

● 宛 玲 张晓林

数字资源长期保存过程中的知识产权问题分析*

摘 要 根据开放档案信息系统参考模型,图书馆需要解决长期保存资源摄取环节、存储环节和服务中的知识产权问题,需要建立相应的管理机制。参考文献 13。

关键词 图书馆 数字图书馆 数字资源 知识产权
分类号 G253

ABSTRACT According to the reference model for an open archive information system (OAIS), library should solve the problem of intellectual property in the ingestion, storage and service of the archival storage resources, and should establish related management mechanisms. 13 refs.

KEY WORDS Library. Digital library. Digital resource. Intellectual property.

CLASS NUMBER G253

众多图书馆已经采购或生产了大量数字资源,它们大多具有长期流传和持久使用价值,需要长期保存。然而,数字资源长期保存面临许多知识产权问题。数字资源长期保存,其目的是要保证未来用户可以检索、获取、解读、浏览、理解、解释各种数据和记录^[1]。数字资源长期保存是一动态过程,涉及的知识产权问题和如何解决这些问题与整个保存过程密切相关。开放档案信息系统参考模型(Reference Model for an Open Archive Information,简称 OAIS)为描述数字资源长期保存的整个过程提供了一个基础框架^[2]。OAIS 将数字资源长期保存过程分为摄取、数据管理和长期存储、获取 3 个主要环节,并提出长期保存计划和系统管理两个功能模块。如果我们将长期保存权定义为图书馆对特定数字资源进行收集保存、管理并在一定条件下提供服务的权利的话,那么 OAIS 所提出的每个环节的行为都涉及相应的知识产权问题,可以相应地归纳为存档权、存储管理权和服务权,需要在计划过程和管理机制中通过相应的法律、技术和管理措施予以保障。

1 长期保存资源摄取环节的知识产权问题

“摄取”是指图书馆从资源生产者或所有者处获得或接收要长期保存的数字资源过程。从知识产权管理的角度,长期保存过程中首先要解决的是,图书馆是否获得对数字资源进行收集、复制和长期保存的权利,即所谓长期保存中的存档权。存档权不仅涉及

直接的内容数据,还往往涉及以下两类数据:①与内容数据的存储和使用相关的“元数据”(此处我们对元数据做广义的理解),例如关于数字格式与处理方式的信息,关于编码、格式、标记、结构、压缩、加密的数据,关于内容校验、身份认证、版本演变、知识产权管理方面的数据等,确保能够完整识别和解析内容数据;②支撑运行系统,包括相关的软件甚至硬件系统、知识组织体系、利用与服务环境(例如链接环境)等,确保能够有效利用内容数据来“复原”原有的数据服务。后两类数据的保存往往成为有效使用内容数据的基本条件。

数字资源的存档权与其生产方式、获得方式和资源体系制度安排相关。

(1) 自建数字资源,包括对馆藏文献数字化和创建自主知识产权的数字资源。对于没有版权的馆藏资源数字化后形成的资源,图书馆拥有长期保存的权利;对于有版权的馆藏资源,在资源本身属于独有资源、资源的保存和获取难以通过其他渠道予以保障的情况下,可以引用我国著作权法第 22 条第 8 款规定,“图书馆、档案馆、纪念馆、博物馆、美术馆等为了陈列或者保存版本的需要,复制本馆收藏的作品”,对有版权的馆藏资源合法地数字化,合法取得存档权。但这时对存储管理权和服务权可能仍有限制。如果数字化授权只是为了在一定范围提供网络传播,或者数字资源本身有其他多种途径进行保存或获取,数字本身(即使合法)也不一定自然就包括存档权,图书

* 本文受国家科技图书文献中心“数字化科技信息资源长期保存体系与政策机制”项目资助。

馆为了进行长期保存必须在获得数字化授权的同时寻求存档权。对于自主知识产权的数字资源,包括自己制作的网页、数据库、元数据、软件系统等,图书馆应有长期存档的权利。

(2) 外采数字资源,主要指图书馆通过订购获得数字资源的使用权,这些资源可以存放在图书馆本地镜像系统、联合采购集团的异地镜像系统、资源商委托的第三方异地镜像系统、或者资源商自身的远程数据库系统。多数情况下,图书馆对外采数字资源只有网络使用权(包括那些装载在本地镜像的资源),一旦采购合同终止或发生天灾人祸等,都可能造成图书馆丧失对数字资源的可持续使用权,因此这些资源是长期保存的重点对象。外采数字资源的存档权,包括所采购数据的永久使用权和数据备份权,获取和保存“与内容数据的存储和使用相关的数据”和“支撑运行系统”的权利。这里存在3个问题:①“永久使用权”并不等于存档权,图书馆需要明确寻求本地数据备份权和在一定情况下对这个备份数据的备份权(尤其是多个图书馆合作进行长期保存时)。②常规的采购合同往往只涉及内容数据,权利实施又要求对“元数据”和“运行系统”有缜密的技术分析、检验和实现方式,所以需要通过专门机制来系统地诉求相应的权利。③由于外采数字资源本身规模庞大,采购单位众多,长期存储成本很高,所以每个或单个图书馆独立地对外采数字资源长期保存往往不可行,需要在集团采购基础上建立联合存档系统甚至国家存档系统来合理有效地进行长期保存。

(3) 网络公开资源,一般指可以在网络上公开使用的数字资源。可以分为两类:①未在版权保护期内的作品以及按照著作权法不给予保护的公共资源;②开放资源,指著作权人允许人们通过网络使用的资源,但“允许使用”一般在使用方式上和使用目的上有限制,例如“供个人学习、研究使用”、“不能生成衍生作品”等^[3]。对于第一类资源,理论上图书馆可以对其进行下载存档,但一则需要有“符合公共利益”的责任体系来支持存档的合理性,二则由于网页设计可能有版权,因此图书馆可能需要获得网站版权所有者授权才可以下载数据内容予以保存。对于第二类资源,一般需要获得著作权人专门授权才能进行长期保存,但由于它数量多,本身又缺乏“采购”或者“存缴”这样的有形机制,比较难以安排相应的存档授权。一种可能的途径是

争取作为这类资源的镜像站点(可能需要支付一定费用),通过获得网络传播权来争取存档权;另一种可能是,在特殊情况下,根据“公共利益”原则可以建立公共服务镜像站点,收集和存档开放资源,但这样做的法律依据、“公共利益”条件和权益安排还有待进一步讨论。还有,开放资源本身在理论上已经是公共资源,图书馆对其长期保存后的服务权也可能需要限定。

(4) 缴送数字资源,即按照有关法律或行政规定由资源生产者、拥有者向指定机构缴送的数字资源,包括按照法定缴送制度向相应级别公共图书馆或国家图书馆缴送的电子出版物,也包括职务作品(或参照职务作品处理的其他作品)著作权人向职务所在单位缴送的电子版作品,例如向学位授予单位缴送的学位论文、向本单位的机构存储库(Institutional Repository)缴送的论文和报告等。对于这类数字资源的存档权,首要的问题是缴送制度是否包括数字版本的缴送。目前许多国家还没有形成完善的制度来保障数字资源的缴送。2002年欧洲国家图书馆议会对欧洲100多家出版商进行了问卷调查,反馈数据显示1/3国家现行的缴送制度中没有包括数字资源^[4]。对于职务作品的缴送制度,一般通过行政法规或机构规章予以确定,目前国内外已有部分科研机构和政府提出了一些规定,例如我国部分高校建立了学位论文电子版存缴制度,美国国会一个小组委员会要求国家卫生研究院资助的研究项目论文应该在发表后一定期限内向公共PubMed数据库缴送^[5],但在职务作品数字版本缴送方面,总体上仍处于制度缺失、权利规定粗糙、保障机制不完善的状况。在已经建立包括数字版本缴送制度前提下,缴送制度已经规定了所指定的存缴本图书馆(包括本单位职务作品的存缴本图书馆)对缴送数字资源的长期保存权,但实际收集、元数据收集或制作、资源整理和服务等方面,还存在大量的操作与权利问题。

从上面分析已经看出,在许多情况下图书馆必须得到授权才可获得存档的权利。但数字资源的版权关系比较复杂,图书馆有必要针对具体情况分析具体的权利归属,例如作品本身的复制权、作品数字版本的复制权、作品或作品数字版本的信息传播权和邻接权、作品或作品数字版本的支持软件的著作权等。

2 长期保存资源存储环节的知识产权问题

图书馆在获得合法的存档权、对数字资源进行收集或采集后,需要利用数据管理模块对它们整理并编制长期保存元数据,然后将数字资源和元数据存入长期存储系统,并根据需要维护被存储数据的持久可利用性。在最初的存入过程中,图书馆需要对数字资源进行拷贝、析取元数据并可能进行整理,而在长期保存管理过程中,图书馆可能需要对数字资源进行刷新(Refresh)、迁移(Migration)、仿真(Emulation)等处理。在整个存储和管理过程中,必然涉及一系列的知识产权问题,我们称之为存储管理权。

数字资源的长期存储管理可以采取多种技术策略,例如拷贝数据方式、保存管理技术方法、保存内容深度、保存系统模式等,它们从不同侧面涉及到不同的知识产权问题。

(1) 拷贝数据方式的影响。①原样拷贝方式,不做任何改变地将数据对象复制到简单的文档存储系统中,主要涉及复制权问题,尤其是周期性拷贝过程产生的多次复制权。②将数据拷贝到数据库系统中,如果存储系统采用专门的数据格式、数据库系统和软件系统,涉及复制的同时还往往涉及代码、格式、结构、标记的转换和对元数据的析取,还可能由于更换了系统平台,需要解决软件兼容问题,可能需要解析版权保护技术措施来保证资源的可靠保存和读取。③将拷贝的更新数据组织到已经存储的资源体系中,可能需要对数据进行对比过滤,在新老数据间建立关联,对部分数据内容进行删减,对部分版权信息进行修改等,从而影响保护作品完整权、修改权和版权信息完整性。

(2) 保存管理技术方法的影响。目前主要的保存技术方法包括仿真、迁移、更新等。仿真方法是通过详细描述数字信息利用的各种技术参数和环境条件,使未来计算机系统可以模仿原始系统环境来读取和使用数字信息内容^[6]。从工程角度,仿真有可能使功能和显示方式产生丢失^[7],这就涉及保护作品完整权问题。迁移方法是将数字资源迁移到不同的软件或硬件环境中,从而保证数字资源可以在发展的环境中被识别、使用和检索。这涉及复制权以及过量复制问题,而且可能涉及对代码、格式、结构、标记等的转换,可能涉及新的元数据析取。迁移技术在保存复杂数字对象时,容易导致外观、感觉、功能、超文本链接甚至内容方面的一些丢失,

这些都涉及保护作品完整权和修改权。如果版权所有者的版权管理信息出现丢失,还涉及保护版权管理信息权等。更新方法主要是从现有存储媒体上将数据拷贝到同类型的新存储媒体上,或者从一种存储媒体拷贝到另一种存储媒体上,这里主要涉及复制权以及过量复制问题。版权保护技术措施问题使保存管理复杂化。从功能上讲,版权保护技术包括访问控制型、使用控制型、使用监督型和安全保护型^[8]。无论采取哪种保存管理技术,都可能受到原有作品版权保护技术的影响。

(3) 保存数据内容程度的影响。①对数据的保存,按保存程度可以分为3个层次:一是仅保存数字比特流,即纯粹的内容数据;二是与此同时还保存与内容数据的存储和使用相关的元数据;三是同时还保存数据的支撑运行环境。②由于数字资源的利用依赖对其格式和处理方式的理解,所以仅仅保留纯粹的内容数据不能够保证长期保存的可靠性。极有可能一段时间后就没人能够清楚知道被保存数据的内容结构、加密方法、压缩方式、标记方式等,被保存数据变成“死数据”。因此,收集和保存与内容数据的存储使用相关的元数据是长期保存权的基本要求和组成部分。实际上,提出并建立这些元数据的获取权利和获取机制的过程本身,会很自然地引发对存储处理权的几乎所有方面的系统讨论和协商。图书馆应充分利用这个过程来澄清自己在长期保存中的存储管理要求和权利,避免权利缺失或侵权。③由于建设或购买数字资源的目的是利用它们提供相应的信息服务,所以数字资源与相应的服务系统已经很难区分开来,特别是当前数字信息系统利用其知识组织机制、链接机制、交互处理机制等提供了丰富的服务功能,仅仅保存纯粹的内容数据往往不能“复原”原来系统提供的信息服务,从而影响真正意义上的数字资源的可持续使用。鉴于此,数字资源长期保存还需要保存数据的支撑运行环境。但支撑运行环境的保存本身还存在很多挑战:首先,什么构成一个服务的支撑运行环境?这个问题的答案在不同的数字信息系统中往往是不同的;其次,如何析取“支撑运行环境”的组分?在许多系统中这些组成被互相“硬绑定”(hard-binding)在一起,难以区分、确认、复制、检验、保存和重新装载;再者,这些组分(即使能析取)可能涉及多个版权所有者的(实际上元数据也可能如此),资源商可能没有权利将它们的长期保存权利授予图书

馆,而且“支撑运行环境”更容易涉及资源商或服务商的技术秘密,更容易被用来对内容数据进行整理或修改,获取对“支撑运行环境”的长期保存权反过来也需要图书馆对自己的责任进行明确界定。

(4) 保存数据系统模式的影响。根据长期保存系统“复原”服务的程度,可将它们分为隐暗存储系统(Dark Archive)、半透明存储系统(Semi-Dark Archive)和透明存储系统(Light Archive)^[9],它们也会带来不同的知识产权问题。①隐暗存储系统直接利用资源商所提供的备份数据或下载拷贝作为长期保存机制,提供一种简单的数据备份存储能力,以便在需要将备份数据上载到一定系统上提供服务。这种方式主要是涉及内容数据的复制权。②半透明存储系统,在提供内容数据备份存储能力的同时,往往提供与内容数据的存储和使用相关的元数据,提供对内容数据进行定期上载、确认、使用检验的机制,往往要求资源商允许图书馆为保存需要将备份数据迁移到自己的存储系统上,并定期维护中测试数据及其相关功能的完整性和可使用性。当然,这样一来不仅增加了图书馆的保存成本,而且提出了存储处理权中的许多新方面,包括在网络环境下对数据与服务进行测试时可能涉及的保护数据完整权和网络传播权,需要比较复杂的权利转移和保障机制。③透明存储系统,实际上是一个与资源商服务系统同步的存储环境,具有相同的数据和基本相同的服务能力,从而保证在任何时候能够及时“复原”相同服务,这时,不仅需要解决前面所提到的存储处理权中的各个方面,而且还涉及双方更复杂的权益约束。为了有效实现数字资源长期保存的目的,必须至少建立一个半透明存储系统,保证在需要时能够在一定时间延迟内复原内容数据和基本的数据服务;对于关键资源,应该通过合作建立透明存储系统。因此,建立相应的相关知识产权管理机制非常重要。

3 长期保存资源服务中的知识产权问题

图书馆并不能随意将自己长期保存的数字资源不加限制地提供服务。利用长期保存数据提供服务不但将可能侵犯版权所有者的复制权、信息传播权等,而且还往往影响到他们的经济利益和其他利益(例如数据的“惟一发布者”和“核心服务者”权利)。因此,需要平衡知识产权、商业利益、公共利益和公平机制等方面的因素,明确界定、获取和

保障图书馆在一定条件下将自己长期保存的数字资源向一定范围的用户提供一定形式的服务权利,我们称这为服务权。根据知识产权法律的精神,在市场经济环境下,利用长期保存数字资源提供公共服务时,不能损害版权所有者的合法权利,不能破坏数字资源和信息服务市场的正常运作,不能利用所保存的数字资源不当得利,不能阻碍数字资源的正常传播和信息服务的正常发展。因此从一定意义上讲,利用长期保存数字资源提供服务应该是一种特例。服务权需要界定:在什么情况下能够向用户提供服务?在这种情况下能够向什么范围的用户以什么方式提供服务?这些服务还必须受到什么限制?

(1) 提供服务的许可条件及许可范围。①当采购合同终止时,图书馆需要拥有继续利用已采购资源提供服务的权利。由于图书馆借采购合同获得的长期保存权在用户范围上与合同本身规定的使用权一脉相承,由于资源商可能还与其他图书馆维持合法、正常的合同关系,因此这个图书馆只有权利利用长期保存资源向原采购合同规定的用户范围提供服务。②当因天灾人祸出现供给失效时,图书馆需要拥有利用长期保存数据提供服务的权利,同时为了保护市场机会、促进市场发展,也为了社会公众的利益,如果市场上还有同类资源及其服务(简单供给失效)时,图书馆应该拥有利用长期保存数据向原采购合同规定的用户范围提供公共服务的权利。当市场上没有同类资源及其服务(完全供给失效,也称市场失效)时,作为公共服务机构,尤其是那些在全国或某一领域具有全局性公共服务地位的图书馆,或者是承担存储资源长期保存任务的图书馆,应该拥有利用长期保存数据向社会公众提供公共服务的权利。③存缴本图书馆或负责联合存档系统、国家存档系统的图书馆,往往可以提供有限范围的“公共服务”,例如荷兰国家图书馆与 Elsevier 签订的长期保存协议规定,荷兰国家图书馆可以向到馆读者提供所保存的 Elsevier 数据服务^[10];又如美国缴送制度规定,缴送到国会图书馆的 CD-ROM 出版物可以且只能在国会图书馆使用,可以同时有 5 个并发用户^[11]。④比较复杂的是多个图书馆合作建立的长期存档系统的服务权。假定该合作系统委托其中的某个成员(保存馆)负责长期保存备份数据,当其中一个图书馆(需求馆)与资源商终止采购合同但合作的其他图书馆可能并没有终止采购合同时,需求馆(假定非保存馆)在理论上有权利用长期保

存系统中自己采购合同期内数据向自己用户提供服务,但需要从整个长期保存数据中析出该馆“合同期数据”,而且可能不能简单将析出数据拷贝到该馆系统来提供服务(因为这可能形成多套备份数据,违反存档授权)。因此可能要在保存馆技术系统上载这些数据并提供向需求馆的服务,这显然需要得到保存馆的支持。

(2) 提供服务的限制条件。即使图书馆拥有合法的理由利用长期保存数据进行服务,这些服务还受到一定的限制,以保护版权所有者的合法权利和公共利益。这主要涉及两方面的问题:①图书馆在提供服务时,必须保护该数字资源整体署名权,保护作品完整性,保护版权管理信息,尽力保护相关元数据和“支撑运行环境”的完整性,尽力保证数据服务功能的完善性,保证数据不被恶意修改和下载,对用户合理使用数据尽到合理的管理责任,同时保护资源商的有关商业和技术秘密。②利用长期保存数据提供服务的基础是这种服务符合公共利益,图书馆在提供服务(尤其是利用网络公开资源和缴送资源提供服务)时不能以商业营利为目的,不能借提供公共服务来捆绑促销自己商业化的数据产品或服务。合作存档或国家存档系统中的保存馆,必须尽责履行自己对整个合作系统或国家的公共服务义务,不能利用所保存的数据及其服务获取商业利益。

4 建立长期保存资源知识产权管理机制

妥善解决数字资源长期保存过程的知识产权问题,是建立可信赖的数字资源长期保存机制的基础,否则图书馆难以获得资源商和其他版权所有者的积极合作,也难以形成可靠的长期合作保存关系。如何建立长期保存资源知识管理机制,我们将在以后结合长期保存实践进行深入研究。就目前而言,我们认为作为公共的信息资源保存和服务机构的图书馆应该充分利用知识产权法中的合理使用与法定许可、与数字资源版权所有者或拥有者签订长期保存许可协议等方式保证图书馆的保存权利,同时建立可靠的用户使用管理机制^[12],加强对图书馆人员和用户的教育^[13],并积极解决或协助解决可能出现的侵权行为,从而建立良好的法律环境、管理机制和合作氛围。

参考文献

- 1 Thomas Severiens. Preservation Planning, Institutional Strategies and Policies. 中欧数字资源长期保存国际研讨会 (Chinese-European Workshop on Digital Preservation) 会议报告. <http://www.csdll.ac.cn/meeting/cedp/schedule.html> [检索时间:2004-09-16]
- 2 Reference Model for an Open Archival Information System (OAIS) CCSDS 650. 0-B-1. <http://www.ccsds.org/documents/650xObl.pdf> [检索时间:2004-09-16]
- 3 <http://www.creativecommons.cn/index.htm> [检索时间:2004-09-16]
- 4 Catherine Ayre. The Right to Preserve—The Rights Issues of Digital Preservation. D-lib Magazine. 2004,10(3). <http://www.dlib.org/dlib/march04/ayre/03ayre.html> [检索时间:2004-09-16]
- 5 National Institutes of Health (NIH). Notice: Enhanced Public Access to NIH Research Information. <http://grants.nih.gov/grants/guide/notice-files/not-04-064.html> [检索时间:2004-09-16]
- 6 张晓林. 数字信息的长期保护问题. 图书馆, 2001 (5)
- 7 Adrienne Muir. Copyright and licensing for digital preservation. <http://www.cilip.org.uk/update/issues/jun03/article2june.html> [检索时间:2004-09-16]
- 8 李祖明. 互联网上的版权保护与限制. 北京: 经济日报出版社, 2003
- 9 张晓林, 宛玲. CSDL 外购数字文献资源的长期保存策略建议. 大学图书馆学报, 2004 (6)
- 10 National Library of the Netherlands and Elsevier Science make digital preservation history. <http://www.kb.nl/kb/resources/frameset-nieuws.html> [检索时间:2004-09-16]
- 11 Mandatory Deposit of Copies or Phonorecords for the Library of Congress. <http://www.copyright.gov/circa/circ07d.html#deposit> [检索时间:2004-09-16]
- 12 张晓林. 虚拟信息资源体系的用户使用管理. 现代图书情报技术, 2000 (5)
- 13 宛玲, 孟连生. 论 e 学术资源永久使用权体系的建设. 见: 图书馆可持续发展与创新研究文集. 成都: 西南交通大学出版社, 2003

宛玲 河北大学管理学院副教授, 中国科学院文献情报中心及中国科学院研究生院在读博士。通信地址: 北京中关村北四环西路 33 号。邮编 100080。

张晓林 中国科学院文献情报中心博士生导师, 教授。通信地址同上。 (来稿时间: 2004-10-15)