

**2013 年第 11 期（总第 32 期）**

# 长期保存跟踪扫描

**主办单位：中国科学院国家科学图书馆**

**2013 年 11 月**

**为传播科学知识，促进业界交流，  
特编译《长期保存跟踪扫描》，仅供个  
人学习、研究使用。**

## 目 录

<b>【动态追踪】</b> .....	1
Portico 电子期刊保存服务的里程碑.....	1
《熟知所需：优化图书馆及档案馆馆藏的保存》 .....	1
<b>【专题报道】</b> .....	2
2013 年保存存储规划会议综述.....	2
<b>【重要文献摘译】</b> .....	4
《美国国会图书馆国家录音保存计划》摘译.....	4
DPC 技术观察报告：《电子期刊的保存、信任及持续性获取》摘译.....	9
<b>【资料推荐】</b> .....	26
《数字保存的法律信息》 .....	26
对图书馆、博物馆及历史遗迹保存机构保存活动的调查.....	27

**【动态追踪】****Portico 电子期刊保存服务的里程碑**

2013 年 11 月 4 日, Portico 机构宣布其保存了 2500 万篇期刊文章, 这些文章覆盖超过 287000 卷、100 多万种期刊。对于 Portico 以及支持数字资源保存的图书馆和出版商来说, 这是一个重要的里程碑。

从 2005 年开始, Portico 保存的数字资源的数量和类型呈现出大幅增长的趋势, 目前有超过 1000 个图书馆和出版商选择与 Portico 合作。约有 17000 种电子期刊、超过 220000 种电子书及 72 种数字化的历史藏品已委托 Portico 来进行长期保存。Portico 所保存的文档、书籍及数字藏品接近 4.15 亿个文件, 存储量近 60TB。

编译自:

<http://www.portico.org/digital-preservation/news-events/news/general-news/portico-e-journal-service-reaches-historic-milestone>

(陈瑶编译, 唐果媛 吴振新校对)

**《熟知所需: 优化图书馆及档案馆馆藏的保存》**

《熟知所需: 优化对图书馆及档案馆馆藏的保存》这篇报告是英国国家图书馆在 2006 年报告的成果基础上形成的, 它所选用的数据来自 86 份保存评估调查, 这些调查总结了在 2006 年到 2011 年间保存工作中的优势和不足及为确保能长期使用所保存资源所采取的措施。

这篇报告建议从三个方面采取行动, 以提高所保存资源的长期可用性及优化保存策略应用:

- 1、组织机构应对他们所保存的资源有更为清晰的理解。
- 2、指出在保存时间中存储环境领域存在的缺陷, 以及通过应急预案弥补这些缺陷。应急预案应该被定期检查并进行员工实战培训。
- 3、保存替代品应作为一种保存策略而受到重新检查。

报告全文: <http://www.bl.uk/blpac/pdf/ktn.pdf>

编译自:

<http://www.bl.uk/blpac/ktn.html>

(陈瑶编译, 唐果媛 吴振新校对)

**【专题报道】****2013 年保存存储规划会议综述**

美国国会图书馆每年都会举办数字馆藏的存储架构设计会议即保存存储会议。2013 年该会议于 9 月 23 日-24 日举行，会议中一系列的演讲和讨论令人印象深刻。

标准是今年会议的主题。该主题不仅适用于媒体和硬件，还适用于接口。在保存领域，接口是一种用户与工具之间进行存储文件交互的软件和操作系统机制，而这种接口是消失得最快的，至少它的变化要跟上不断变化的需求。引用 Instrumental 公司 Henry Newman 的一句话来评价这次会议，即“这些都不是新问题，只有解决旧问题的新工程师。”

美国国会图书馆人员讨论了图书馆的基础设施和需求，以此拉开了此次会议的序幕。图书馆通过目录清单服务来记录文件，其中包括要用于未来审计的不变性。但是，不同的内容有不同的需求。存取文本和图片副本的数据中心要在 10 秒内存储管理 PB 级的数百万文件。存取视频和音频副本的数据中心要在 10 秒内存储管理 PB 级的约 70 万文件。不同规模的文件数量和大小意味着不同阶段需要不同的硬件来传递这些内容。在图书馆存储采购方面，30%的采购纯粹是为了扩大存储容量，70%的采购是因为持续更新的技术，通常也包括扩大存储容量。

在本次会议中，磁带技术始终是一个热门话题。T10K 磁带迁移一直在持续。希捷公司的 Dave Anderson 介绍，在 5-10 年内将不再支持磁带环境的接口。IBM 公司的 Gary Decad 认为，磁录密度的增长率在减慢，而且 PB 级存储制造的年增长率也不再增加。

磁带是迄今为止在生产使用中有最高 MSI（百万平方英寸）的存储介质。磁带、硬盘驱动器和固态存储是表面区域密集型技术。许多参会人员相信，固态存储提高了硬盘驱动器技术。NAND 闪存对硬盘驱动存储使用的影响不太明显。替换企业硬盘驱动的费用相当大，而且在短期内不会替换。

为了市场稳定，董事会必须把技术授权给多个制造商和供应商。但是为 NAND 闪存等新技术建造生产设备是非常昂贵的。LTO 磁带设备也昂贵，但不是因为建造设备昂贵，而是因为制造业中缺乏盈利能力。参会人员在参加完此次会议后，对制造企业存储标准的许可和现存的垄断有了更加深入的认识。

“云”小组展开了热烈的讨论。期间，三句名言脱颖而出：第一句是来自电影艺术与科学学院 Andy Maltz 的“云是不错，但有时会下雨”；第二句是，来自国会图书馆的 Fenella France 说“我与某人交流，他说我们的对话发生在云端，我问云在哪。云仍然在某处的物理服务器

上。”；第三句是来自 Tessella 的 Mike Thuman“云和本地之间的双向箭头不是基于 Kryder 定律或摩尔定律，而是基于墨菲定律，其需要把数据传回来。”。

斯坦福大学的 David Rosenthal 提出了一些关键主题：

- 1、 什么时候云比你自己做得更好？当你有强烈需求而不是稳定的使用的时候
- 2、 云的使用是采用“毒贩”法则。首次免费，但是由于下载/退出/迁移会产生费用，使得用户很难离开云。
- 3、 云不是一项技术，而是一种商业模式。技术是可以让你使用自己的东西。

亚马逊的 Jeff Barr 说，“我想我是官方主导市场的正式毒贩。”但是亚马逊也迫切想知道，在 AWS 环境下，需要给保存社区提供什么样的存档文件的保存行为报告系统。

在标准这部分，不仅用附有剪贴画的精美幻灯片介绍了 NISO 和标准的制定过程，还包括标识符和文件系统，以及一个新兴标准 AXF 的说明。

今年会议出现的一个相对较新的主题是使用开源的解决方案，比如在 OpenStack 项目中组件工具的使用范围。基于 HTTP 的 REST 是日渐重要的文件接口，其技术是从基于系统的文件接口转向基于对象的文件接口。目前一切都可以从厂商获取自定义存储管理层。

会议还讨论了其他形式的媒介。有两个最具创新的媒介，一个是在密封的盒上刻有激光的不锈钢卷带，另一个是记录视频的金属合金媒介。光学媒介还没有消亡。日立集团的 Ken Wood 指出，已存活 30 年的商业音频 CD 仍在硬件市场立足，并将继续存在。从技术上来说，纠错的软件接口与硬件支持同样重要。但是随着移动设备和无光学播放器或硬盘的轻薄笔记本电脑的兴起，机械光盘播放机和存储正在消失。

来自数字保管和保存社区的代表们的 ppt 占了会议内容的很大比例。较有特色的包括：数据保护的项目以及洛斯阿拉莫斯国家实验室、国家人文基金会和美国国会图书馆所付出的努力。他们不止一次的指出，内容和数据的创造者常常不觉得保存是他们的一部分责任。有一句重要名言是，“比起实际保存所花的时间，你可以花费更多的时间来弄清楚保存的内容。用于评估保留哪些内容所需的保管费用可能是巨大的。”

编译自：

<http://blogs.loc.gov/digitalpreservation/2013/10/planning-for-preservation-storage/>

（唐果媛编译，王敬 吴振新校对）

**【重要文献摘译】****《美国国会图书馆国家录音保存计划》摘译****文献摘要:**

美国全国的图书馆、档案馆和博物馆约有四千六百万件录音制品,其中有数百万的音频记录需要长期保存。唱片公司、表演艺术家、广播公司和收藏家持有的数百万独特的音频记录也需要长期保存。在数字化时代,新技术为保存措施提供了保障。然而,对于有责任为后代保存历史音频记录的机构来说,将数字资源保存对象扩展到数字音频已带来了重大的技术、组织和资金挑战。

美国国家录音计划已经制定并提供了一份蓝图以“全面实施国家录音保存规划”,制定了“2000 国家录音保存行动”。国会指出,美国国会图书馆依据本法律建立的该规划应以教育为目的增加音频记录的可获得性。只有公众可以听到保存的音频,公众才会从中受益。拓宽访问的技术、体制和法律障碍是国家录音保存工作面临的严峻挑战。该计划识别了音频领域最重要的保存和访问问题,并提出了相应建议。

国会认识到,国家录音保存计划应通过共同努力来实现,除了美国国会图书馆,还应包括音频存档工作者、教育者和历史学家、版权所有者、唱片界的代表,以及其他与录音保存相关的工作者,并且应该考虑到录音保存委员会所做的研究。该计划在很大程度上源自美国国家录音保存委员会委托进行的一项具有里程碑意义的研究,并作为国家录音保存行动的成果发布,该计划强调公私利益相关者之间的协调合作是成功实现国家录音保存项目必不可少的条件。

国家录音保存计划对协调实施保存工作的建议有以下相关联的四点:保存基础设施,保存策略,访问的挑战,以及关于保存和访问的国家长期战略。有些建议将在不久的将来实现。长期计划的实现可能需要一代人的努力。

**正文摘译:****1、建设国家录音保存的基础设施**

新的国家保存基础设施应以协调的方式开发,实现三个主要目标:

- (1) 扩展物理和数字的基础设施以妥善存储和保存有风险的音频材料。
- (2) 发展教育计划,包括开设学位培训课程,培养具备能对音频记录进行数字化并保存的全能专家,同时建立一个统一音频保存资源目录的专业网站,以支持教育并让专业人员了解最新的进展。

(3) 建立一个协调的国家研究日程来支持新技术的开发，以解决该领域最具挑战性的问题。

## 物理和数字基础设施

### 建议 1.1

音频记录的媒体存储设备：构造环保的受控存储设备，为音频记录媒体的长期保存提供最佳条件。

### 建议 1.2

提高国家保存音频的能力：制定策略以增加当地、地区内和国家的音频记录保存资源重新布局，以此来满足预期的需求。

### 建议 1.3

数字存储：制定策略确保各种类型的图书馆和档案馆都能访问数字仓储库中长期保存的音频内容。

## 教育和专业培训

### 建议 1.4

大学课程和学位课程：建立基于大学的音频存档和保存学位课程。

### 建议 1.5

音频保存的继续教育：为音频实践工程师、档案管理员、馆长、图书管理员等开展继续教育项目。

### 建议 1.6

音频保存资源目录：创建一个协作的在线资源集，来收集、审查、传播音频保存领域的知识和最佳实践。

### 建议 1.7

音频保存的新技术：鼓励科学技术研究，开发新技术来恢复、重新格式化并保存音频记录媒体。

### 建议 1.8

传统技术的文档：研究、收集、记录并保存用传统做法和技术记录的信息。

## 2、实施保护策略的蓝图



### 建议 2.1

音频保存指南：编制一个基本的音频保存手册，用它来指导非专业人士管理音频收藏品。

### 建议 2.2

评估保存的音频藏品：制定措施协助音频藏品管理者全面评估藏品，以确定优先保存的对象。

### 建议 2.3

公私合作：通过音频保存资源目录网站宣传指南，在公共机构、私营企业、私人收藏家和其他相关者之间建立合作伙伴关系来保护濒危音频。

## 保存数字音频文件的新工具和指导方针

### 建议 2.4

音频材料的保存工作流程：开发工具，实际执行模型，最佳实践，以及数字化模拟音频记录的高效率工作流程。

### 建议 2.5

数字音频文件的元数据标准：结合标准建立和最大化互操作性，开发与数字音频文件相关的元数据指南和最佳实践的建议方案。

### 建议 2.6

支持整个内容生命周期保存的工具：鼓励开发符合创作录音和保存管理的标准和最佳实践的工具。

### 建议 2.7

创建和保存原生数字音频文件的最佳实践：研究、开发以及提升改进的和可扩展的流程以封装拥有多个部件和丰富元数据的数字音频对象，并开展（或更新）长期存档对象转换和管理实践。为了使生产音频记录的机构创建的文档即作为原始存档文件（不再需要转换），需要定义首选格式（其中包括元数据），或者为首选文件格式、嵌入式或相关元数据和对象包制定建议方案。

## 3、出于教育的目的应促进广泛的公共访问

### 建议 3.1

国家录音作品：通过扩大现有的录音作品项目来鼓励持续开发权威的国家录音作品。

### 建议 3.2

本作品采用[知识共享署名-非商业性使用-禁止演绎 2.5 中国大陆许可协议](#)进行许可。

国家录音集合名录：创建一个可公开访问的机构、企业和私人录音藏品的国家名录。

### 建议 3.3

为音频编目建立最佳实践：建立一个由美国国会图书馆领导的专家工作组，以确定编目记录所需的数据，以及为音频编目开发最佳实践，音频编目要以精简的标准为目标。

## 版权立法改革

### 建议 3.4

联邦版权保护 1972 年前的录音制品：把 1972 年 2 月 15 日前的录音制品纳入联邦版权法保护范围内。

### 建议 3.5

孤儿作品：确保不能识别版权主体的录音制品能更容易保存和合法使用。

### 建议 3.6

美国版权法第 108 条：为了便于保存和扩大公众访问录音制品，需修改美国版权法（1976）的 108 条。

## 提高公众访问录音制品的合法性

### 建议 3.7

音频流许可协议：构建一个许可协议的基本模型，允许图书馆和档案馆按照研究员和大众的需求以安全的方式复制音频记录。

### 建议 3.8

把保存的音频记录接入网络：创建一个共享保存网络，通过它可以访问多个图书馆和档案馆的录音制品。

### 建议 3.9:

音频记录标注所有权数据库：为音频记录标注的所有者创建一个在线公共登记中心。

### 建议 3.10:

扩大公众访问美国国会图书馆保存的音频记录：探索如何使全美国的研究员都能访问到美国国会图书馆数字化的音频记录。

### 建议 3.11:

合理使用最佳实践的准则：制定一个图书馆和档案馆能够接收和使用的合理使用音频记录的最佳实践准则。

#### 4、国家长期战略

##### 建议 4.1

国家音频保存委员会：国家音频保存委员帮助美国国会图书馆协调和执行国家音频保存工作，以及提高国家规划的公众意识。

##### 建议 4.2:

音频记录保存的行政领导委员会：在国家音频保存委员会和美国录音艺术与科学学会的共同支持下，组织一个由业内高管和档案馆负责人组成的咨询委员会，以解决需要公私部门合作才能解决的音频记录保存和获取的问题。

##### 建议 4.3:

协调的国家馆藏政策：为音频记录制定一个协调的国家馆藏政策，包括建立一个“合作伙伴档案”来接收版权费；共同努力收集原生数字音频；同时制定策略来收集、编目和保存当地产生的音频、电台广播内容、那些易被忽视的新出现的音频格式和类型以及公司文件。

##### 建议 4.4:

二十一世纪的音频保存：制定战略，使图书馆和档案馆能收集和保存目前受到用户许可协议限制的具有深刻文化意义的音频。

##### 建议 4.5:

为保存音频筹款：通过协调国家音频保存委员会、国家音频保存基金会和赠款项目的筹款策略，来确保国家保存工作的发展和成功。

##### 建议 4.6:

评估国家音频保存项目：由美国国会图书馆国家音频保存委员会主持召开美国音频遗产保存的会议。

编译自：

<http://www.loc.gov/rr/record/nrpb/PLAN%20pdf.pdf?loclr=blogsig>

(唐果媛编译, 王敬 吴振新校对)

## DPC 技术观察报告：《电子期刊的保存，信任和持续性获取》摘译

### 报告摘要：

这份报告讨论了图书馆、出版商、中介机构和服务提供商在数字保存以及有关电子期刊的信任度和长期获取等领域的发展状况和问题，还包括从广泛的数字保存社区学到的有关外包业务和信任度的一般性教训和建议。报告不仅仅关注技术问题，同时还涉及到相关法律、经济和服务的问题。

### 正文摘译：

#### 1、背景

2012 年 1 月，DPC 举办了关于“信任度和电子期刊”的论坛。该技术观察报是受 DPC 的成员及其它参会成员的委托并在他们的讨论结果上形成的。

这份报告讨论的主题是：电子期刊的持续性访问、信任度及保存。目前期刊和文章出版已经从印刷格式转向电子格式，这种转换使传统出版商业模式和关系发生了重大的转变，对研究性图书馆来说，这些问题变得越来越重要。

目前，这个领域已经出现了许多重大的变化。例如，由图书馆采购期刊并将其存放（和保存）在本地磁盘上（图书馆之间产生了多个冗余副本）的模式向租用（授权）远程访问保存在出版商的平台上的电子期刊的模式转变，这些平台通常位于国际上其他的管辖区内。

除了这些变化，还有不断发展的电子期刊开放获取运动，这种开放获取试图取消获取时费用。订购期刊、开放获取期刊及两者的混合（即包括期刊的文章有一部分是开放获取的，有一部分是需要订购的，也包括期刊在固定的一段时间后才能免费获取）的出现构成了复杂的出版状况，基于此，必须考虑电子期刊的保存和长期访问。

电子期刊的这种状况是随着电子出版本身开始由静态向动态内容的转变而演变的，在主要学科中，数据及与文章相关联的补充材料的重要性在不断增加。

所有这些改变反过来使电子期刊的保存更具挑战性、更加国际化并依赖于其它事物，同时也把有关信任的问题推向了前端。在这样的环境下，信任不仅仅是保存的技术问题，还有谈判的权利（保存谈判的记录以供未来使用）问题，以及获得关于机构所存资源及如何被保存、如何及何时才可以获取这些资源的所有透明信息。

这使得电子期刊成为数字保存最具活力和挑战的领域之一，特别是在关于分享或是外来资源保存服务的商业模式和信任机制中。

因此，这份报告很可能会得到数字保存机构的广泛关注，而不仅限于从事电子期刊生产、获取及保存的机构。

## 2、问题

### 2.1 印刷到数字的转变

电子期刊越来越流行，它作为学术界研究论文的一种传递机制，正大规模地取代纸质期刊。尽管出版和获取电子版的学术期刊能带来许多显而易见的益处，但是为了能保持长期访问它们，必须处理好相关的费用问题及挑战。

电子期刊的转换仍在继续，它们不再是早年流行的“印刷期刊电子化”模式。电子期刊的底层数据和其他相关文章之间保持了更多复杂的联系，提供了更多的功能用于获取和解释内容本身。不断增加的电子版本可以看作是“记录的版本”，比任何纸质副本更丰富和完善。

电子期刊或许比其它任何形式的内容发展得更为迅速，这给数字保存提出了新的挑战，这就需要持续的研究与探索以解决这些问题。就期刊的拥有者和用户访问出版平台的所属地域而言，电子期刊还延伸到了国外，这给呈缴立法的起草及既含有电子版又含有印刷版的期刊的转换阶段带来了挑战。这种转变对图书馆工作流的重建也有影响。在印刷界，期刊的选择、收集维护和保存由不同的单位（每个单位都有特定但互补的角色）负责监督工作。在电子界，需要更为完整的决策框架，如，选择期刊时，期刊的影响因子是决定保存期刊的重要考虑因素。

### 2.2 出版业的结构

向电子出版的转变加速了出版业的整合和全球化，这种趋势在 STM（科学、技术和多媒体）出版商中尤为明显。学术期刊目前涵盖了大量的出版物和出版商。2009 年，Swets 估计全世界大约有 52730 种具有商业授权的学术期刊，其中 24369 种学术期刊可以在大型的 STM 出版商中获得，大约有 28361 种学术期刊可以从小型 SMT 出版商中获得，这些小型 SMT 出版商拥有一种或多种期刊。

对电子期刊进行数字保存的挑战之一是属于少数的大出版商和 700 多个属于长尾的小出版商（多数为只有一个期刊的社会出版商）之间不同的标准流程及格式、技术能力和创新能力之间的巨大差异，这种两极化趋势不断发展，相对而言，因为大型出版商（或出版商的大型主机平台）的存档程序更容易建立，并且规模经济也发挥了作用，其存档时的服务费用要比小型出版商便宜。

大型出版商已经能够提供“大宗交易”，这种“交易”把订阅的电子期刊与它们的其它众多产品捆绑在一起。尽管这种交易可能扩大获取的范围，但捆绑后的过期停定获取权没有被平等对待，通常有一个核心的订阅期刊，它们有过期停定获取权，而余下的期刊拥有当前订阅获取权但是没有过期停定获取权。

有许多集成商,他们从众多不同的出版商获取当前的期刊,并以极具吸引力的价格提供给用户。然而,必须声明,集成商不能提供过期停定获取权,他们仅仅适用于临时获取需求。

最后,开放获取,通过互联网来评价同行的学术文章以提供不受限制的获取权力,对出版业具有重大影响。反过来也影响了电子期刊及期刊中学术论文的数字保存。这里出现的挑战包括:明确对开放获取文章适当的存档权限及保存的资助,或者如何选择要保存的开放获取期刊。

## 2.3 保存

或许有三个问题需要注意:保存方法,保存范围,商业模式。商业模式会在经济因素之后(Section 2.7)再讨论。

### 2.3.1 保存方法

保存电子期刊有两种广泛使用的技术:

- (1)、出版商提供源文件给存档机构,源文件来源于已产生的期刊呈现文档;
- (2)、存档机构从出版商的网站上获取呈现文档,包括便于浏览器呈现内容的所有辅助文件。

这两种技术各有利弊。

源文件保存的优势是:与期刊呈现文档相比,它包含了更多的内容(但是考虑到一些情况,如出版商和存档机构达成的相同的传递格式与元数据可能会因实际的源文件不同而不同);存档机构将保留这些源文件,并将它们变成标准的存档格式,随着时间的推移能对它们进行管理、迁移或者模仿;它不需要为了准确地描述网络内容之间差异性而不断进行更新的证据和网络爬虫。劣势是:如果有些材料错误性的缺失,它就有可能不完整;前期需要大量投资;呈现文档也会和出版商的呈现文档大不相同。

获取呈现文档的优势在于可以得到关于出版物的视觉和感觉,初期的费用也很低;这种技术的劣势是:存档机构必须适应出版商的网站,保证可以获取到相关的内容;通常,出版商更新他们网站的频率要比改变源文件交付流程的频率要高;存档机构必须仔细地搜寻出版商的网站,在网站使用高峰时,要限制提取率,同时避免爬行陷阱,等等。

这些技术都存在潜在的局限性:随着内容更加动态化和网络化,每种技术都试图通过不同方式适应这种变化;它们都需要在未来进行格式的迁移和模拟,这是潜在的挑战。

### 2.3.2 覆盖范围

所有期望使用第三方服务或建立自己存档的图书馆的一个重要考虑是它所使用题名的覆盖范围及如何选择及存档这些新题名的优先程度。可以依据规模、价值或者损失风险度来本作品采用[知识共享署名-非商业性使用-禁止演绎 2.5 中国大陆许可协议](#)进行许可。

优先考虑出版商和期刊，通常这些是截然相反的，最小或者占据较小市场份额的期刊通常也是风险最大的（至少在需求方面）。主要存档馆的题名覆盖范围大不相同。最近一项对康奈尔大学和哥伦比亚大学图书馆的研究发现，LOCKSS 和 Portico 仅保存它们电子期刊的 15%。同时他们注意到，尽管 LOCKSS 与 Portico 之间期刊的保存范围有重叠，但这两个机构保存的期刊并不相同。

近来一项关于 LOCKSS、CLOCKSS、Portico 保存系统保存内容的比较研究发现，合作出版商与期刊间存在显著重叠。LOCKSS 存储了许多小型的、处于危机中的出版商的期刊，这些出版商认为 LOCKSS 对他们来说也许是唯一经济可行的选择。

### 2.3.3 呈缴

国家图书馆的主要功能是保证已出版的国家遗产已被保存并且可被访问。呈缴制度是许多国家存档的重要工具。把呈缴制度从印刷环境延伸到电子期刊和其它电子出版物是全球趋势。

存在许多实际性的挑战。在印刷环境中，概念、规则和实践已被接受和理解，在网络环境中，出版物、出版商和出版地可能会表达新的意思或是原意思不适用于网络环境。国家存缴图书馆携手其它图书馆、出版商和技术商积极探讨了这些挑战及相关技术问题。在长期的磋商及数次延期后，2013 年 4 月，英国制订出了电子出版物的呈缴准则。

呈缴立法（或类似的自愿上缴安排）一般涉及订购期刊，这些期刊被认为是已出版的国家文化遗产的一部分。另外，在很长一段时期内，它会在很大程度上限制异地获取保存的电子资源，目的是保护出版商的商业利益。

通常这种国家依法呈缴馆藏不能适用于国际范围的要求，这些订阅期刊的永久获取权即订购授权属于其它图书馆及其用户。荷兰国家图书馆已经制订政策来使用国家授权保存电子期刊，另外还包括一系列在外国出版的电子期刊（包括 DOAJ 的开放获取资源）。

## 2.4 授权

在简介中已提到，对许多商业出版商而言，图书馆付费订阅电子期刊的授权访问是主要的商业模式。代表图书馆的联盟在授权模式和条款的谈判中通常扮演了重要的角色。对今后的获取来看，保留这些条款及权利的记录对以后取消获取权是至关重要的，特别是涉及到许多不同的权利的“大宗交易”。Jisc Collections 最近的一项研究强调了随着时间的推移出版商和图书馆维护这些权利记录所面临的一些挑战。

存档条款、持续使用或过期停定获取权的条款及保存这些条款的详细具体的存档机制都很重要。一个例子是：UK NESLi2 2013 Model Licence 相关条款。

出版商可以通过自己的主机服务和平台提供过期停用获取权,但是订阅用户仍然担心他们长期访问的能力。为了解决电子期刊长期保存和持续获取中出现的具体问题,特别是订阅商业性的电子期刊,已经出现一大批新的存档机构。它们寻求实现由出版商、图书馆及他们的专业组织或者协会制订的授权协议的相关条款。

## 2.5 “触发事件”

对出版商的商业模式而言,重要的是,特定的存档机构不直接与出版商自己的平台竞争。同样重要的是,尽管随着时间的推移,有些事件可能会影响到出版商,但是图书馆也必须确保获取的持续性。已有机构在很努力地探索“触发事件”或所有权转让以及相关流程,并在出版商与存档机构达成的协议中声明。

在与电子期刊及其持续传递相关的特殊条件被满足时才会激活“触发事件”。如果因种种原因,一些出版商或其他资源的电子杂志无法被用户获取,这时触发事件就已经发生了。这时可以通过已实施数字化保存的存储机构为用户提供动态访问。

## 2.6 期刊转移

在电子出版界,出版商或者所有者的改变以及由此而产生的角色和职责转化,必须被“转移出版商”和“接受出版商”接受,用以保证期刊订阅用户和使用者能永久访问内容。已经创建的转移行为准则(现在已经是第二版,2013年正在磋商形成更新和扩展后的第三版草案,第三版试图制定可接受的操作标准),鼓励出版商签署此准则以实现“转移兼容”。

## 2.7 信任

从上面可以看出,在电子期刊的数字保存过程中,不同的部门和组织间存在着复杂的相互依赖关系,同样也获得了许多获取电子论文的好处。只有相关团体之间建立一定程度的信任,这种依赖关系才能形成。图书馆工作人员数字保存的知识和意识水平同样重要。因此,电子期刊数字保存中的信任是多层次的。它涉及到不同利益相关者间的信任,以及它们之间协议及实现过程强烈的信任。在第六节将进一步讨论保存机构中信任的定义并提供信任的指标。

## 2.8 经济考虑

### 2.8.1 研究型图书馆的经济考虑

图书馆正面临着资金限制和不断增强的空间压力,它们对快速转变为提供纯电子资源表现出越来越大的兴趣,从而缓解这些压力并提供给用户全新的电子服务。在向纯电子资源的转变过程中,图书馆及其职工面临的最常见的障碍和问题之一是维持和确保电子内容的长期访问。



出版机构投资电子期刊存档的主要动机来自向电子期刊转变过程中所产生的成本和服务的利益。这大致分为两个平行互补的区域:

(1)、当前许可杂志的印刷版或者印刷版+电子版向纯电子期刊过渡过程中产生的利益;

(2)、购买或者授权以往电子期刊和(或)历史悠久的印刷版期刊数字化后带来的利益。

对图书馆来说,与纸制版期刊相比,电子化期刊显著、潜在地减少了成本。因此,减少在接收处理印刷版期刊的过程中的图书馆人力资源和资金的预算,例如,改变近期杂志的陈列、书脊标签、条形码、插入和应用藏书标签、装订和排架;减少整架过程中图书馆人力资源和资金的预算,例如,巡视和整理书架,藏品顺架,藏品清理和清洁;减少在使用本地工具(如联机公共检索目录)帮助用户直接获取期刊过程中的图书馆人力资源和资金的预算;由于不再需要存储印刷期刊而减少图书馆空间租金成本;释放低使用率的印刷期刊所占的图书馆核心空间,用于满足图书馆的其它需求(并避免新图书馆扩展中周期性运行成本和资金的消耗);纯电子期刊比“印刷+电子”期刊更有利于合作采购。

这些节约非常可观,但是他们必须意识到,这也需要非预期的成本:通过对电子期刊存档方案进行投资,来为授权期刊的持续/永久的获取提供额外的担保和机制,通常,方案的年度成本相对能接受;在授权中,持续性获取权利可能有附带成本:例如,机构必须保证订阅者能够获取之前订阅的期刊;或者确保取消持续性获取权中所产生的获取成本或最小成本(往往没有迹象表明这将是什么);可能产生额外的计算机硬件和软件、用户在线管理和维护的费用;欧盟在向电子期刊转变过程中,图书馆承受了较高的增值税(印刷期刊则会承受较低的增值税);与印刷期刊相比,授权的电子期刊在实施和监控更复杂的本地访问条款和条件过程中产生了更多的花费;用户在图书馆和(或)家里所花费的打印费用;其它研究估算了向纯电子期刊的转变和合理化的印刷集的潜在的成本节约。

他们认为,电子期刊存档解决方案的成本通常可以并入印刷期刊成本节约中的主要现金预算(例如绑定)。释放/减少印刷期刊的存储可以显著地降低和控制成本,以支撑向纯电子期刊转变并投资于电子期刊存档方案。

## 2.8.2 保存型图书馆/服务商的经济考虑因素

出版商用不同的方式展示他们的电子产品。在英国图书馆电子期刊研究中,作者研究了出版商产出和集成商(期刊引文数据库)必须要处理的元数据。他们发现,在这中间存在了不同的处理过程,这“依赖于出版商的技术成熟度和购买的服务水平”。

总之, 电子期刊的任何集成过程都可能是复杂且耗费大量劳动的。这项工作不是一蹴而就的, 在期刊出版和处理、出版商合并或出售期刊过程中一直需要这项工作。

如 2.3.2 节所述, 技术的复杂性和成本导致出现严重的困境, 因为在有限的预算内很难获得规模经济或大量的资源。

论文版权的任何复杂性也意味着相关保存活动很难规模化。即使对开放获取的文章和期刊的保存来说, 也很难规模化, 除非有永久的授权, 允许在没有协商 (例如, 使用知识共享许可) 的情况下存档, 因为, 分别与版权持有人协商的费用是惊人的。

最后, 对电子期刊存档服务来说, 长期持续性资金支撑仍是基本的挑战。正如关注数字资源长期保存和访问的 BRTF 指出的那样, 高教领域的收藏利益意味着学术存档给每个人都带来了好处。但对于数字信息, 如电子期刊, 单独的机构几乎没有动力去保存这些信息, 他们强烈渴望其它机构能存储这些信息。

因此产生了搭便车的问题。缺乏认识和理解加剧了问题的严重性。存档服务在鼓励众多图书馆成员保持它们资源的持续性方面面临着重大的挑战。

### 3、 标准和最佳实践

本节探讨一些在电子期刊保存领域有影响力的关键标准和最佳实践做法。这次是有选择性的讨论, 并非探讨数字资源长期保存方面的所有标准和最佳实践, 其中许多内容已经在 DPC 技术观察报告系列中的其它报告中探讨了。读者可参阅“web 存档的报告”(Pennock, 2013), 讨论了与 web 内容以及出版商网站的电子期刊呈现文件相关的采集、其它标准和最佳实践。

这里主要从以下三个方面进行探讨: 对象结构和元数据、标识符和服务、正式定义的最佳实践。第 6 节探讨了信任度和保存仓储库的标准, 以及这些标准如何在电子期刊存储中实施。

#### 3.1 对象结构和元数据

随着期刊从纸质到电子格式的转变, 它们密切反映了彼此的外观和结构。为了维持电子期刊的静态视觉呈现而做出的出版和保存工作主要集中在便携文档格式 (PDF) 的使用上。由于电子期刊不断进化、数量的不断增多及内容的动态化, 电子期刊的出版和保存已越来越多的采用可扩展标记语言 (XML)。XML 是一种标记语言, 它定义了一套文件编码规则, 其格式具有人机均可读的特点。对于在科学、技术、工程和医学领域的许多用户, XML 的机器可读性胜于 PDF, 这正成为一个显著的优势, 也是电子期刊存档机构面临的一个重要问题。

### 3.1.1 PDF 和 PDF/A

Adobe 的便携文档格式 (PDF) 于 90 年代初作为发布、共享文档的一种方式首次推出, 它维持其静态视觉外观并独立于呈现这种格式文件的工具和系统。PDF 起初只是一个专有的格式, 2008 年将其作为 ISO 开放标准发布, 被称为 PDF/A 标准的 PDF 版本, 目前已定义了一个电子文件长期存档的 PDF 文件格式标准, 如果忽略其中不适于长期存档的特性, 可将其作为 PDF 的一个子集。但是, 最新的 PDF/A 第三版 (ISO, 2012a) 收到了图书馆保存领域的一些警告。致力于 PDF/A 的 DPC 技术观察报告预计将在 2014 年出版。

### 3.1.2 XML 和期刊文章标签套件

基于 XML 的格式已成为许多办公生产工具的默认格式, 其中包括微软办公软件 (Office Open XML)。

XML 规范定义 XML 文档为一个结构良好的文本, 这意味着它满足规范中提供的一系列语法规则。除了结构良好, 一份 XML 文档应该是有效的, 这意味着它包含某种参考模式及符合某种模式。

最初是文档类型定义 (DTD)。XML 模式是一种较新的模式语言, 通常称为 XSD (XML 模式定义)。XSDs 在描述 XML 时, 功能远比 DTDs 强大, 并且它允许对 XML 文档的逻辑结构进行更详细的约束。

期刊文章标记套件提供了一组 XML schema 模块, 用它来定义元素和属性, 这些元素和属性用于描述期刊文章的文本和图形内容, 以及一些非文章材料, 如信件、社论、书和产品的评论。该标准包括对元素和属性的描述, 以及三个期刊文章标签集——存档和交换标记集、期刊出版标记集、文章创作标记集 (Beck, 2011; NISO, 2012)。它由美国国家医学图书馆 (NLM) 文档类型定义套件 (DTD Suite) 演化而来。当美国国家医学图书馆 (NLM) 发现在公共医学中心需要一种通用格式的单个文档类型定义 (DTD) 来实现出版商和存档机构之间的数据交换时, NLM DTD Suite 应运而生。NLM DTD 的第一版于 2003 年年初公布, 2006 年, 英国国家图书馆和美国国会图书馆宣布了它们的期刊内容交换也遵循这一标准。它迅速成为 XML 中标记期刊文章约定俗成的标准, 甚至在 NLM 之外也是。随着使用率的提高, 用户和潜在用户开始寻求文章模式的正式化, 以作为国家信息标准组织 (NISO) 的标准。NISO 的标准工作开始于 2009 年年底, 2012 年 8 月完成并发布, 即为 NISO JATS——期刊文章标签套件, 版本 1.0。

### 3.1.3 用于保存馆藏的 ONIX (ONIX-PH)

在线信息交换 (ONIX) 是一个用电子形式表示著作和系列作品信息的国际标准。

用于保存馆藏的 ONIX (ONIX\_PH) 是一种 XML 信息结构, 用于传播已被保存下来或本身为保存措施对象的电子资源馆藏。ONIX\_PH 传递馆藏声明快照, 并显示电子资源覆盖的各个版本的保存工作状态。在结构上, 该消息是一个为系列网络馆藏 (SOH) 专门发展的 ONIX 格式, 它运用 ONIX 申明馆藏范围, 允许复杂的、多层次的覆盖和枚举的交换。

ONIX-PH 适用于出版商、维护保存资源注册信息的保存机构和电子资源供应链中的其他中介机构。Jisc 资助的 PEPRS (试行电子期刊保存注册服务) 项目的合作伙伴在协助定义 ONIX-PH 的业务要求中发挥了重要作用, 尝试在 Keepers 注册系统中交换从保存机构获得的信息 (见 4.8 节), 它由 EDINA 主办, 并与 ISSN 国际中心合作开发。目前 ONIX-PH 在 V0.21 版本草案中可用, 它正在进行试用交流。在试点结束后, 将根据需要对该信息进行修改并发布正式版本 V1.0.6。

#### 3.1.4 Ajax 和 HTML

超文本标记语言 (HTML) 用于在 web 浏览器中显示文本和图像, 长期以来一直作为基于 web 交互的标准。Ajax (异步 JavaScript 和 XML) 已经成为越来越受欢迎的设计方法, 并且可作为流行的 HTML 浏览器中的 web 应用技术, 可以提供更具互动性、类似桌面的用户体验。当前版本的 HTML 标准 (HTML4) 预计将在 2014 年正式被 HTML5 (目前正在经过正式筛选采纳过程) 取代。HTML5 旨在改善 HTML 以支持最新的多媒体和设备, 它扩展并提供合理可用的文档标记; 可以处理多媒体和图形内容而无需使用专有的插件; 可被制定用在如智能手机和平板电脑等低功率设备上运行 (W3C, 2012)。这些发展反映了电子期刊是如何随着 web 技术的发展, 由静态到更多的动态内容, 并在发展中提出新的保存挑战。Web 存档和有关技术问题的更广泛的讨论请参阅 DPC 技术观察 web 存档报告 (Pennock, 2013)。

### 3.2 标识符和标识符服务

持久性标识符的概念已经演化成尝试通过一个查找过程来提供直接访问的一种方法, 因此, 即使随着时间的推移, 他们的位置发生改变, 对于电子出版的论文或数据集的引用也保持不变。电子期刊最广泛采用的持久性标识符一直是数字对象标识符 (DOI), DOIs 也越来越多的用于研究数据集的引证。对于电子期刊保存服务, 另一个关键标识符的作用就像 ISSNs, 是作为期刊及其不同媒体版本的唯一标识符。

#### 3.2.1 数字对象标识符

数字对象标识符 (DOI) 是永久标识数字内容的管理服务。标识符 (DOI 名称) 被解析至注册人指定的位置。一种可扩展的元数据模型用于将描述性以及其它元数据元素与 DOI 名称关联起来。

根据国际 DOI 基金会 (IDF) 提供的公共基础设施和相关政策, DOI 系统通过注册机构联合实施。

### 3.2.2 国际标准刊号

国际标准刊号是由八位数组成, 用于唯一标识纸质或电子期刊出版物。以纸质或电子形式出版的期刊可能有两个 ISSN 号, 纸质版一个 ISSN 号, 电子版一个 ISSN 号 (e-ISSN 或 eISSN)。ISSN 或 ISSN-L 联系着连续性资源的不同版本, 以保证其连续性。

ISSN 代码是由通常位于国家图书馆的 ISSN 国家中心分配的, 并由位于巴黎的 ISSN 国际中心负责协调。国际中心是一个政府组织, 通过联合国教科文组织和法国政府间的协议创建于 1974 年。国际中心维护国际连续出版物数据注册系统 (ISDS), 或者称为 ISSN 注册系统——世界范围内已分配的所有 ISSN 号的数据库。ISSNs 和 ISSN 注册表已被证实是管理员注册表的关键要素 (见 4.8 节), 并且 Burnhill 详细讨论了它们在保存活动中的作用 (Burnhill, 2013 年)。

## 3.3 最佳实践方法

### 3.3.1 实践中的转移编码

目前使用的转移编码 (Transfer Code) 是第二版本 (2013 年正在进行第三版草案更新和扩展), 它创建了一系列标准, 适用于所有期刊由一个出版者传输到另一个出版者。出版商公开注册转移码, 并将它应用到实践中, 被视为“传输标准”。

作为业内人士自愿遵守的“最佳做法”守则, 它并不取代合同条款或知识产权。在所有期刊转移中, 都以参与转移双方协商的合同条款为准。然而, 签署转移编码的出版商, 选择在商业上做适当的努力, 以确保期刊转移的合同与该守则一致。

英国期刊集团 (UKSG) 经营免费的增强传输报警服务, 出版商和图书馆均可使用。它旨在作为一个基本的报警机制, 以促进正在计划或进行中的期刊转移项目间的沟通。最近, 为了减少在转移通知过程中的延迟, 及手动跟踪所需的努力, 知识库和一群研究合作伙伴一直在监测出版商和期刊以研究自动保存。

### 3.3.2 在线期刊论文补充材料的推荐做法

越来越多的补充材料被添加到期刊论文中, 但到目前为止, 没有一套公认的做法来指导如何选择、传递、发现和保存这些资料。确保期刊论文补充资料的有效访问、使用及长期保

存, 需要前期规划好永久标识符、元数据、文件格式和封装。NISO 和 NFAIS (全国先进信息服务联盟) 的联合项目, 就出版商收录、处理、显示和保存期刊论文补充资料, 制定了一个推荐实施规程。在线期刊论文补充资料的推荐做法(The Recommended Practice on Online Supplemental Journal Article Materials)、元数据架构、标签库和标记示例都可以从 NISO 网站上获得。

### 3.3.3 知识库和相关工具 (KBART) 和开放链接 (OpenURLs)

基于 OpenURL 标准的链接解析器工具, 可以帮助图书馆用户连接到他们机构的电子资源。驱动这些工具的数据被存储在知识库中。

为了使他们的用户最大限度地有效访问他们所授权的电子资源的副本, 图书馆建设、购买或许可链接解析器知识库。在网络环境中, 给最终用户提供适当副本的内容, 链接解析器在这方面发挥着关键的作用, 尤其是电子期刊过期停定或原来的发行者不再提供内容访问后。在 2006 年, UKSG 委托进行的研究, 确定并描述了在现在知识库中, 存在的影响开放链接系统效率的一系列问题。该报告 (SIS, 2007) 建议创建一个小组, 以推动“最佳实践”解决方案的确定。

知识库和相关工具 (KBART) 工作组由 NISO 和 UKSG 成立, 以完成包括代表出版商、图书馆、链接解析器和 ERM 供应商、订阅代理及电子期刊存档机构的报告。它于 2010 年首次发布建议和最佳实践。

## 4、服务

### 4.1 综述

本章提供了电子期刊保存和持续性访问的简要综述, 也包含了英国国家图书馆和英国国家版权图书馆的一些详细说明。

### 4.2 英国国家图书馆、英国及爱尔兰版权图书馆

英国国家图书馆是世界上最大的研究型图书馆之一, 在英国和世界上都非常活跃。它是英国和爱尔兰六个依法存缴图书馆 (英国国家图书馆, 苏格兰国家图书馆, 威尔士国家图书馆, 剑桥大学, 牛津大学和都柏林三一学院) 网络的一部分。

### 4.3 CLOCKSS

CLOCKSS (Controlled LOCKSS) 成立于 2006 年, 它是图书馆和出版商进行合作的非盈利组织。它是一个以 LOCKSS 软件 (LOCKSS 将在下节介绍) 为基础的 dark 存档系统, 该系统允许数量有限的图书馆代表广大社区扮演存档的角色。

#### 4.4 KB e-Depot

KB 是荷兰国家图书馆的简称, KB 管理着 e-Dpot 系统的运转。e-Dpot 保存了荷兰电子出版商的国家呈缴集和其它电子内容(如, 荷兰国家网站及来自荷兰国家图书馆数字化工程的主图像文件)。

#### 4.5 LOCKSS

LOCKSS (Lots of Copies Keep Stuff Safe)项目给图书馆提供了开源的工具及支持, 这样, 图书馆就可以在本地对各种各样的材料进行监管, 这些材料包括订阅的和开放获取的学术资源(书, 期刊等)。

#### 4.6 Portico

Portico 成立于 2005 年, 为以电子格式发表的学术文献提供第三方服务, 目前分别为电子期刊、电子书和数字化历史集提供准专门的保存服务。

#### 4.7 联合托管

小部分地区联盟也为电子期刊的保存和访问提供并组织了他们的托管服务。

#### 4.8 Keepers 注册系统

Keepers 注册系统是一个试点服务, 在保存服务方面它提供了易于访问的包含电子期刊的信息, 并且突出了那些没有存档安排的电子期刊。

### 5、情形

为了阐述本篇报告中讨论的潜在挑战与问题, 这里列出了六种情形(不作具体介绍)。

第一种情形: 图书馆取消订阅一些期刊

第二种情形: 图书馆取消订阅那些捆绑在一起的期刊

第三种情形: 无法从出版商那里获得过去的期刊

第四种情形: 出版商停止运营

第五种情形: 出版商运营过程中出现灾难性的故障

第六种情形: 缩小印刷电子期刊的范围

### 6、信任

#### 6.1 背景

在这篇报告之前, DPC 论坛上已经讨论了关于数字保存机构信任需求的问题, 参与讨论的人员认为保存机构建立信任需要做到以下几点:

- (1) 有一个清晰的保存定义和任务;
- (2) 在服务交流中明确 DP 意味着什么, 管理部门和运营部门应该支持这样一种观点——保存与获取这二者中, 保存应该处于一个优先地位;
- (3) 公开所保存的资源以及机构是如何保存这些资源的, 期刊的保管状态应始终可见。应该参考利益相关方的意见, 提供相关工具满足上述要求;
- (4) 由第三方审核并把审核结果呈报给利益相关方, 以此扩大信任的开放性;
- (5) 机构能够随着时间推移持续地提供可靠的分发服务, 并能够有效回应相关事件需要的解决方案;
- (6) 说明机构的目标及如何跟进。这意味着能够应对濒危的内容, 能够证明在紧急时刻迅速采取有效措施;
- (7) 展示研究未来需求的承诺和能力;
- (8) 证明能够提供有效和批准的对于各种新内容的解决方案, 同时新方法也在研究中, 这需要研究机构相互合作。

信任可以认为是理念与现实的结合体 (其实这二者之间往往仅存在微小的差别)。

## 6.2 数字保存仓储库的信任度

一些论坛的讨论一直聚焦在存档的信任度问题以及建立评估保存仓储信用程度的标准与客观方法, 信任度已经成为近年来一系列行动的焦点。这些行动可以帮助图书馆、学者、出版商等选择他们要依靠或者支持的数字保存仓储。现在有两个著名的信任度审计报告: 一是 TRAC 及其后来的审核报告; 二是 TDR。

TRAC 包含了帮助判别保存仓储在管理与政策、对象处理及技术领域能力的指标。TRAC 是由 OCLC 下属的图书馆研究小组 (RLG) 和美国 NARA 赞助的一个专案小组发展而来的。TRAC 也借鉴了 CRL 的保存仓储在 2005-2006 年度关于信任评测审核的一些经验和发现, 这些实验审计的对象包括一系列的电子期刊存档服务系统, 特别是 LOCKSS、KB 和 Portico。我们能获取 LOCKSS 在 2006-2007 年度的评测审计报告(CRL, 2007)、Portico 在 2010 年第二次评测审计报告(CRL, 2010)和 Scholars Portal(CRL, 2013)这三次审核结果。TRAC 最终的审核清单于 2007 年发布。

TDR 采用 ISO 16363 标准, 它于 2012 年 2 月份发布, 是 TRAC 审核清单的修订版。相较于 TRAC 审核清单, TDR 的许多变化仅是一些结构性的调整, 但它所强调的核心范围与 TRAC 是一致的。目前 CCSDS 工作小组领导正在准备一份补充标准 (ISO16969), 是关于提供可信赖仓储系统审计和验证服务的机构的要求。



在存档认证这一领域还有一些其他的正式措施,如 DSA 和德国保存认证信任度标准(DIN31644)。还有一些仓储内部自我评估的工具和方法,包括基于风险评估的数字保存机构审核方法(DRAMBORA)、可信赖的数字保存仓储的计划工具(PLATTER)及由 IBM 发明的数字资源长期保存的自动化评估工具。

2010 年,建立了关于数字保存仓储审核与认证的欧洲框架,它是由 DSA、CCSDS 的保存仓储审核与认证工作组、以及德国信任保存认证标准小组合作建立的,它的目的是支撑一个由下面三个层面组成的关于审核与认证数字保存仓储的完整框架,以此来增加保存仓储的信任度:

基本认证颁发给获得了 DSA 认证的保存仓储;

有了基本认证之后,如果一个仓储能够有一个结构性的外部审核以及基于 ISO16363 标准或 DIN31644 标准的公开化的自我审核,那么它将会有有一个进一步的认证;

当一个仓储在得到基本认证之后,又充分地得到了基于 ISO16363 标准或 DIN31644 标准的外部审核,那么它将得到正式的认证。

除了这些认证之外,还有大量基于保存仓储审核过程自动化和“信赖工程”方法的项目。在这些项目中,值得注意的是安全存档系统(SafeArchive System),它被设计用来建立位于 LOCKSS 点对点复制网络顶部的虚拟覆盖网络,也被用于支持配置、监控以及基于 TRAC 的审核。安全存档系统可以通过识别自动验证的具体的 TRAC 标准来自动遵从正式的复制与保存政策,也可以通过完整文档的支持来遵从一些其他的标准。

### 6.3 经验教训

APARSEN 项目总结概括了保存仓储认证的过程中的经验教训。早期开发仓储认证的工作往往依赖于大量的捐赠资金补贴和机构的重大费用,从大量的实验项目及通过当前有关认证分层和简化的、标准化的方法的努力所得来的经验减少了这种现象的发生。然而,尽管取得了相当大的进步,但是一些富有争议性的审核程序还没有完全消失,在审核人员与保存仓储中都还存在这一些问题。

审核保存仓储的工具与程序对于提高数字保存仓储的质量规格和明确最佳实践有很大的帮助。可信赖的公开审核报告能为用户和投资者提供关于数字保存仓储在特定时间、特定存储库的活动的信息和保证。

审核报告与工具可以在不同的环境(内部、外部)中产生或运用。Seadle 提出了一套有用的定义会对分类工作有益:

公开测试：公开测试是指 a) 资源没有原始参与的人群也能复制；b) 其测试结果已经发布在公开的期刊与网站上。

私人测试：私人测试是指 a) 如果没有最初搭档帮助的话，那么就不能够复制；b) 其结果也仅仅只有一小部分人能够获取。

半公开测试：半公开测试是指 a) 只有得到最初搭档帮助才能复制；b) 其结果则已经发表在公开的期刊和网站上。

正如在 6.1 提到的那样，关于电子期刊保存信任度的讨论很大程度上集中在讨论保存仓储的信任度上。然而，就如在前面几节陈述的那样，存档仓储只是提供电子期刊长期获取的一个要素。许多关于制订协议透明度和规范步骤的相同要求在作者、图书馆和出版商之间同样适用。其他的一些措施，如转换行为规则 (Transfer Code of Practice)、Keepers 注册系统 (Keepers Registry) 及在许可证中的示范条例与制定机制的发展同样是电子期刊保存信赖的关键因素。为了建立信任度，应该建立关于长期保存的明确的协议以及清晰的步骤和机制，确保这些协议能够有效实行和能够验证供应链上的所有元素。

## 7. 进展扫描

自始至终，这份报告都试图呈现现有环境中电子期刊保存的发展情况，以及突出了发展过程中正在出现或现有的对未来多年内电子期刊发展具有影响的问题。有两个可能影响未来电子期刊存档发展趋势的问题值得反复思考。

早期电子期刊的单一、静态的 PDF 格式文件正在逐步被分布在互联网上的绑定了动态且相互联系的资源所取代。现在，在电子期刊上涉及的那些复杂的实体为了便于理解而自动扩展显示图表或图片，那些增添到文档中的辅助素材可能链接到或分布在文章中。这些辅助素材包括从实验中得出的研究数据、图像、视频、音频记录及演示的材料。在一些社交媒体上 (如博客、线上讨论小组和 Twitter 等)，对一些学术工作所做的摘要性内容或讨论的内容会极大地扩大这些学术工作的知名度，也可以让人们更好地评估这些学术工作以及把这些学术与其他的一些信息资源关联起来。从长远来看，可能更多的重点会集中到绑定了链接内容的研究对象上。

这些改变呈现出关于信息互操作与长期保存的技术挑战以及管理、学术记录描述甚至是记录本身版本概念有关的新挑战。

把网络完全认同为学术交流的基础设施的趋势要求我们相应地调整未来的保存工作。这也就凸显出了为什么参与 DPC 电子期刊保存工作的人员一直强调在电子期刊保存工作中把对获得信任度的研究开发工作放在重要位置。

第二个值得注意的问题是电子期刊的开放获取及其影响的发展趋势。这里出现的挑战包括明确开放获取文档的适合的存档权限及保存开放获取期刊的所需经费。最近一项研究调查了在开放存取目录中列出的期刊的存档。研究发现在 DOAJ 所列期刊名称中只有百分之八出现在 LOCKSS 或 CLOCKSS 中, 只有百分之五的名称出现在 Portico 中。荷兰国家图书馆与 KB 签订了一项协议, 会把 DOAJ 中有全文的期刊的百分之五十存储在 KB e-Depot 中, 然而, 完成这项工作估计得花八年的时间。这也即意味着在 DOAJ 中列举出的绝大多数开放获取期刊并没有得到有效的长期保存。

2009 年, DOAJ 收集的期刊 (4000 种期刊) 由大量的出版商 (2000 多家) 来出版, 每个出版商在不同的平台、用不同的格式及超过 50 种语言出版一小部分期刊。许多出版商在面临需要可持续性的财务、技术及管理支持时显得非常脆弱。为了保存这些开放获取期刊而做更多的努力, 以及与此同时维护支持保存仓储运转的商业模式是在未来所存在的另一个主要挑战。随着文章获取和权限水平的提高, 保存将会变得越来越复杂, 例如, 那些糅合了能开放获取与用户专属文章的混合型期刊。

## 8、 总结

电子期刊的持续获取和保存的活动行动还包括组织机构之间的合作、发展第三方服务及在不同的利益相关方之间建立关于长期保存与获取的信任度。完美地解决这些问题不仅仅需要科学技术, 法律、经济发展、服务发展对解决这些问题具有同等重要的作用。

电子期刊保存需要在未来得到进一步的发展, 以解决目前学术交流模式以及出版商业模式的持续变化问题。

对图书馆来说, 能够提出一些能满足他们的电子期刊保存要求的归档办法的标准草案:

- (1)、出版商、期刊的数目和类型以及保存资源的年度范围。
- (2)、存档的覆盖面与所需订阅的期刊内容的匹配程度
- (3)、摄入的保存内容应该有质量保证, 以确保没有遗失的问题或是损坏的内容。

服务条款及质量:

- (1)、对订阅期刊内容所需要的连续获取的条款的支持程度。

在取消订阅或其他的一些突发事件条件下, 那些对确保继续访问的可接受的一些条件。最小的时间延误和清晰的结果界定。

(2)、存档机构明确承诺学术同行评议期刊的长期保存。开放数字存档的正式认证, 并满足广泛认可的最佳实践标准。

(3)、应当提供适当水平的支持。期望从图书馆或机构获得的 IT 以及保管的任何水平的支持必须是可以接受的。

管理与经营模式:

(1)、利益相关方参与到保存机构的管理当中。

(2)、有足够说服力的组织, 有长期、可持续性的投资, 有足够的人力资源和足够的专业技能。

(3)、能够承受且物有所值。

(4)、在电子期刊存档中仍然存在许多挑战, 但是对广泛的数字保存社区以及图书馆和出版商来说, 还是获得了显著成功以及经验教训。

## 9、 建议措施

### 9.1 对于学术图书馆、图书馆协会与图书馆联盟

如果在本地保存电子期刊内容或合作保存的话, 应该对存档提供一个合适的数字保存策略和实施步骤; 从图书馆信息技术系统或工具 (如关联解析知识库) 来整合获取存档的数字内容。

如果把保存业务外包的话, 需要评估电子期刊存档方案如何匹配你对覆盖程度、服务项目和条件的需求。

评估电子版潜在的利益和节省的资金, 计算电子期刊保存的费用, 去除现金预算中 (如绑定) 的节省资金。

制定执行条款以使图书馆收藏、发展战略与政策的具有延续性。

分配员工尽职完成上述活动, 并确保人员的相应知识及工作流程到位, 能够执行所制定战略与政策。

与同行一起保存电子期刊, 在选择和服务问题上投入以服务提供者和成员机构。

为了取得对持续访问和相关电子期刊保存的解决方案的许可而不断努力。

依照电子期刊资源许可的持续/永久的获取权力不断收集数据。

### 9.2 对于出版商

本作品采用[知识共享署名-非商业性使用-禁止演绎 2.5 中国大陆许可协议](#)进行许可。

至少采取一个来自外部的电子期刊保存方案。

制定和公布关于永久获取和长期保存的政策。

在图书馆和出版商之间签订的许可协议中, 应该明确过期停定获取权力的条件。

当一些期刊转让给其他出版商时, 应该遵循 Transfer Code of Practice 标准。

如果是开放获取出版商, 可以运用 CC-BY 许可条例 (这个许可条例对用户和保存机构都具有最佳的友善度), 这个条例允许在其他方面做长期保存和编辑文档与数据挖掘的工作。简单、一致的许可条例实质上减少了保存开放获取资源的费用。

### 9.3 对于电子期刊保存机构

不断提供关于所保存的资源和新内容的最新信息。

向其他机构提供一种可以向保存机构推荐新资源 (内容) 的方法。

在馆藏发展政策方面 (这些政策依据被评估和优化的新采购需求设立了标准) 咨询相关利益方。

与出版商保持 (专指那些有权对大量期刊进行管理和提取相关内容的出版商) 正式关系。

对学术具有长久的、国际性重要影响的电子期刊保存应该至少具有两个国际节点 (镜像服务), 这些节点应该分布在不同的大洲和地区, 这样就可以最小化保存的长远威胁。

制定公开获取的政策与程序文件。在这些政策和文件中应该对获取的条件有明确的陈述和定义, 包括过期停定获取与突发事件等。

在工作中, 尽量多地寻找一些小型或中型的出版商和开放获取期刊。

保持积极的研究开发工作以更好地应对未来在电子期刊出版、学术交流、保存工作中出现的挑战。

编译自:

<http://www.dpconline.org/newsroom>

(陈瑶 王敬编译, 唐果媛 陈瑶 吴振新校对)

## 【资料推荐】

### 《数字保存的法律信息》

在经济和文化领域, 数字保存是不可或缺的。对每个公司来说, 不管部门多少和规模的大小, 保存业务过程都是一个新兴的挑战。作为TIMBUS项目的一项成果, 《数字保存的法律信息》这本书关注数字保存的相关法律信息并提供数字保存领域的法律指导。

这本重要的书揭示了保存活动在知识产权和数据保护方面的影响。包括：存储个人数据时可能会违反数据保护法，复制和存储有版权的数据时可能会侵犯版权所有人的专利权。

这本书考虑了在 IP 和数据保护权利下的保护范围，同时提供了避免侵权的策略。技术合作问题和履行现有的法律义务来保存数据被认为是数字保存的两个重要方面。

清晰明确的法律框架以及对数字保存法律信息的详细分析将会给正在给保存机构执行数字保存业务流程的咨询公司带来巨大的用处，同时也会给那些公司自身、保存系统的开发人员以及数字保存领域的研究者们带来很大的用处。

目前该书无免费版本，购买地址：[http://www.e-elgar.com/bookentry\\_main.lasso?id=15401](http://www.e-elgar.com/bookentry_main.lasso?id=15401)

编译自：

<http://timbusproject.net/news/news/232-book-on-legal-aspect-of-digital-preservation>

(王敬编译，陈瑶 吴振新校对)

## 对图书馆、博物馆及历史遗迹保存机构保存活动的调查

Primary Research Group 这篇报告的数据来自于 55 个机构，包括博物馆、图书馆、大学、历史遗址及其他主要活动集中在保存稀有数据或历史书籍、文档、电影、视频记录等等的机构。

这项研究密切关注上述机构是如何开展保存活动的，包括人员编制、资金、预算、授权策略、职员培训、合作、设备采购、业务外包、收入、经营、社交媒体与网络的使用，及在图书馆、博物馆、政府机构、历史遗迹等机构的其他保存管理与修复活动。

下面是重要研究发现中的一些方面：

大约有四分之一的受访者有机器保存和修复的经验；

对 30.91%的机构来说，保存或改进现有的数字格式或媒介是一个重要的因素；

博物馆平均每年的保存预算是 45250 美元；

机构越大，对资助依赖就越少，来自于自身的保存经费所占的比重就越高；

21.82%的机构有预算用于对员工进行保护和保存技术方面的培训和教育。

该分析报告目前还不能免费获取。

编译自：

<http://www.primaryresearch.com/news.php?id=118>

(陈瑶编译，唐果媛 吴振新校对)