

5-2016

Institutional Repositories: Characteristics, Benefits and Software Platforms

Ping Fu

Central Washington University, pingfu@cwu.edu

Liu Hua

Shanghai University

Follow this and additional works at: <http://digitalcommons.cwu.edu/libraryfac>

 Part of the [Library and Information Science Commons](#)

Recommended Citation

Fu, P. & Hua, L. (2016). Institutional repositories: Characteristics, benefits and software platforms. *Research on Library Science*, 531-36.
DOI: 10.15941/j.cnki.issn1001-0424.2016.05.007

This Article is brought to you for free and open access by the Library at ScholarWorks@CWU. It has been accepted for inclusion in Faculty Scholarship from the Library by an authorized administrator of ScholarWorks@CWU.

机构库：特性、利益和软件平台选择

傅平 刘华

【摘要】立足机构库理论研究和应用发展现状，就机构库的定义、基本特性、发展现状、受益者和软件平台进行系统梳理，希望给国内同行和正在建设机构库的图书馆带来启发和参考。以美国华盛顿大学图书馆机构库的建设为案例，加以具体说明。机构库不仅可集中管理一个机构的所有学术成果，并可免费开放获取，从而扩大机构和学者的知名度和影响力。机构库软件主流平台有4~5种，用户可以根据自己的实际情况加以选择。每个软件平台各有优缺点，但基本能满足用户的需求，各机构可以根据自己的侧重点和优先选项决定自己的软件平台。

【关键词】机构库 软件平台 美国华盛顿大学 图书馆

Abstract: Based on the theories and practices of Institutional Repository (IR), the paper reviews and summarizes IR's definitions, features, development, benefits and software systems selection process. Hopefully it will provide insights and bring inspiration to those domestic libraries and interested parties planning to work on this area. The authors take Central Washington University Library as a case study to illustrate an approach to implement an IR within a short time frame. IR has many merits including collecting and showing an institution's scholar outcomes. There are four or five main software platforms for institutions to choose. An institution can decide to adopt one depending on its own interests and priorities.

Key words: institutional repository software platform Central Washington University (CWU) library

DOI:10.15941/j.cnki.issn1001-0424.2016.05.007

1 机构库定义

机构库是一个以数字化方式收集、保护和传播一个机构，特别是研究机构的知识成果的在线档案库。广被学术界采纳的一则定义是 R. Crow 于 2002 年提出的“机构库是用来捕获和保护一个或多个大学社区所有知识产出的数字化集合”。R. CROW 进一步指出：机构库拓宽了学者获取研究成果的手段，变革了学术交流系统，加强了学术领域对学术研究的控制，降低了商业期刊的垄断力量，从而增加了竞争性，对有机构库的机构和图书馆来说，既缓解了经济上的压力，又提高了机构的相关知名度^[1]。

再则，机构库有潜力成为一个大学质量的可测量指标，并有潜力成为展示与其研究活动相关的科学数据、社会和经济等方面的可测量指标，从而增加了机构的知名度、地位和公共价值。

另外一个定义是 C. Lynch 于 2003 年提出的：一个大学机构库是该大学提供其社区成员来管理和传播这个机构及其社区成员创建的数字资源的一整套服务，它充分体现在一个机构对这些数字资源的承诺，包括长期合理保存、组织、获取或传播。按照 C. Lynch 的说法，基于大学的机构库是一组服务，即一个大学提供其社区成员来管理和传播这个机构及其社区成员创建的数字资源。在大学环境中，机构库可以为教师的科研成果、学生的学位论文、电子期刊、数据集等提供一个集中展示和保护的平台^[2]。

2 机构库发展现状

机构库是目前图书馆和开放社区（open access communities）领域的一个热门话题。根据开放获取知识库网站（ROAR, Registry of Open Access Repositories）的统计，截至目前，全球在 ROAR 注册登记的机构库有 3 606 个，其中非洲有 93 个、亚洲有 591 个、欧洲有 1 293 个、北美有 649 个、大洋洲有 98 个、南美有 288 个。与 2007 年 9 月只有 900 多个机构库注册（Dora, 2007）相比，8 年来在 ROAR 注册的机构库翻了 4 倍多。从下表可看出，2012 年是机构库建设的高峰年，注册的机构库达 653 个。

目前机构库数量占前三位的国家是美国（693 个）、英国（231 个）、德国（199 个），中国有 90 个。机构库构

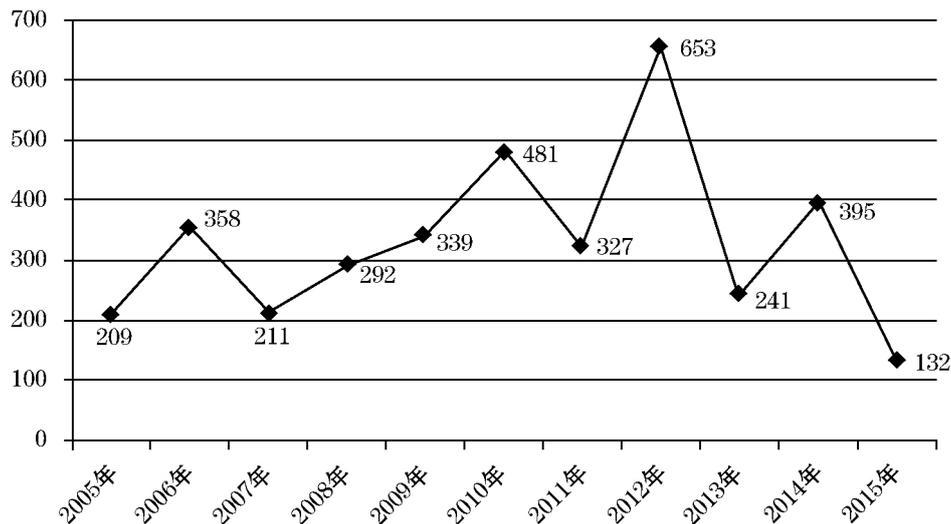


图1 机构库年度注册数量统计表 (数据截取日期 2015-08-26)

建平台和软件超过 30 个, 前 5 位的软件平台分别是 DSPACE (1 581 个)、EPrints (568 个)、bepress Digital Commons (366 个)、Fedora 含 Fez (68 个)。

目前, 已注册的机构库已逾 4 000 多个, 以大型研究机构的机构库为主, 占到 70% 以上, 其次是数字期刊和学位论文机构库, 机构库类型大致如下:

表1 机构库类型分类统计

机构库类型	英文原名	数量
研究型机构与部门机构库	Research Institutional or Departmental	2 714
研究型多重机构库	Research Multi-institution Repository	78
研究型交叉机构库	Research Cross-Institutional	264
数字杂志出版	e-Journal/Publication	113
数字学位论文库	e-Theses	292
数据库和索引	Database/A&I Index	57
研究数据	Research Data	24
开放和关联数据	Open and Linked Data	23
学习和教学库	Learning and Teaching Objects	60
展示库	Demonstration	20
其他	Other	385
合计		4 030

3 机构库的基本特性

根据 Crow^[3]、Shearer^[4]和 Ware^[5]的研究成果, 机构库主要集聚了以下特性:

信息形式是数字化的: 机构库收集的主要是数字化信息, 这与大学档案馆的职能并不重叠, 档案馆是收集涉及到大学的所有类型的东西。

定义是机构层面的: 不同于传统的特定学科和数字图书馆, 机构库捕获的是该机构的所有成员的研究成果。

特定收集范围是学术资料: 从这个范畴来讲, 学术著作包括 (预印本、出版后文稿、技术报告等)。

收藏内容是累积的、永久的: 这是指机构库数据可长期保护和使用的特性。

服务是开放存取的: 机构库的一个重要特色, 是自由和开放存取其中的内容。

技术上的互操作性: 这指的是机构库既可以作为各种形式数据提供者, 也可以作为服务提供者来加工和增值

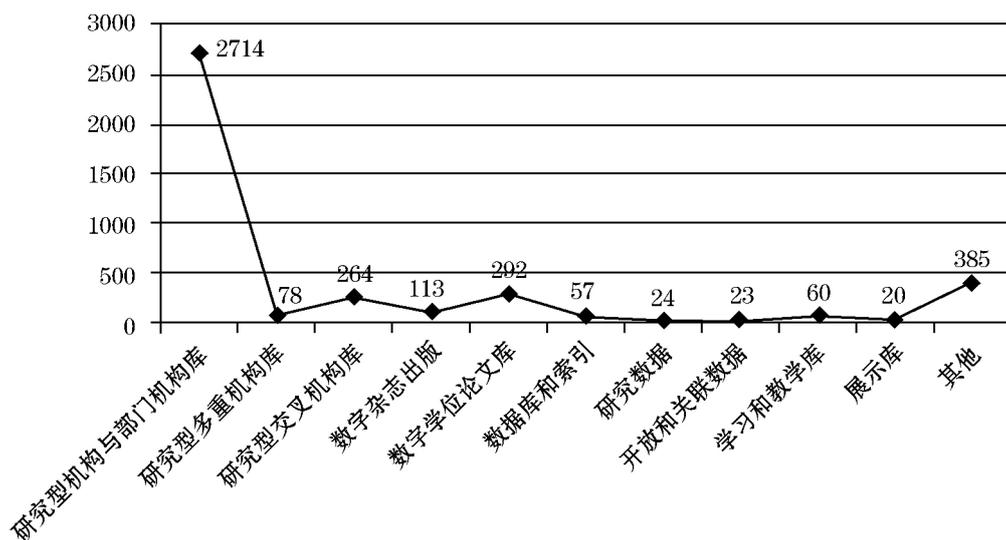


图2 机构库类型分类统计

数据，支持 OAI-PMH 协议。

自我存档的能力：大多数机构库要求作者或与作者相关的人员直接存入他们的内容。

换句话说，传统模式如期刊平台需要支付费用来访问阅读它们的出版物，而机构库则对需要存取文献的研究者开放，即实现了自由和即时的在线访问^[6]。机构库对学术交流起到了补充、增强的作用，并提供额外的获取途径。

机构库还有一个特点是多样性。尽管机构库都以数字化的形式体现^[7]，但各个库的政策、工作流程、功能、服务和元数据可能不同，机构库运营模式和资金来源也不同^[8]，机构库所包含的内容也各不相同。

4 谁得益于机构库

研究人员、机构和国际社会都能从机构库获益。新加坡管理大学图书馆 2008 年发布了一份关于机构库研究的报告^[9]，他们总结分析了英国联合信息系统委员会（JISC, Joint Information Systems Committee）所支持的机构库项目后得出以下结论：

4.1 机构库首先使研究人员受益

它可以提高作者的研究成果被发现和引用的机会及该作者在该机构的影响力。

在快速发展的领域如电子技术，研究人员可以通过机构库及时发表预印本（preprint），相对于同行评审的论文缓慢，以证明他们是第一时间做了这方面的研究，并及时获得反馈意见。

帮助管理和存储与研究相关的数字化的研究数据和内容。

帮助研究人员管理资助拨款机构对出版的要求。

提供规范的学术化的简历和所发表的论文清单。

允许创建个性化的出版物清单。

提供论文被下载和使用的数据，使研究人员能够随时知道具体的论文被下载和点击次数。

通过各种链接增加其他机构库里相关论文被发现引用的机会。

4.2 机构库使研究机构受益

机构库可以与其他大学的系统进行互联，通过信息共享以达到效率最优化。

增大了机构的知名度和威望，当然这取决于收录的内容。

机构库内容很容易被本地及全球检索。

提高机构对知识产权问题的认识和管理，并促进机构对有关版权的保护。

含有高品质内容的机构库可以作为一个“橱窗”或营销手段来吸引员工、学生，并获取研究机构的认可，从而争取更多的研究资助。

机构库可以存储其他类型的内容，如不一定公开发表的有时也称为“灰色文献”的内容。

以前分散在虚拟学习环境中的学习和教学材料可以集中存储在机构库，以增加潜在的重复使用、再利用和共

享的机会。

机构库可作为一个机构向所属研究评估机构 (Research Assessment Exercise, RAE) 提交绩效报告的一个重要工具。

从长远来看, 机构库可以节约大量成本, 存储内容越多, 越节约成本。与静态的网站相比, 机构库更具动态灵活性, 数据库采用标准化元数据, 安全可靠, 便于检索和长期保存。

4.3 机构库使全球学术界受益

由于机构库具有 OAI 兼容的元数据捕获机制, 从而可以促进学术信息自由交流, 增加合作研究的机会。

增强公众对研究工作和活动的认识。

总的来说, 研究人员试图尽可能广泛地传播自己的研究并希望得到承认和引用。机构则寻求学术地位和威望。国际社会则希望对研究人员的职责有更多的了解。机构库使研究成果在全世界范围公开出版, 在互联网上可以搜索和共享, 其显著的社会效益是无可争辩的。

5 机构库软件平台评估、选择和实施

加拿大英属哥伦比亚大学对四种常用软件平台 DSpace, EPrints, Digital Commons, Islandora 和 Hydra 在以下三个领域作了详细分析和比较^[10]:

安装/运行管理包括元数据、互相开放、可操作性;

内容管理包括禁止条款、版本和存储、统计、文件格式和批输入、用户界面、搜索;

服务支持;

结合 Castagné 的研究, 本文作者得出的结论是:

DSpace 是开源软件, 已被证明是自 2002 年推出以来, 最为广泛使用的机构库软件平台。2013 年 12 月最新发布了 4.0 版本, 增强了后台管理和对研究数据的更强有力支持, 所以 DSPACE 仍处于强有力的竞争地位。

EPrints 是开源软件, 主要优点是它友好的用户界面, 易于实施。但是, 从其他系统移植到 Eprints 比较困难。EPrints 成本较低。Godfrey 把它描述为“对一个财力和 IT 力量有限的大学来说, 是一个快速实施的理想的解决方案”^[11]。

Digital Commons 是商用软件, 用户界面非常友好, 可以降低学校的技术责任, 可使用户得到全面的支持, 实施容易。Digital Commons 是一个有吸引力的选择, 内置同行评审期刊的出版系统和会议, 事件管理系统还可以创造更多的附加价值, 附加选项如 Selectworks 能建立作者个人的出版物专集和规范化的简历, 因此会吸引更多作者使用机构库。

近年来, Fedora3 已经开始走下坡路。但 Fedora4 的持续发展, 吸引了来自机构库社区新的兴趣^[12]。Fedora 的强有力的功能使 Hydra 和 Islandora 在机构库软件平台开辟了新的应用前景。与 Archivematica 搭配, 这两个平台可以为用户提供一个单一的解决方案, 满足大存储、显示和保存的需求。其主要缺点是, 因为两者都是相对较新的开源软件工具, 和 DSpace 相比, 还没被更多的用户完全测试。

由 Jean Gabriel Bankier 和 Kenneth Gleason 共同为联合国教科文组织撰写的机构库软件比较一文^[13], 则从体系结构 (Infrastructure)、前台设计、内容组织和控制、内容发现、出版工具、报表、多媒体、社交特点和通知提示 (Social Features and Notifications)、用户登陆管理、数据可取性 (Accessibility) 和长期保存等角度对 Digital Commons、Dspace、Eprints、Fedora、Islandora 这 5 个主流平台的 11 个方面进行了更详细的分析和比较。一般认为, 这篇报告可以作为用户选择软件平台的参考指南。

6 一个案例

本文作者之一所在的美国华盛顿大学从 2011 年 9 月成立了一个专门委员会, 开始对现有的机构库软件进行调查研究和评估。这个委员会在比较了目前在美国最流行的三个机构库建设平台 DSPACE, EPRINT 和 bepres Digital Commons 以后, 向图书馆推荐 bepres Digital Commons 作为机构库的平台软件。下面是采用 bepres Digital Commons 的主要依据:

支持云计算环境;

客户服务优良;

柔性的机构库结构;

简单的高质量的 DublinCore 元数据;

客户可以个性化的元数据;
 厂商负责系统升级;
 管理员可以进行系统设定;
 支持标准界面, 用户界面非常友好直观;
 支持移动设备;
 个性化的出版设计——支持期刊出版和会议组织出版;
 支持开放获取出版 (Open access publishing);
 提供 IP 范围、用户和域名控制;
 支持各种文件格式;
 支持图像及多媒体;
 支持嵌入式视频音频连续播放;
 容易被 Google, Google Scholar 等搜索引擎发现和优化;
 引文链接 (Link resolver);
 遵守 OAI-PMH 可以被其他系统采集;
 可以被发现平台集成;
 数字对象唯一标识 (DOI) 和永久 URL;
 实施简单。

中华盛顿大学于 2013 年 12 月正式与 bepress 签订了合同。2014 年 1 月 2 日正式成立项目组, 启动机构库的建设。只花了 2 周时间, 他们和厂商就一起设计好了用户界面。bepress 又通过远程会议, 向中华盛顿大学项目组展示了 Digital Commons 的样板工程。在此基础上, 中华盛顿大学项目组拟定了自己的社区 (Communities) 和系列 (series)。社区 (Communities) 和系列 (series) 是 Digital Commons 的术语。比如一个二级学院的著作集合可以看作一个社区 (Community), 二级学院下面的每一个系的学术著作集合可以看作一个系列 (series)。他们把社区 (Communities) 还扩展到民间社会团体, 支持民间社团的数字化项目, 比如华盛顿州当地的口述历史等等。他们计划把现有的数字档案库全部移植到机构库。学生学位论文和优秀的学生科研项目报告也是他们关注的一个重点。项目组同时和中华盛顿大学教授们主办的刊物、会议召开者联系, 向他们介绍期刊和会议管理出版功能, 结果, 有教授主编的两个期刊和一个学生主编的期刊, 以及一个全校学生科研报告的年会对机构库有浓厚兴趣, 现在正在做期刊和会议的界面设计和数据移植。3 月 14 日机构库实现了初步开通上线。允许作者自己存放自己的著作。其间他们通过公共宣传部门、社交软件平台、介绍会、海报、电子邮件、网站等多种形式向学校师生推荐机构库平台。他们计划在 4 月 14 日正式启用机构库。整个实施过程简单顺利快速。厂商也给予了大力支持和培训。到 2015 年 8 月 26 日, 机构库已经收录了 12 544 篇中华盛顿大学教师和研究人员文章, 总下载次数为 23 778 次。经过一年多的运行和实践, 学校已经看到机构库给学校学术研究和扩大学术影响带来的好处。学校当局已经同意图书馆专门设立一个机构库馆员的岗位, 来负责与教师和员工的交流以及机构库的日常管理运行。详细情况可以查看网址 digitalcommons.cwu.edu。

7 结束语

机构库经过十几年的发展, 到今天已经进入了成熟发展的阶段。包括中国在内, 很多学术机构都看到了机构库的重要性。机构库不仅集中管理一个机构的所有学术成果, 而且通过开放平台让机构外的更多的学者和公众免费获取、下载和应用。从而大大增加了机构和学者的知名度和影响力。机构库软件平台虽然多样化, 有开源软件、商用软件, 但主流平台就有 4~5 种, 用户可以根据自己的实际情况加以选择。每个软件平台各有优缺点, 但几乎都能满足用户的需求, 每个机构可以根据自己的侧重点和优先选项决定自己的软件平台。

本文作者结合所在机构的实施实例, 就机构库的定义、发展现状、性能、受益者和软件平台作了简单系统的综述。如果本文能给将要或正在建设机构库的图书馆和同行带来一些启发的话, 就达到了我们写这篇文章的目的。

注释

[1][3]Crow, R. The Case for Institutional Repositories: A SPARC Position Paper. ARL Bimonthly Report, 2002, 223: 1-37.

[2]Lynch, C. Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age. ARL Bimonthly Report, 2003, 226: 1-7.

- [4] Shearer, K. Institutional Repositories: Towards Identification of Critical Success Factors. *The Canadian Journal of Library Science*, 2002/2003 (3): 90-108.
- [5] Ware, M. Institutional Repositories and Scholarly Publishing. *Learned Publishing*, 2004 (2): 115-124.
- [6] Olivier, E. The Efficacy of Institutional Repositories: Reflections on the Development of a Personalised Collection on UPSPACE. *Perspectives in Education*, 2007 (1): 123-128.
- [7] Smith, K. Institutional Repositories and E-Journal Archiving: What Are We Learning. *Journal of Electronic Publishing*, 2008, 11 (1).
- [8] Swan, A. & Awre, C. Linking UK Repositories: Technical & Organisational Models to Support User-Oriented Services Across Institutional & Other Digital Repositories. Tech. rep. 2006, JISC, London.
- [9] Martin, P. & Page, R. SMU Institutional Repository: Knowledge Dissemination of Research and Scholarship. Singapore Management University. [EB/OL]. [2014-10-22]. http://ink.library.smu.edu.sg/library_research/4/.
- [10] Castagné, M. Institutional repository software comparison: DSpace, EPrints, Digital Commons, Islandora and Hydra. University of British Columbia. [EB/OL]. [2014-10-20]. http://circle.ubc.ca/bitstream/handle/2429/44812/Castagne_Michel_LIBR596_IR_comparison_2013.pdf?sequence=1.
- [11] Godfrey, N. Informal Comparison of Some Institutional Repository Solutions. Metalogger. [EB/OL]. [2014-10-22]. <http://metalogger.wordpress.com/2008/10/19/comparing-some-institutional-repository-solutions/>.
- [12] Shin, E. & Woods, A. Building the future of Fedora. [EB/OL]. [2014-10-22]. http://or2013.net/sites/or2013.net/files/slides/building_fedora_futures_0.pdf.
- [13] Gabriel, J. B. & Gleason, K. Institutional Repository Software Comparison. Educational, Scientific and Cultural Organization, United Nations. [EB/OL]. [2014-10-20]. http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/news/institutional_repository_software.pdf.

傅平 美国华盛顿大学图书馆技术部主任。

刘华 上海大学图书馆研究馆员。

(上接第22页)

- [3] 方婧, 桂思思, 陆伟. 美国图书情报硕士教育质量保障体系调查与分析. 第四届中美数字时代图书馆学情报学教育国际研讨会论文集[C]. 武汉: 武汉大学出版社, 2015: 111-124.
- [4] 桂思思, 方婧, 陆伟. 中美图书情报硕士专业学位教育质量保障体系比较分析及启示. 第四届中美数字时代图书馆学情报学教育国际研讨会论文集[C]. 武汉: 武汉大学出版社, 2015: 274-283.
- [5] 希蒙斯学院[EB/OL]. [2015-12-08]. <http://www.simmons.edu/>.
- [6] 孙占山. 国外图书馆员职业资格认证情况比较[J]. *新世纪图书馆*, 2008 (4): 89-91, 46.

王燕红 上海大学图书情报档案系, 硕士研究生。