全文的权限外,还可根据读者所在的单位,为读者揭示可获取该电子期刊的年代和卷期范围。

(3) 支持情景敏感的集成浏览功能。提供电子资源与印本资源的集成化的字顺浏览功能。并判断读者 IP,显示是该电子资源是否可访问全文。

## 5 中国科学院电子资源知识库的发展方向

中国科学院电子资源知识库无论资源类型还是收录范围都处于起步阶段,只集成了中国科学院集团订购及院内外少数单位自主订购的数据库及电子期刊资源。资源类型单一,集成范围较小,尚未形成规模。下一步电子资源知识库将集成更广泛的电子资源类型,并构建可由发布系统或其它相关系统自由调度的电子资源知识库。

(1) 资源类型的扩展。中国科学院电子资源知识库目前集成的资源范围主要是中西文电子期刊,电子图书尚未纳入到知识库中。电子图书无论从描述方式和服务都与期刊有很大不同。电子资源知识库正致力于电子图书的集成,在现有知识库的基础上继续扩展适用于集成电子图书的系统功能,制定电子图书元数据描述与组织规范,集成中国科学院集团订购的电子图书资源。

(2) 资源范围的扩展。中国科学院电子资源知识库同时在加强院外及成员单位自主订购电子资源的收集与整理,与院外机构和单位加强合作交流,实现在全国范围内电子资源知识库的共建共享。目前已与 BALIS 集团签定共建共享协议,为最终实现建立全国数字资源知识库的目标提供了有力的资源保障。

(3) 开放获取资源的集成。在网络技术飞速发展的今天,网上开放获取资源越来越多。中国科学院电子资源知识库同样关注开放获取资源的集成与组织描述。目前知识库已集成了部分开放获取的电子期刊,并为科研人员提供服务。知识库将针对开放获取资源的获取、资源组织描述及数据更新维护等方面开展研究和探索,实现开放获取资源在知识库中的最大集成,为用户提供更广泛可用的资源信息。

## 6 结语

中国科学院电子资源知识库经过近一年时间的建设,实现了电子资源与印本资源的有效集成,电子资源知识库支持动态更新、动态发布、情景敏感的集成服务,为多馆联盟模式下电子资源的集成组织与服务提供了一种新的思路和方法。但电子资源知识库的建设毕竟是在摸索中前进,在使用和服务的过程中也暴露了一些问题,如数据结构与电子资源元数据的设置不够合理,造成数据的冗余;在数据处理过程中系统速度较慢等问题,还需我们在实践过程中逐步完善。

## 参考文献:

- [1] 联合目录集成服务系统 [DB/OL]. [2010-03-20] <http://union.csd1.ac.cn/index.jsp>
- [2] 中国科学院集成期刊系统 [DB/OL]. [2010-04-10] <http://www.las.ac.cn/ejournal/ejournals.jsp>
- [3] 吴振新,张智雄,李春旺,等.情景敏感的集成期刊服务系统建设与实现[J].现代图书情报技术,2006,(12):1-4.
- [4] 王利蕊.高校图书馆数字资源整合方式研究[J].现代情报,2009,(2):75-77.
- [5] 孙晓丽.公共图书馆对电子资源的组织与揭示引发的思考[J].图书馆理论与实践,2009,(2):69-70.
- [6] 熊大红,戴小鹏,罗影等.组织知识共享模型与技术研究[J].图书情报工作,2009,(20):114-117.

- [7] 李广建. 面向信息机构的嵌入式 NSTL 资源集成服务系统的设计与实现[J]. 现代图书情报技术, 2009, (6): 2-7.
- [8] 吴常青. 图书馆知识集成理论与实践探讨[J]. 图书情报工作, 2009, (19): 32-35.
- [9] SFX[R/OL]. [2010-04-12] <http://www.exlibris.com.cn/product/sfx/index.asp>
- [10] 1cate[R/OL]. [2010-04-05] <http://www.openly.com/journalseek/>
- [11] Serials Solutions[R/OL]. [2010-04-10] <http://www.serialssolutions.com/chinese-simplified/>

## 作者简介:

路纳新, 1975 年生, 馆员, 中国科学院国家科学图书馆, 北京海淀区北四环西路 33 号, 010-82626611-6111, lunx@mail.las.ac.cn, 100190