

与科学数据库结合的机构知识库建设探索

刘国俊 许哲平

(中国科学院植物研究所, 北京 100093)

摘要: 我国机构知识库建设取得一定的成果, 但通过对OpenDOAR的数据分析和对比发现, 我国在全球机构知识库中的影响力还比较小。文章提出了一个与科学数据库相结合的IR建设新框架, 并对功能和特点进行了介绍。IR建设需要继续紧密结合科研实际需求, 补充和挖掘科学数据和特色资源, 加强相关学科的信息资源整合, 特别是科学数据库, 拓展服务方式, 提高服务的深度和专业性。

关键字: 科学数据库; 机构知识库; 科研活动; 大数据; 数据共享; 访问量

中图分类号: G203

文献标识码: A

DOI: 10.3772/j.issn.1674-1544.2013.04.005

Research on Construction of IR with Scientific Database

Liu Guojun, Xu Zheping

(Institute of Botany, Chinese Academy of Sciences, Beijing 100093)

Abstract: The construction of IR has been achieved a lot of fruits in recent years in China, however, it also encounters some problems. The era of Big Data brings huge opportunity and challenge for future development of IR. From the analysis of OpenDOAR data, China has less IRs than many other countries in the world. The development of current IRs in China is not good as well. A new architecture of IR development has been introduced and described in this article. In the future, IRs should introduce new concept, develop new tools, build more connections with scientific data and focus on their own special resources and make them available online. Furthermore, different IRs should improve and extend current information service based on the features of different disciplines and institutes.

Keywords: scientific database, institutional repositories, scientific researches, big data, data sharing, access statistics

尽管机构知识库(IR)的定义还有一定的分歧, 但是比较一致的观点是: IR是各类机构(主要是科研或教育机构)对自身的知识产出以数字化方式进行收集、保存和传播的相关工作^[1]。其主要的资源类型包括各类期刊文章(已发表或未发表)和各种论文(学位论文和专著等), 也可能包括其他的科研活动产生的数字对象, 如实验室数据、野外观测数据、中间处理数据、数据处理软件、管理文档和各类学习资源等, 数据格式包括文本、多媒

体(音频、视频)、数据库和软件等。因此, 广义上来讲, IR应当是包含一系列科研活动产生的科学数据, 但是从目前来看并没有很好地将其整合进来, 反而是科学数据的独立建设和共享做得有声有色。2012年, 包括美国《大数据研究和发展倡议》在内的一系列报告将以前学术界讨论的“大数据”概念进行了落地, 将在今后的几年里对相关的学术界和产业界产生重大影响。在这种大环境下, IR建设也面临着更大的机会和挑战。

第一作者简介: 刘国俊(1980-), 女, 中国科学院植物研究所馆员, 研究方向: 机构知识库建设、数字图书馆。

基金项目: 国家科技基础条件平台项目“植物标本标准化整理、整合及共享平台建设”(2005DKA21401)。

收稿日期: 2013年5月30日。

1 国际机构知识库发展现状

在国际上, OpenDOAR (The Directory of Open Access Repositories) 是首个对全球范围内的开放存取知识库进行搜集、描述和组织的项目, 由英国的诺丁汉大学和瑞典的伦德大学图书馆于2005年共同创建^[2-3], 并在2006年正式发布上线。目前, 全球有2283家开放知识库在该网站进行了注册。通过对这些知识库进行分析, 能有助于了解当前国际机构知识库的发展现状(以下OpenDOAR的统计数据均截止到2013年5月28日)^[4]。

按国家分布的机构数来排名, 前十名分别是: 美国(395家)、英国(210家)、德国(164家)、日本(138家)、西班牙(98家)、波兰(75家)、意大利(72家)、法国(71家)、巴西(67家)、加拿大(58家)。而中国只有33家机构注册在案。当然, 这与我国实际情况还是有一定出入的。仅仅中国科学院的IR建设目前就有83家单位(<http://www.irgrid.ac.cn/>), 即便考虑些未统计的数量, 与美国、英国、德国和日本等国还是有一定的差距。

按存储的数据类型来看, 2283家机构存储的数据类型包括期刊论文(1549家)、学位论文(1220家)、未发表报告(823家)、专著(817家)、会议论文(804家)、多媒体和视听材料(536家)、参考文献库(422家)、其他特殊类型(361家)、学习材料(350家)、数据集(85家)、专利材料(65家)、软件(33家)。对比张冬荣等人的研究成果^[5], 可以发现中国科学院的IR存储内容多样性太低, 多集中在期刊、学位和会议论文。而且, 这里很多是可以通过ScienceDirect、Wiley、SpringerLink或CNKI、万方、维普等数据库得到的, 并没有明显的机构独有特色, 自然也无法对用户形成很大的吸引力, 而对数据库等科研信息资源则基本没有涉及。而科学数据库恰恰能对IR的建设起到很好的补充作用^[6]。

为了更好地体现数量上的对比, 这里从OpenDOAR上选取拥有50万条以上数据记录的机构知识库, 见表1。可以看出的是, 表1中除了传统的实体机构之外, 也包括许多类似PubMed的在线数据库平台, 作为一种虚拟机构也包含在这个列表中, 这无疑释放一种信息: 无实体的虚拟知识库也是一种客观存在的资源, 并且可能会蚕食或取代没有特色和稀缺资源的实体机构知识库。另外, 在众多欧美国家中, 也有2个中国机构入围。表1还列

出了这些组织机构知识库建设过程中使用的软件, 可以看出, 要想更好地挖掘和应用知识库资源, 推出特色服务, 在确保资源能够有效利用的前提下, 还需要加大自身平台的研发力度。

通过对OpenDOAR的资源进行简单分析和对比, 可以发现, 尽管我国机构知识库的发展取得了一定成绩。但是, 欧美等国的IR资源无论在资源的丰富度、多样性以及广度和深度方面都有了长足的发展。这与我国IR建设过程中出现同质化严重、资源类型单一、关注度低等方面形成了鲜明的对比, 这也为今后我国的IR建设指明了一定的方向。

2 中国科学院机构知识库现状

我国目前主要的机构知识库为CALIS机构知识库(中国高等教育数字图书馆, 网址为<http://ir.calis.edu.cn/>)和中国科学院机构知识库服务网格平台(<http://www.irgrid.ac.cn/>)。前者的元数据总量为7.9万条, 28家IR(2012年11月18日)。后者为43.2万条资源, 83家IR(2012年11月18日)。这里重点对后者进行介绍和分析。

在对中科院机构知识库的79家参与单位进行在线调研工作(2013年5月4日)中发现, 有24家单位的IR网址打不开, 占总数的30%。其中79家单位的资源分布情况如图1所示。

为了相对客观地了解IR资源的访问情况, 从中选择31家单位做进一步的分析。选择的原则是: (1)3000条记录以上; (2)能够在线访问; (3)页面上有资源总量、访问量和下载量的统计数字。然后利用机构编号、资源总量和访问总量3个参数进行图表分析(表2)。

从上述图表中可以看出, 资源总量与访问量(使用情况)有一定的正相关性, 但并不完全如此。如资源量在1万条以上的IR, 访问量都在100万人次以上。而类似水生生物研究所、沈阳自动化研究所、上海光学精密机械研究所、国家科学图书馆等单位的IR资源量虽然不足1万, 但其访问量也在百万人次以上。而植物研究所、南海海洋研究所、山西煤炭化学研究所等单位虽然资源量和访问量都不太突出, 但是单一资源的使用情况(访问量/资源数)却比一些资源量或访问量靠前的IR单位有更好的表现。尽管访问量可能存在的问题^[7]。但是, 这样的多种指标综合对比确实能够对IR的运行和访问情况进行一定的定量评估, 然后从中发现问

表1 OpenDOAR上拥有50万条以上记录的机构列表(截至2013年5月27日)

组织机构名称	国家	记录数	软件
Papers Past	新西兰	36215172	Greenstone
Europe PubMed Central	英国	26500000	PMC
Research Papers in Economics	美国	12000000	未知
English Heritage ViewFinder	英国	8000000	未知
Internet Archive	美国	7419856	HTML
DIALNET	西班牙	3826518	未知
Geograph British Isles	英国	3442540	未知
PubMed Central	美国	2700000	PMC
University of Michigan Library Repository	美国	2269339	Perl-based
Entrepôt OAI-PMH numéro 1 de la Bibliothèque nationale de France : collections numériques	法国	1785056	未知
Education Resources Information Center	美国	1479798	未知
CERN Document Server	瑞典	1269877	invenio
NASA Technical Reports Server	美国	1069140	未知
Predicted Crystallography Open Database	立陶宛	1062771	HTML
Biblioteca Virtual de Prensa Histórica (Virtual Library of Historical Newspapers)	西班牙	1049482	Digibib
J. Willard Marriott Library Digital Collections	美国	1000000	CONTENTdm
Bayerische Staatsbibliothek – Münchener Digitalisierungszentrum	德国	943977	未知
Social Science Research Network	美国	875700	未知
Marshall Technical Reports Server	美国	847651	未知
arXiv.org e-Print Archive	美国	843658	arXiv
Korea Thermophysical Properties Data Bank	韩国	837789	未知
OpenGrey Repository	法国	690609	Exalead
中国科学院西安光学精密机械研究所机构知识库	中国	580495	DSpace
University of Cincinnati Digital Resource Commons	美国	567978	DSpace
中国科学院近代物理研究所机构知识库	中国	547815	DSpace

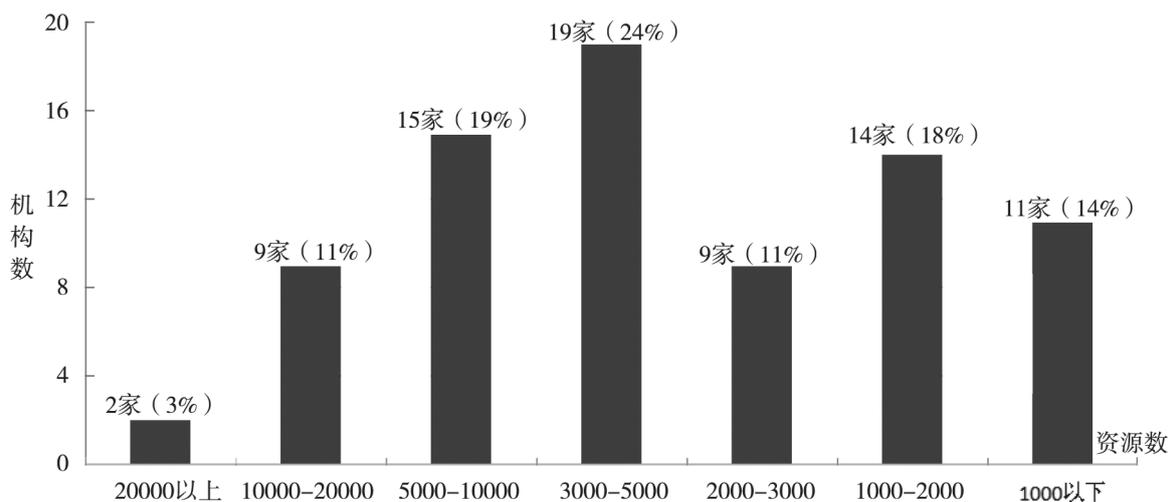


图1 中国科学院 IR 资源数量分布图

表2 中国科学院 IR 资源与访问情况图 (截至2013年5月4日)

编号	机构名称	资源数	资源访问量	访问量/资源数
1	工程热物理研究所	20581	1012209	49.182
2	高能物理研究所	19917	1043075	52.371
3	大连化学物理研究所	18674	2779787	148.859
4	长春应用化学研究所	18514	2132137	115.163
5	地理科学与资源研究所	16340	2741307	167.767
6	上海微系统与信息技术研究所	13145	1132096	86.124
7	力学研究所	11691	3180071	272.010
8	半导体研究所	11141	4588213	411.831
9	上海有机化学研究所	9059	146807	16.206
10	水生生物研究所	8937	1945102	217.646
11	合肥物质科学研究院	7509	1080558	143.902
12	昆明植物研究所	7426	923630	124.378
13	沈阳自动化研究所	7006	1811248	258.528
14	软件研究所	6500	239002	36.770
15	上海光学精密机械研究所	5608	1833969	327.027
16	近代物理研究所	5225	670615	128.347
17	国家科学图书馆	5156	3118475	604.824
18	昆明动物研究所	4841	363349	75.057
19	山西煤炭化学研究所	4780	946771	198.069
20	上海应用物理研究所	4479	429587	95.911
21	广州地球化学研究所	4304	518138	120.385
22	心理研究所	4139	741610	179.176
23	理论物理研究所	4036	258185	63.971
24	植物研究所	3903	881837	225.938
25	成都生物研究所	3827	637853	166.672
26	西双版纳热带植物园	3806	630695	165.711
27	亚热带农业生态研究所	3664	210063	57.332
28	上海技术物理研究所	3331	2719	0.816
29	紫金山天文台	3235	392766	121.411
30	新疆生态与地理研究所	3105	492707	158.682
31	南海海洋研究所	3017	668233	221.489

题。而不是单纯地从一些定性或仅仅从资源量来进行服务评估。

3 我国科学数据库现状

我国目前的科学数据主要集中在国家科技基础条件平台中心 (<http://www.escience.gov.cn>) 支持下的23家首批国家科技基础条件平台中的部分平台,如国家标本资源共享平台等。这些平台由多家具有雄厚科研实力的机构和高校共同参与建设和维护,都有着多年的建设和运行经验,而作为国家层面的总体部署来看,也同时考虑到了政府机构、科研部门和社会公众等方面的需求,将成为我国今后长期基础信息化工作中的重要一环。此外,还包括一些

以学科主题为主线的在线数据库系统,其中就包括基础科学数据共享网 (<http://www.nsd.cn>)。该平台由中国科学院计算机网络信息中心牵头组织,以中国科学院、国内重要高校和其他科研院所的基础科学数据资源为基础,充分利用中国科学院已有的数据应用环境基础设施条件,开展基础科学领域的数据资源整合集成与共享服务,主要部署了物理、化学、生物、天文、空间等多个学科以及特定地区的专题服务,其长远目标也是成为国家科技条件平台的认定平台之一。

由于基础科学数据库共享网汇总了多个基础科学数据库,有较好的学科代表性,并且有很好的在线资源统计和访问统计数据,所以本文以此为依据

来探索科学数据库的建设和使用情况。为了更加具体地说明该平台的建设和运行效果,同时也为了与IR建设进行对比。这里也选择了资源记录数、访问量和访问量/记录数3个指标进行综合制表(表3)。

从表3中可以看出与IR建设类似的情况:资源量大的平台,其访问量不一定大。而总体访问量大的平台,单一资源访问量也不一定大。但是,表3还可以用于分析不同学科的资源建设与访问情况:生命学科的总体访问和单一资源访问量比其他学科要高不少,然后是化学和物理,其次是天文、微生物和核物理。

还可以看出的是,与科学数据的访问量相比,IR的访问量还是太低:表2中31家IR机构所有的访问量总和才4170万人次,而表3中仅植物主题数据库(4.7亿人次)、系统生物学多组学综合数据库(3亿人次)、高能物理数据库(1亿人次)这3个平台的独立访问量就都在1亿人次以上,而这还只是一个年度的访问统计。实际上,中科院IR和基础科学数据共享网数据库都有众多的单位参与,涉及不少高校和科研机构,有很多是参与到科学院IR建设和高校的CALIS建设中,而且有些科学数据资源本身就是各个参与单位的优势学科资源或者专业文献资料。如果这两者能够有效结合起来的话,不但可以使信息资源相互补充,而且极大地提升IR的资源和服务内涵,还可以打通目前高校和科学院两大独立的IR知识库的资源流通渠道。

4 与科学数据库结合的IR建设方案框架

从上面的分析来看,我国目前的IR建设,由于意识不到位、领导不重视、项目支持少、人员队伍水平不高、资源扩展不足等原因,还多数停留在传统文献的收集和保存层面,属于比较低层次的建设和服务。通过与科学数据库的结合,IR能够从科学数据库项目中获取更多的关于本机构的信息资源,扩展自身的资源来源渠道,收获社会效益。而科学数据库可以从IR中获得更多有特色的专题数据库。理想的情况是在源头上就将科学数据库纳入IR的范围内,这样会使目前IR建设突破窘境。不过,由于科学数据库大多数是得到各类项目或课题的资助,其影响力和重视程度比IR要高得多。而IR除了研究机构的积极主动重视之外,并没有太多的外部资助,同时一般研究机构也没有专人从事这方面的工作,自然其影响力就要小得多。另外,受资助经费来源和项目成果归属问题的限制,IR和科学数据库的有机结合还有一定的困难。

但是,要看到是,大数据时代来临的趋势不可阻挡,如果IR的建设不能很好地跟上这股潮流的话,势必会被主流社会越拉越远。虽然单个机构存在人力、物力、财力方面的种种困难,但是通过前面对国内外IR建设和国内科学数据库建设的现状进行分析和对比不难发现:只要合理规划、有效利用内外部资源、辅以学科背景的分析和挖掘,还是能

表3 基础科学数据共享网数据库状况列表(2012年1月-2013年1月)

数据站点名称	记录数	访问量	访问量/记录数
天文学基础数据库	3375189849	453752	0.0001
微生物主题库	409183003	205956	0.0005
系统生物学多组学综合数据库	87312400	20260521	0.2320
高能物理数据库	6987105	7146693	1.0228
植物主题数据库	5782225	31355391	5.4227
毫米波射电天文数据库	5340246	1125996	0.2109
化学数据库	2646179	4702689	1.7772
青海湖区域研究专题数据库	1244518	1112771	0.8941
空间数据库	1148936	347775	0.3027
原子分子物理主题数据库	579301	119852	0.2069
中国病毒资源基础数据库	446298	2535086	5.6803
动物主题库	256859	4844274	18.8597
植物园主题数据库	138145	1750718	12.6730
聚变数据库	3671	536895	146.2531

注:只列出记录数10万条以上的数据库。

够在IR建设方面取得成果的。

结合大数据时代的特点^[8]，同时根据IR建设过程中存在的问题^[9-12]，对当前的IR建设做进一步的扩展，形成图2中的建设框架。整个IR建设由4个中心（科研活动中心、发布出版中心、教育学习中心和归档备份中心）组成，同时还需要有四大保障：政策、人才队伍、硬件基础和软件技术。

4个中心的主要特点如下所述。

科研活动中心：科研人员和学生进行试验、调查和数据分析等科研活动的实体或虚拟场所，包括实验室、野外台站、野外调查、数据分析、建模和数据挖掘等。这些活动中产生的中间过程信息，具有明显的机构知识库特征，对于今后从事类似研究工作的人员来说，都是非常有意义的。当前常提到的协同创新和虚拟科研环境即属于该类中心。该中心主要是需要开发适合专业需求的信息平台为各类科研活动提供虚拟的在线活动场所。

教育学习中心：机构知识库很重要的职能即为专业学生提供教学服务，包括教学资源、课堂笔记、野外实习笔记以及数字化本机构的历史文献古籍或前人的专业笔记，然后辅以数字化呈现。这种专业化学习中心不具有替代性，能够有足够的特色资源吸引各类用户。

归档备份中心：目前的IR建设起到的就是一个归档备份中心的作用，不过这还远远不够。更多的精力应该放在异质化资源的收集和整理方面，因

为同质化资源的加工本身就意味着重复建设。除了传统的文献类电子资源之外，还要多多加强其他类型信息资源的备份和归档，如多媒体材料、科研活动中过程数据的归档和永久性保存。需要有大容量的存储空间做支撑，同时有便利的存储平台，为用户提供归档、备份、不同权限间的信息共享的功能。

发布出版中心：该中心主要是通过以期刊文章、专著、科研报告以及科学数据库为表现形式的成果，对外形成发布出版中心。其中，科学数据库完全符合IR的特征，是IR的一部分，不但能够对外进行公益共享，还能够直接为出版业服务，是目前IR建设一个极为重要的补充。例如为了顺应科学研究和数字出版的发展趋势，中科院植物研究所和科学出版社于2013年4月签署了《战略合作协议》。除了在期刊、专著的数字出版领域进行合作，双方还计划面向市场推出一系列植物数字资源开发产品，如手机植物志、手机词典等，通过不同的终端设备或媒介，来展示和推送植物所IR资源和科学数据库中的优势资源，面向社会各类需求服务。

大数据时代下的IR建设有了上述框架，但是也要有实际的保障措施，才能真正落地和推行，这就是前面提到的4个保障。在当前的国内体制下，政策支持是基础，硬件设备是前提，人才队伍是关键，软件技术是核心和提升。4个保障是相互配合

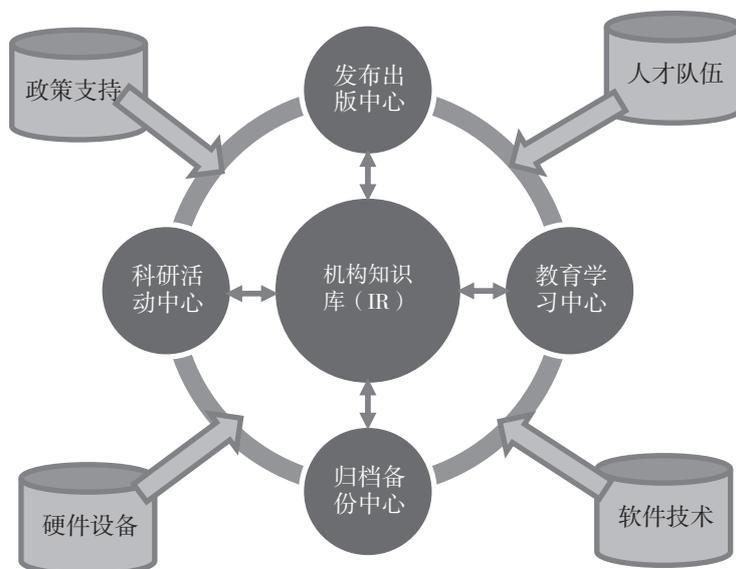


图2 大数据时代下的IR建设框架

的，缺少其中的任何一环，都可能会出现不好的局面。

5 实例分析

为了更好地解释上述提到的IR建设方案，这里以中科院植物所正在推动建设的数据平台工作为例进行说明。由于该平台将整合现有的植物所IR资源、数字化图书馆信息、出版期刊信息、科学数据库信息，大多数信息资源还是植物所各项科研活动的成果，其数据记录已经超过千万条以上，容量超过60TB。广义上来说，可以认为是大数据时代下的IR建设。

该数据平台的目标是一站式地整合植物所产生的各类信息资源（来源包括各类项目、各课题组、各行政部门和科研单元、研究人员和学生），并为他们提供符合工作需求的平台和存储空间，以及部分虚拟的服务器资源和FTP存储空间。数据平台自行开发的数据管理和共享平台主要利用MySQL、Drupal、Solr、GeoServer等开源软件技术，为广大在所职工和学生提供数据（物种名称、标本、文献、图片、生态观测数据、仪器实验数据、空间数据等）的在线存储、在线备份、在线管理、在线交流的平台，用户能够自己对数据进行权限管理，为不同用户分配权限，尽可能为用户提供便利，制定相应机制吸引用户的参与和分享。通过调研和意见征集，重点进行数字图书馆、数字实验室、数字标本馆、数字野外台站、数字科研平台、数字交流平台、数字管理平台等若干平台的建设和研发。这些平台产生和管理的数据如果能够直接对接到研究所的IR资源库，则无疑会大大丰富IR的数量和质量。

在具体的方案实施方面，主要是紧密结合植物所现有的硬件基础设施和信息化建设成果。硬件方面包括在2012年刚获批购置的“十二五”修购专项设备（16个刀片服务器和180TB的存储设备）。为了使设备的运行和服务正常化和制度化，该平台已经初步制定了硬件设备资源所内共享管理办法以及硬件资源申请表，在全所范围发送和回收了信息资源需求调查问卷，这些为平台更好地开展打下了良好的需求基础。信息化建设成果包括科技部基础条件平台项目——国家标本资源共享平台（NSII）及其相关的生物多样性信息网站群，如关注生物物种名录的Species 2000中国节点、关注植物图像的中国自然标本馆、中国植物图像库，关注植物标本的PE

标本馆、中国数字植物标本馆（CVH），关注生物多样性文献的BHL（生物多样性历史文献图书馆）中国节点，仅这些以植物所为主要力量建设的科学数据库其记录就超过千万条，容量超过60TB，而且数据增长正在逐年加快，大数据深层次分析的需求越来越紧迫。

在机构组织和人员配置方面，拟建的数据平台隶属于文献与信息管理中心，主要是因为文献与信息管理中心已经参与了多项科学数据库项目，并且也是植物所IR的负责部门，在两方面都有较好的经验。数据平台有一个专门技术人员负责管理，有较好的计算机技术水平和生物多样性信息学背景，特别是熟悉数据库处理技术和海量数据挖掘相关的工具，能够结合各研究部门和支撑部门的业务需求，配合研究人员对各类IR资源进行深层次的挖掘和展示。所内各相关部门的通力合作也是必要的支持条件之一。否则，很多的资源连汇总都成问题，更别提统一分析了。因此，研究所在这方面也要有相应的政策支持。

在管理机制和理念方面，数据平台负责人除了积极学习各类相关技术之外，还入选中科院计算机网络信息中心与微软研究院联合举办的“云资源学术合作计划”。因此，通过这个机会学习和了解微软的Windows Azure云平台的设计理念和管理机制，也可以较好地将这些经验应用到植物所的大数据管理方案和实施工作中。所内用户根据自己的需求，进行不同类型硬件资源的需求申请，负责人通过统一的协调，对这些申请进行落实，并在后期的利用和管理过程中进行相应的协助。

6 结论和讨论

我国机构知识库的建设既取得一定的成果，也遇到了一系列问题。大数据时代的到来为机构知识库带来了巨大的机遇和挑战。根据我国IR建设现阶段的情况，结合大数据时代下的IR建设框架，今后可以从以下几个方面加大力度，以使IR建设取得更好的成绩。

（1）加强顶层汇交平台的双向作用。如中国科学院机构知识库服务网格平台（<http://www.irgrid.ac.cn/>）除了对各个IR进行资源收割之外，还要加强资源从汇交平台到各个IR的反向推送作用。很多资源，特别是开放的资源，总平台可以利用技术和人员的优势进行总体汇总，再推向各个机构的平台。

另一方面,有的文章可能是多家单位共同发表的,如果有一家单位发布在自家的IR上,总平台就可以设法推送给其他几个合作单位的IR。这会极大地提高各机构知识库的信息量和交互性,吸引更多的用户来使用。

(2)加大异质化资源的建设,注意各方的实际权益。一些通用的期刊和文献,能够比较方便地在CNKI、万方和维普上面检索得到,用户就很少再到各个单位的机构知识库上面来搜索,毕竟信息量不是同一个级别的。但是,有些特色的资源是这些通用的文献平台所不具有的,如一些古籍、专业课题验收报告、内部课件、项目汇报材料等。这样才会使资源做出新意,不然很难有所突破。可以适当鼓励跨机构的合作共建IR资源。

(3)加大与科学数据库平台之间的合作与联合。借助目前国家在基础条件平台上的扶持力度和大众的关注度,IR建设完全可以参与进去,承担部分课题任务,前提就是要资源互补。在这个过程中,IR可以从科学数据库平台上吸引优质资源和用户,自身的资源也可以在数据库平台上得到充分展示。更长远地,可以加强与通用搜索引擎(如百度、Google等)和国际开放IR平台的合作,使更多平台上都能搜索到IR的资源,产生更大的影响力。

(4)加强IR资源在科研和教学活动中的推送服务,构建专业信息平台。通过专题订阅、关键字订阅、最新优质论文提醒和推送、著名期刊杂志的定制等服务进行信息推送。目前应用比较多的DSpace平台,在功能定制和交互性等方面都还存在一定的问题,特别是对于以用户为中心的未来IR建设不太适应,可以适当鼓励有基础和人力IR研发符合个性化需求的信息平台,然后加以推广。

(5)多方寻求政策支持,使IR融入机构日常工作。当前体制下,只有纳入机构的日常管理工作,才能得到更稳定的长久支持,包括人、财、物等方面。因此,可以依托传统图书馆或信息中心,扩展IR的工作,努力成为机构信息化工作的一部分。不然,很可能会昙花一现。

参考文献

- [1] Institutional repository [EB/OL]. [2013-05-20]. https://en.wikipedia.org/wiki/Institutional_repository.
- [2] 何琳.OpenDOAR和机构知识库发展现状[J].图书馆工作与研究,2009(2):30-33.

- [3] 田质兵.全球开放存取知识库发展现状分析与启示[J].图书馆建设,2010(12):43-47.
- [4] The Directory of Open Access Repositories-Open-DOAR [EB/OL]. [2013-05-20]. <http://www.opendoar.org/>.
- [5] 张冬荣,祝忠明,李麟,等.中国科学院机构知识库建设推广与服务[J].图书情报工作,2013,57(1):20-25.
- [6] 张晓林.颠覆数字图书馆的大趋势[J].中国图书馆学报,2011,37(5):4-12.
- [7] 许哲平,覃海宁,马克平,等.基于专业日志分析的科学数据库平台用户活动研究——以中国数字植物标本馆(CVH)为例[C]//中国科学院科学数据库办公室.科学数据库与信息技术论文集(第十一集).北京:科学出版社,2011:205-211.
- [8] Michael Schroeck,Rebecca Shockley. Analytics: The real-world use of big data [EB/OL]. [2013-05-20]. <http://www-935.ibm.com/services/us/gbs/thoughtleadership/ibv-big-data-at-work.html>.
- [9] 张晓林,张冬荣,李麟,等.机构知识库内容保存与传播的权利管理[J].中国图书馆学报,2012(4):46-54.
- [10] 朱梦皎,武夷山.中、日、印高校机构知识库建设现状比较分析[J].图书与情报,2012(6):69-72.
- [11] 曾苏,马建霞,汤天波,等.国内科研机构 and 高校机构知识库规划建设现状与问题研究[J].现代图书情报技术,2009(1):50-57.
- [12] 李霞.关于机构知识库构建的若干思考[J].中国科技资源导刊,2009(6):47-52.

欢迎订阅 欢迎赐稿

中国科技资源导刊

(双月刊)

《中国科技资源导刊》是中国科学技术信息研究所与南京大学共同主办的科技管理类学术性期刊,是中国科技核心期刊(中国科技论文统计源期刊)。主要刊登科技资源(尤其是科技物力资源、科技信息资源和科技人力资源)管理领域的学术论文、研究报告、综述评论,宣传和探讨科技资源管理的战略政策,探索和揭示科技资源管理领域的基本原理和规律,展示技术创新实践经验等,促进我国科技资源管理领域的理论研究与实践管理水平的不断提升,为科技资源管理者和研究者提供高水平的学术交流平台。

《中国科技资源导刊》为双月刊,大16开112页。刊号:CN11-5649/F。每期定价15元,全年90元。欲订阅者请与编辑部联系。

通讯地址:(100045)北京市西城区三里河路54号《中国科技资源导刊》编辑部
电话:010-68514086/68571416 联系人:莫妍
E-mail:zgkjzydk@istic.ac.cn 网址:<http://www.zgkjzydk.com.cn>