

Overlay 期刊系统构建模式研究

祝忠明 王楠 卢利农

(中国科学院国家科学图书馆兰州分馆 兰州 730000)

【摘要】通过对相关研究项目和应用的调研,对基于知识库软件和基于期刊出版软件构建 Overlay 期刊系统的两种主要模式的技术特点和应用趋势进行简要分析和说明,并提出一种基于 OJS 和 OAI 收割服务相结合以构建 Overlay 期刊系统的基本框架。

【关键词】Overlay 期刊 Overlay 期刊系统 知识库 构建模式

【分类号】G250

Construction Models for Overlay Journal Systems

Zhu Zhongming Wang Nan Lu Linong

(The Lanzhou Branch of the National Science Library, CAS, Lanzhou 730000, China)

【Abstract】 From an investigation of some related studies and projects, the paper concludes two main construction models for building Overlay journal systems, which are digital repository system based model and electronic journal system based model. The technical features and usage scope for each model are also presented. Based upon a trend analysis, the paper proposes a construction framework for Overlay journal systems, which suggesting a method of extending OJS with an OAI harvesting service plugin.

【Keywords】 Overlay journal Overlay journal system Digital repository Construction model

1 引言

Overlay 期刊 (Overlay Journal) 是学术出版和交流过程逐步向解构化或分布式方向演化发展过程中出现的一种新的出版模式^[1]。尽管 Ginsparg^[2]早在 1996 年即提出了 Overlay 期刊的概念和设想,但直到开放获取运动的兴起,Overlay 期刊才真正得到了关注和应用。因此,目前比较普遍的看法是^[3,4],Overlay 期刊是一种基于开放获取知识库的具有学术质量保证的开放获取期刊,它按照一定的内容筛选标准,从一个或多个数字知识库(机构知识库或学科专题知识库)中选择论文,经过编辑评审或同行评审,以在线期刊的形式进行发布和提供服务。与一般的开放获取期刊相比,Overlay 期刊并不直接发布论文的全文内容,而只发布和提供论文在来源知识库中的链接,提供基于知识库内容的增值服务。这种运作模式使得 Overlay 期刊成为一种低投入、易实现的开放获取期刊形式,也为文献信息机构介入数字学术出版领域、深度组织与开发利用开放获取知识库的内容提供了新的机遇和模式。在 Overlay 期刊的建设和应用中,如何构建可靠的 Overlay 期刊建设的平台,支持与相关知识库之间的良好互操作,是 Overlay 期刊顺利建设和运行服务的重要保障。本文将对 Overlay 期刊系统构建的主要模式进行分析和比较研究,以期为国内 Overlay 期刊的建设和发展提供借鉴和参考。

收稿日期: 2009 - 03 - 02

收修改稿日期: 2009 - 04 - 01

2 当前 Overlay 期刊系统构建模式分析

根据对一些典型 Overlay 期刊建设平台和应用研究项目的调研和分析,目前 Overlay 期刊系统的构建模式主要有基于知识库软件的模式(即以数字知识库软件为主构建 Overlay 期刊的编辑和出版平台)和基于期刊出版系统的模式(即以电子期刊出版平台为主构建 Overlay 期刊的编辑和出版平台)。

2.1 基于知识库软件系统的构建模式

Overlay 期刊系统基于成熟的数字知识库软件,根据 Overlay 期刊的编辑出版流程和组织形式,对其进行适当的定制扩展后,形成 Overlay 期刊的建设平台。如 Web Science Overlay Journal^[5]即采用 EPrints 3 作为其出版和发布平台。目前来看,以这种模式构建 Overlay 期刊系统的应用相对较少。采用这种模式构建的 Overlay 期刊平台,其内容编辑、组织和发布功能的实现都依托知识库软件固有的功能来实现。在编辑工作流的实现方面,基于知识库软件内在的提交 workflow 机制,将一定的编辑 workflow 关联和映射到提交 workflow 上,建立起支持作品的筛选、提交、审核等步骤的编辑 workflow 控制。在内容组织和发布方面,主要利用知识库软件支持按照时间层级(如年、月的层级)进行内容组织的功能,模拟完成“按期”形式的内容组织和发布,向用户提供基于时间序列组织的期刊浏览服务。依托知识库软件提供的丰富的其他形式的用户服务,可以使 Overlay 期刊向用户提供多途径浏览、多条件检索和全文检索、以及内容推送服务等多方面的功能和服务支持。

一些项目则从基于数字知识库的学术交流体系构建方面,对构建 Overlay 期刊的技术流程和模式进行了研究和试验。如美国康奈尔大学和洛斯阿拉莫斯国家实验室的合作项目 Pathways^[6],试验性地设计和构建了基于跨知识库内容构建 Overlay 期刊的流程和工具:首先,它基于 OAI-PMH 和 Nutch 建立了一个开放知识库内容搜索系统,供 Overlay 期刊编辑者检索和筛选目标文章;其次,使用开源数字知识库软件 Fedora 作为 Overlay 期刊编辑和发布的平台。同时,它提出了一个跨知识库的互操作层设计,使得任何符合 Pathways 数据模型和接口的知识库和 Overlay 期刊系统之间,在 Live Clipboard^[7]等 Web 应用间交换数据技术的支持

下,可比较容易地实现双向的数据交换。如图 1 所示。

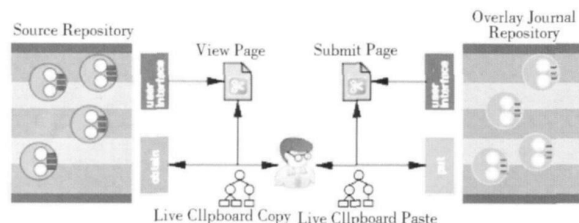


图 1 Pathways 的 Overlay 期刊建设流程示意图

2.2 基于电子期刊出版系统的构建模式

通过知识库系统与期刊出版系统之间基于流程的互操作能力的扩展,将作者投稿和学术质量控制机制(如同行评审)引入 Overlay 期刊的创建过程中,使得基于开放知识库能够容易地构建起具有学术质量保证的 Overlay 期刊。知识库系统的重点在于为作者提供快速登记和发布学术成果的渠道,保障作者学术观点和学术成果优先权的有效建立,并提供学术成果的存档功能。期刊出版系统基于知识库系统,重在提供学术质量控制的机制支持和编辑出版功能,并实现为作者提供学术成果的公共接受和认可的证明机制。知识库系统和期刊出版系统之间,基于一定的编辑出版流程,实现数据交换和互操作。

JISC 先后支持的项目 OJMS (Overlay Journal Infrastructure for Meteorological Sciences)^[8]和 RDJA (Repository Interface for Overlaid Journal Archives)^[9]是对这种模式进行了比较深入的研究和应用实践的代表。特别是 RDJA,它试图通过把支持 Overlay 期刊建设过程中知识库系统和期刊出版系统之间的互操作过程以公共抽象接口的形式一般化和规范化,设计和发布了 RDJA API 接口规范^[10],使任何支持该接口规范的知识库系统和期刊出版系统,可以很容易建立互操作关系,实现基于知识库的 Overlay 期刊建设。在 RDJA API 的支持下,属于投稿性质论文的状态在知识库和 Overlay 期刊之间是透明的,期刊对通过同行评审处理的论文,直接从知识库中抽取论文的元数据,附加指向知识库的链接后,在期刊系统中予以发布。

RDJA 目前已提供了对上述接口规范的参考实现^[11]。在期刊出版系统方面,选择开源软件 OJS 进行了接口扩展,以支持从知识库中接收作者的投稿,并优化了评审和出版过程。在知识库系统方面, RDJA 通过对 EPrints 3 进行扩展,增加了 RDJA API 接口,使

基于 EPrints 软件建立的知识库可以支持向 Overlay 期刊提交论文。这两种经过 RDJA 修改和支持 RDJA API 的期刊出版系统和知识库系统都提供公开下载和使用。同时, RDJA 按照 RDJA API 的技术路线, 也进行了示范性 Overlay 期刊的出版试验。通过与 arXiv 知识库合作, 使 arXiv 支持 RDJA API, 并利用支持 RDJA API 的 OJS 期刊出版系统, RDJA 目前编辑和出版了基于 arXiv 学科知识库的 Achyon Continuous, Dark Matter Continuous, Cosmology Continuous 等多种示范性的 Overlay 期刊^[12]。

2.3 两种模式的比较分析

基于知识库软件系统的构建模式, 通常情况下需要做定制和扩展开发工作, 使知识库软件能够支持一定的编辑出版流程, 以及支持 Overlay 期刊能够比较符合传统“期刊”的样式进行内容组织和发布。基于期刊出版系统的构建模式, 充分利用期刊出版系统在编辑出版流程方面的支持能力, 支持 Overlay 期刊根据需建立优化的内容筛选、评审、编辑、组版和发布流程, 按期组织和出版内容, 能够比较容易地实现将作者投稿和同行评审机制引入 Overlay 期刊的组织和出版过程, 以完全符合传统(电子)期刊的出版模式, 创建和出版学术质量得到严格保证的 Overlay 期刊。而且, 就两种模式的实际应用来看, 大多数的 Overlay 期刊都采用了基于期刊出版系统的模式, 而采用知识库系统作为编辑出版平台的应用则要少得多。特别是随着 RDJA API 的发布, 以及在对开源知识库软件 Eprints 和电子期刊出版系统 OJS 所做的支持 Overlay 期刊建设接口的扩展工作, 对于促进开放知识库系统在支持 Overlay 期刊应用的标准化和规范化方面将起到积极的推动作用, 有助于基于扩展后的 OJS 系统很容易地与相关开放知识库系统进行交互, 快速地构建支持 Overlay 期刊建设的技术平台。

因此, 基于期刊出版系统, 特别是开源的电子期刊出版系统进行 Overlay 期刊技术平台的构建, 从目前及长远来看, 都是比较简便易行的选择。

3 Overlay 期刊系统构建的现实模式

根据上文的分析结果, 基于开源电子期刊出版系统的模式将是 Overlay 期刊系统构建模式的基本发展趋势, 并从某种程度上符合和强化了 Overlay 期刊的开

放存取特性。从长远来看, 开放获取的知识库和 Overlay 期刊系统之间能够在一定互操作标准和规范基础上, 实现无缝的数据交换和建立协同的工作流, 对于 Overlay 期刊的发展将起到极大的促进作用, 这也正是 RDJA API 推出的用意所在。但是, 当前大部分知识库并不支持 RDJA API, 仅有 arXiv 在与 RDJA 的合作过程中, 提供了对 RDJA API 的支持^[11]。在 Overlay 期刊还不能成为一种普遍性应用的条件下, 开源知识库软件和电子期刊软件是否能够逐步支持 RDJA API 还处于不确定的状况。尽管 RDJA 为开源知识库软件 EPrints 和电子期刊软件 OJS 扩展了 RDJA API 接口, 但随着 RDJA 项目的结束, 后续的支持也是一个未知数。因此, 本文通过综合 Overlay 期刊平台建设的现有模式和一些应用案例所形成的良好实践经验, 给出一种当前比较可行的 Overlay 期刊平台建设基本框架和模式:

(1) Overlay 期刊平台的建设以开源电子期刊系统为基础, 通过扩展对开放知识库内容进行筛选采集的功能, 支持“文献采集 - 编辑评审 - 集成组织与发布”的 Overlay 期刊编辑出版完整流程的实现。鉴于 OJS^[13] 在开源电子出版系统领域的影响及其在技术架构上的良好可扩展性, 优先选择 OJS 作为 Overlay 期刊的基础编辑出版平台。

OJS 从发展之初即定位于为促进同行评审和开放获取期刊的出版提供可靠的技术平台, 目前已有超过 2 000 家用户。它支持包括投稿、多轮次的同行评审、组版发布、索引和检索服务等完整的电子期刊编辑管理工作流程。技术上, 同一些基于社群发展的优秀开源软件一样, 具有基于插件的功能扩展架构, 存在较多的功能插件选择, 也易于根据需要扩展开发必要的功能插件。

(2) 开发基于主题的开放知识库内容搜索与采集功能插件, 实现与 OJS 编辑工作流的有效衔接和融合, 使之满足 Overlay 期刊建设的内容筛选与采集处理需求。

由于开放知识库基本上都支持 OA I 接口, 可以通过开发基于 OA I 的元数据收割服务插件来实现对开放知识库内容的搜索与采集。Overlay 期刊的建设需要按照事先确定的学科领域或主题, 从多个知识库中定期地搜集相关学术文献, 因此, 基于 OA I 的文献搜索采

集插件,其主要的功能和技术要求包括:

支持对来源知识库的灵活配置和管理,满足 Overlay 期刊根据自身的需求,有选择地确定搜索和采集文献的目标知识库,并根据需要动态地进行调整和管理。

支持基于主题的 OA 元数据收割服务。Overlay 期刊需要从多个目标知识库中定期地采集符合其学科领域或主题特征的最新文献,这就要求在标准的 OA 元数据收割服务过程中必须加入基于主题过滤处理。在目前主题筛选和过滤技术尚不成熟的条件下,仍可以通过利用知识库支持具有一定学科或主题特征的 OA 分类集合(OA I Set)^[14]特性、元数据本身可能包含的主题标引信息、以及预定义关键词列表等多种方式的组合来达到基本的过滤和筛选目标,为后期进一步进行人工评价、采选和评审提供具有一定相关性质量保证和适度规模的数据。

筛选和采集获得的数据须与 OJS 无缝集成。OJS 默认提供了基于 XML 的数据批量导入插件^[15],但其导入过程将绕过正常的 OJS 编辑处理流程,使导入的数据直接进入发布状态。因此,采集获得的数据需要模拟作者提交的流程,首先转入 OJS 的稿件队列^[16],与 OJS 的编辑处理流程平滑衔接,融入正常的编辑出版 workflow,等待转入后续的评审、编辑、组版和发布流程。

(3)充分依托 OJS 强大的编辑出版 workflow 支持,通过适度的定制和优化,完善基于知识库采集数据的评审、编辑和集成组织,支持 Overlay 期刊的发布和出版。

(4)在面向用户的服务组织方面,除了利用 OJS 提供的浏览、检索、推送等基本服务外,可以充分利用 OJS 提供的面向不同学科领域的辅助阅读工具(Reading Tools)^[17],根据 Overlay 期刊的学科或主题特点,进行定制和扩展,为用户阅读和使用文献营造良好的相关情境,支持用户方便地基于当前浏览的文献与作者联系、推荐和共享给同事或其他用户、评论、加入相关的讨论组、提取相关信息到搜索引擎进一步检索、了解文献作者的其他作品信息等,丰富用户阅读和使用文献的操作和体验。

4 结 语

尽管基于开源电子期刊出版系统构建 Overlay 期刊建设平台,已经成为 Overlay 期刊建设和服务的主流形式,但是在如何采集和获取知识库内容,或者说与知识库之间开展规范、有效、可行的互操作方面,还存在很多需要深入研究和完善的问题。本文也只是在对

Overlay 期刊系统建设模式分析的基础上,提出了一种基于 OJS 并扩展 OA I 收割服务以支持 Overlay 期刊平台建设的基本框架,希望能够在后续的实践应用中得到进一步的验证和完善。

参考文献:

- [1] Smith J W T. The Deconstructed Journal - a New Model for Academic Publishing[J]. *Learned Publishing*, 1999, 12(2): 79 - 91.
- [2] Ginsparg P. Winners and Losers in the Global Research Village [EB/OL]. *UNESCO Conference HQ, Paris* (1996). [2009 - 02 - 20]. <http://xxx.lanl.gov/blurb/pg96unesco.html>
- [3] Peter S. Guide to the Open Access Movement [EB/OL]. (2004 - 07 - 14). [2009 - 02 - 20]. <http://www.earlham.edu/~peters/fos/guide.htm>.
- [4] Moyle M, Polydoratos P. Investigating Overlay Journals: Introducing the R DJA Project[J/OL]. *D - Lib Magazine*, 2007, 13(9/10). [2009 - 02 - 20]. <http://www.dlib.org/dlib/sepember07/09inbrief.html>
- [5] Web Science Overlay Journal[EB/OL]. [2009 - 02 - 10] <http://journal.webscience.org/>.
- [6] Wamer S, Bekaert J, Lagoze C, et al Pathways: Augmenting Interoperability Across Scholarly Repositories[J]. *International Journal of Digital Library*, 2007, 7: 35 - 52.
- [7] Live Clipboard Technical Introduction [EB/OL]. [2009 - 02 - 20]. <http://rayozzie.spaces.live.com/editorial/rayozzie/demo/liveclip/liveclip/sample/techPreview.html>
- [8] Overlay Journal Infrastructure for Meteorological Sciences (OJMS) [EB/OL]. [2009 - 02 - 20]. <http://proj.badc.rl.ac.uk/ojms/wiki>
- [9] Repository Interface for Overlaid Journal Archives [EB/OL]. [2009 - 02 - 20]. <http://www.ucl.ac.uk/ls/rioja/>
- [10] R DJA Journal - Repository APIs [EB/OL]. [2009 - 02 - 20]. <http://cosmologist.info/xml/APIs.html>
- [11] R DJA Project Software Products [EB/OL]. [2009 - 02 - 20]. <http://arxivjournal.org/rioja/>.
- [12] Overlay Journal System [EB/OL]. [2009 - 02 - 20]. <http://arxivjournal.org/index.php/index>
- [13] Open Journal Systems (OJS) [EB/OL]. [2009 - 03 - 26]. <http://pkp.sfu.ca/?q=ojs>
- [14] The Open Archives Initiative Protocol for Metadata Harvesting [EB/OL]. [2009 - 02 - 20]. <http://www.openarchives.org/OA/OA-OpenArchiveProtocol.html>
- [15] Importing and Exporting Data with OJS [EB/OL]. [2009 - 03 - 26]. <http://pkp.sfu.ca/files/docs/importexport/importexport.pdf>
- [16] OJS/OCS Quick Reference [EB/OL]. [2009 - 03 - 26]. <http://pkp.sfu.ca/files/docs/quickreference/quickreference.Pdf>
- [17] OJS in an Hour [EB/OL]. [2009 - 03 - 26]. <http://pkp.sfu.ca/files/OJSinanHour.Pdf>

(作者 E-mail: zm@lzb.ac.cn)