

数字资源权益状况描述框架研究

刘静羽 黄金霞 王昉 章岑

(中国科学院文献情报中心, 北京 100190)

摘要: 本文通过调研不同类型的数字资源权益状况并结合图书馆业务需要, 提出一套面向服务的数字资源权益描述框架, 包括总框架和针对不同类型资源的补充权益信息, 用于对图书馆所获取的商业数字资源、开放数字资源和自建数字资源的权益状况描述。同时, 应用这一框架提出“建立‘从资源输入到资源输出’的权益信息”响应图书馆服务的工作流程建议。

关键词: 数字资源; 权益; 描述框架

中图分类号: G250

数字资源的权益管理是图书馆业务中的重要组成部分, 特别是当前图书馆普遍重视提升移动获取、精准服务、专业数字资源集成、深度知识标引与检索利用等服务能力, 需要开展对不同数量、类型、来源、深度的数字资源的使用和再利用。为此, 图书馆员需要建立一套面向服务的数字资源权益状况描述框架, 对资源的权益信息予以充分揭示, 从而实现地对所使用的资源权益信息进行登记, 以便快速、全面了解和掌握资源的权益状况, 从而在保障各方权益的基础上制定使用和服务策略。

本研究梳理了当前主要的数字资源类型, 针对每种类型资源, 全面调研了资源内容载体、合同文本、资源所在网站等方面, 提炼出对图书馆开展普遍式文献服务、精准服务、资源深度利用等有影响的关键权益信息, 如著作权人、版权归属、版权时限、授权许可、平台使用限制、元数据版权等, 形成数字资源权益状况描述框架, 并根据这一框架对数字资源服务响应机制提出建议。

1 国内外研究进展与相关实践

随着对信息资源利用程度的加深, 为了保障资源相关主体的利益, 同时最大限度地主张图书馆为公众服务的权益、降低侵权风险, 资源权益状况的确认、权益管理工作受到重视。国内外图书馆界等领域在权益管理领域开展了许多理论研究和实践探索。

国外图书馆和相关组织重视对各类资源的权益状况确认工作, 主要集中在版权方面。例如, 欧盟关注对于孤儿作品的版权状况核实的工作, 并于 2008 年启动 ARROW (孤儿作品版权信息登记与授权系统) 项目。该项目建立了揭示欧盟作品版权状况的分布式网络, 有效解决了对孤儿作品使用的授权问题^[1]。另外, 美

国杜克大学图书馆、加拿大温莎大学图书馆对复杂的 MOOC 资源版权信息进行了清理^[2]。在国内，国家图书馆于 2009 年开展了数字资源版权管理工作，对所拥有的各类资源进行了版权状况梳理。国家图书馆梳理了数字资源建设与服务当中，从资源采选、管理、服务到保存整个生命周期各个业务流程中的版权问题^[3]，建立了版权管理机制、版权管理系统，开展了对国家图书馆各类数字资源的版权状况梳理和确权工作。例如，对自建资源进行版权审查，建立作品版权状态资料库并及时维护、动态更新，对目前还有争议或没有获得有关授权的资源，暂不提供全文服务，仅提供篇名、目次或摘要数据，或严格控制在馆域网范围内使用^[4]。目前国家图书馆还设计并建设了图书馆移动服务版权管理平台^[5]。中国科学院文献情报中心从 2013 年起开展了开放资源的权益管理研究，对开放资源建设过程中出现的关键权益问题和风险规避方案等进行了研究^[6]。同时，新闻出版行业重视版权信息的标准化管理。国家新闻出版广电总局于 2015 年制定发布了《版权信息核心元数据标准（CY/T 134-2015）》；2014 年发布《互联网电视数字版权管理技术规范（GY/T 277-2014）》；2007 年发布《中国广播影视数字版权管理元数据标准（GY/T 261-2012）》。同时，数字出版企业有比较完善的版权信息管理系统，能够帮助企业快速获取有效的版权信息、实现版权信息的共享。

从目前国内外研究和实践的情况看，数字资源的权益管理主要是版权管理。这主要是由于版权状况代表了资源内容的知识产权情况，任何对资源内容的使用和再利用必须按照版权规定、授权许可执行。然而另一方面，随着科技和信息环境的变化，资源再利用除了受到版权约束以外，还受到来自其它权利约束，产生新的权益问题。如数字资源的来源平台主张数据库权^[7]，可能会对资源的获取方式、频次等方面进行限制。又如随着资源开放获取、科研数据共享、科研方式的变革，对资源内容的数据挖掘，引发了各种有关权益问题的讨论^[8]。一方面，国际出版商对数据挖掘行为进行一定限制；另一方面，信息服务机构和科研群体目前出版商对文本和数据挖掘（TDM）施加的限制过于严苛，不利于研究和开发，因而缺乏正当性^[9]。因此，综合当前资源利用和服务中遇到的新问题，需要对数字资源的权益管理，从单纯的版权管理拓展到对其它类型权利的管理，全面梳理资源相关的各种权益状况以支撑图书馆服务。

2 不同类型的数字资源权益状况分析

本文选取了以下十种有代表性的数字资源类型，并对资源协议文本、使用许可和网站政策中涉及的关键权益信息进行了调研和梳理。

（1）电子期刊权益状况

电子期刊的出版模式包括商业订购和开放获取，资源内容层级涉及元数据、

全文、附加数据。电子期刊的权益状况主要取决于论文的版权、使用许可，以及所在平台、数据库的网站政策、元数据收割政策。

版权政策和论文内容再利用政策来说：对于商业订购模式，一般来说资源版权属于出版社，出版社通过商业订购合同对电子期刊论论文的使用进行全面的规划和控制，电子期刊的再利用需要遵循商业订购合同。对于开放获取模式，一般来说金色 OA 期刊的版权归属于作者，出版商享有专属发表权或商业利用权；论文内的图表版权归属作者；论文中所涉及的商标或其它第三方内容版权归原始创作者；对于版权不明的内容，一般情况下归出版商所有。论文再利用政策方面，OA 期刊采用开放共享协议，主要是针对单篇 OA 论文的使用行为做出的规定，这些行为包括阅读、下载、复制、传播、演绎、打印、索引、链接等。

对于电子期刊数据库的网站使用条款中，多数都会声明出版商对于网站（或平台）拥有版权，这其中包括网站的设计、代码和页面内容。网站版权受到所在国版权法和国际知识产权法的保护。网站内容再利用政策方面，该政策与网站版权政策密切相关。声明对网站内容拥有版权的出版商（或平台）通常都会对网站内容的复制、传播、商业利用加以限制。另外，出于网站安全的考虑，部分出版商针对机器抓取网站内容制定了明确的政策，有的允许期刊被机器采集，有的则明令禁止。

元数据收割政策方面，该政策多以提供开放接口的方式实现，实际体现的是出版商对于电子期刊批量获取方面的态度，如提供能下载到元数据，甚至是全文的 API 接口，表明出版社对于期刊论文批量获取的支持。

文本挖掘政策方面，该政策主要是针对批量期刊数据的获取，以及获取之后的重组、分析、挖掘而制定，表明出版商对于期刊批量再利用的态度。目前，绝大多数 OA 期刊并没有针对文本挖掘等批量利用期刊的政策或说明，仅有爱思唯尔（Elsevier）一家出版商提供文本挖掘的 API 接口，并说明遵循 CC 协议的内容可以进行文本挖掘；Wiley 提出遵循 CC 协议的 OA 论文也可以用于文本挖掘，但并未提供批理获取数据的接口。同时结合 OA 期刊“版权政策”“论文再利用政策”，批量利用 OA 期刊的署名方式，并没有在文本挖掘政策中进行规定。这也给图书馆对 OA 期刊的批量再利用带来一定问题。

（2）电子图书权益状况

电子图书的权益，一般分为整体图书的版权和再利用政策，以及章节内容的版权和再利用政策。图书的版权按其出版模式不同，版权持有方可能为作者，可能为编辑和作者共同所有，可能为出版商。商业订购的电子图书在版权和数据库再利用政策方面可参照电子期刊。对于 OA 图书来说，目前开放获取的电子图书相较于电子期刊来说仍是少数，有明确使用许可协议的开放图书则更少。如版权

在公有领域，则资源可以自由利用。除大型出版社以外，很多开放图书的网站或平台没有网站使用条款，对于网站内容的存储、批量获取等方面并没有规定。由于在各国数据库权受到保护，对网站的再利用也会存在一定风险。

（3）电子会议录权益状况

当前大部分会议录的出版等同于电子期刊的出版。电子会议录的版权情况一般比较复杂，可能的情况包括出版社对会议论文享有版权，出版社和作者共同享有版权，或者作者单独享有版权。美国工业和应用数学学会（Society for Industry and Applied Mathematics，SIAM）规定所有电子会议录的版权在出版商，任何未经许可的复制行为都被禁止；德国亚琛工业大学出版的 CEUR Workshop Proceedings 规定单篇会议论文版权在作者，仅限个人和学术使用，整本会议录的版权在出版商；国际光学工程学会（Society of Photo-Optical Instrumentation Engineers，SPIE）会议录版权在出版商，出版商的版权受美国和其它国家法律所适用的版权法、数据库保护；荷兰 IOS Press 出版的会议录，其包含的会议论文版权归属存在多种情况，部分论文版权由出版社和作者共有，而一部分论文的版权归属会议组织机构；世界卫生组织（World Health Organization，WHO）资源网首页标明版权归出版商，作品除另外声明，只能合理使用，WEDC 没有任何版权信息。

电子会议录的再利用政策较为严格，多数要求使用出于个人或者学术目的，未经许可的复制、传播、演绎不被允许。SIAM 规定任何未经许可的复制行为都被禁止；CEUR Workshop Proceedings 复制行为必须出于个人或者学术目的；SPIE 则更为严格，用户被禁止更改、编译、复制、售卖、再传播、出版和再出版文章的内容，或系统下载、传递给非订购者或者机构；IOS Press 规定其会议论文采用 CC BY-NC 协议，禁止商业使用；OpenProceedings.org、Elsevier Procedia 则采取更为严格的开放许可协议，即 CC BY-NC-ND，禁止对作品进行演绎和非商业使用。

元数据再利用政策采用 CC 0 协议，促进用户对开放会议资源的再利用，包括元数据的获取、存储和加工再利用。但有此规定的出版商只有一家。

网站使用政策方面，可以参照电子期刊数据库的使用政策。但大部分开放会议资源平台没有网站使用政策，只有 Elsevier 有比较详细的网站使用条款，包括网站版权、网站内容再利用、网站内容获取、存储、链接等。批量数据获取政策方面，目前大部分开放会议资源没有针对批量数据获取政策，只有一家开放获取出版商提供 OAI-PMH 标准元数据收割接口。

（4）电子课件权益状况

课件的版权政策、再利用政策对于课件资源的采集、存储和服务会产生一定影响。课件资源一般为开放获取资源，对于版权归属为教师或学生、使用 CC

BY-NC-SA 协议的资源，可以在非商业使用和正确标引署名的前提下自由地下载、存储以及从事对内容的编辑加工等再利用；而对于“保留所有权利”的作品，除阅读以外的再利用需要获得版权人的许可，因此会对开放课件资源的采集、存储和服务造成一定阻碍。其次，内容再利用政策中的具体规定对课件资源的具体再利用方式有一定影响。如美国麻省理工学院（Massachusetts Institute of Technology, MIT），要求对原作品或者演绎作品的再传播过程中，要同时标引机构名称、创作者姓名，并且需要保持和原作品同样的使用许可协议平台；此外，平台还推荐了正确的标引方式。这些规定是对开放课件再利用，如图书馆开展服务时应注意的问题。另一方面值得关注的是，对图书馆来说，不存在商业目的的再利用，仅可以收取集成服务中可能出现的少量成本费用，例如用户需要某些开放资源的纸本或者技术提供的费用，但必须告知资源本身是免费开放的。

由于开放课件缺少整个网站的使用条款，以及对于批量获取、再利用开放课件的权益政策，因此也会对资源的采集、存储、服务造成一定影响，用户不易辨别使用行为。当前调研的各开放课件权益政策与平台的权益政策保持一致，权益政策主要包括版权政策、内容再利用政策两类，平台特别强调内容的非商业使用。版权政策和内容再利用政策对于资源的采集、存储和服务都有所影响，主要在于服务方面。开放课件版权情况较复杂，保留所有权利的资源再利用受限较多，图书馆若希望提供基于开放课件内容的服务需提前与版权人联系。

（5）科技报告权益状况

科技报告的版权和使用许可限制相较于其它类型资源比较严格，不利于资源内容的自由和深度再利用。对于开放、免费的科技报告来说，其版权持有情况复杂，版权持有人包括机构、作者、第三方等等。并且此类科技报告都未采用 CC 许可协议，科技报告的复制、传播都必须以个人或学术为目的，必须提前获得版权人的书面许可。美国政府开放的科技报告中，有一部分是公有领域资源，用户可以自由利用。美国国家航空航天局（National Aeronautics and Space Administration, NASA）提供的技术报告服务如果对网站图片、音视频等版权没有特殊声明的，均可认为是公有领域自由使用。

开放科技报告版权的复杂性和严格的再利用政策，包括下载、传播、演绎等，不利于对开放会议资源的自由利用。NASA 明确指出，规定不能批量自动抓取网站内容的原因，是由于每份科技报告版权状况不同，因而批量使用可能会侵权。网站使用条款中，包含了对于资源内容下载方式和使用方式进行了说明。在调研的科技报告中，National Academies of Sciences, Engineering, and Medicine，规定 PDF 需要注册才可下载。在资源内容使用方面，三家机构均规定，下载科技报告不得任意自由使用，需征得版权人同意。同时 NASA 还规定，网站中的图片、音

视频等如果没有特别版权限制,可以自由使用,但对于全文图片不可以批量下载。

网站内容获取方面,一家机构对此做出严格限制,不得采用爬虫等自动获取方式批量下载科技报告内容,因为涉及不同的版权主题,可能面临侵权。

大部分科技报告并没有批量获取的有效途径。

(6) 学位论文权益状况分析

学位论文的版权和再利用政策一般比较严格。Proquest 等商业数据库对学位论文的下载、使用有着明确的限制。我中心的学位论文服务对资源版权、开放内容有着明确的权利规定,如只允许检索和前 16 页的下载。开放学位论文一般存储于机构知识库挂要想获得全文,需要遵循机构知识库整体的政策。

(7) 标准资源权益状况分析

标准类资源和数据库的权益特点是相关权益人众多。在国际、国外的标准化管理体制中,标准版权多为标准制定组织所有,当前标准资源的开放共享成为一种趋势。对于商业标准资源,一般都实行非常严格明确的版权保护。对于免费开放的标准,如 OECD,所有的标准都提供文本免费公开,并不以标准销售为主要收入来源,只需要在使用 OECD 标准内容时,注明内容来源和版权所有者为 OECD 即可。但总体来讲,在标准采用、使用的过程中,要严格遵守版权保护政策才会避免对标准版权构成侵权行为。

(8) 专利资源权益状况分析

专利类资源通常包括各类专利数据库和专利大数据分析平台,如 Innography、TDA、TI、DII、知识产权出版社专利平台、WIPS、ORBIT、Relecura 等商业类专利数据库和专利信息分析工具;也包括互联网免费的专利信息。专利资源受版权保护,一般对于再利用有明确严格的限制,如欧洲专利商标局(European Patent Office, EPO)对版权和网站内容的使用有严格的限制,所有官方出版物的版权属于 EPO,不允许机器抓取、任何数据挖掘、使用数据收集和分析的工具。

(9) 图表数据权益状况分析

附加图表等数据受版权法保护,在资源利用过程中需要判断是否涉及附加数据属于第三方的情况。

综上所述,不同类型数字资源在版权、再利用政策方面的宽严程度不一,但从服务和再利用的角度来说,能够提炼出具有共性的权益信息项,这些信息项对于服务产生影响。因此资源权益状况描述框架,建立保留各个资源共有、对服务产生影响的权益项。

3 数字资源权益状况描述框架的设计

在上述调研基础上,本文对各类型资源涉及的权益信息进行了合并,提出了

数字资源权益状况描述，包括总框架和针对不同类型资源的补充权益信息，用于对资源的权益状况描述。具体内容如下：

表 1 数字资源权益状况描述总体框架

一级信息项	二级描述项	三级描述值	四级描述值	选项属性
基本 信息	展现程度	元数据	(无)	必选
		全文数据	(无)	必选
		附加数据	(无)	必选
	来源渠道	付费购买	文本说明	必选
		协议赠送	文本说明	必选
		开放获取	文本说明	必选
		公开免费	文本说明	必选
		数字化	文本说明	必选
		自建资源	文本说明	必选
		其它	文本说明	必选
	版权归属	公有领域	(无)	必选
		作者	(无)	必选
		出版机构	(无)	必选
		不明确	(无)	必选
	发表年份	(无)	数值数据	必选
	版权联系信息	(无)	文本说明	可选
	授权许可方式	合同授权	文本说明	必选
		开放许可	①非常宽泛(如 Public Domain/CC 0/CC BY/CC BY-SA) ②一定约束(如 CC BY-NC/CC BY-NC-SA) ③较严格限制,不能再演绎(如 CC BY-ND) ④严格使用限制,不能商业利用、再演绎(如 CC BY-NC-ND)	必选
			无明确许可	版权人保留所有权利(All Rights Reserved) 未说明
	权限信息	版权状态	发表权	①授权 ②未授权
署名权			①授权 ②未授权	必选
修改权			①授权	必选

			②未授权	
		保护作品完整权	①授权 ②未授权	必选
		汇编权	①授权 ②未授权	必选
		改编权	①授权 ②未授权	必选
		翻译权	①授权 ②未授权	必选
		复制权	①授权 ②未授权	必选
		信息网络传播权	①授权 ②未授权	必选
		发行权	①授权 ②未授权	必选
	使用目的	允许商业使用	文本说明	可选
		非商业使用	文本说明	可选
	获取权限	在线阅读	①允许 ②禁止 ③未说明	可选
		本地单篇下载	①允许 ②禁止 ③未说明	可选
		开放接口	①有（如API、 OAI-PMH） ②无	可选
		机器爬虫	①允许 ②禁止 ③未说明	可选
	存储权限	（无）	①允许 ②禁止 ③未说明	可选
	文本挖掘	（无）	①允许 ②禁止 ③未说明	可选
	用户范围	院内用户	文本说明	必选
		院外用户	文本说明	必选
		特定用户	文本说明（IP段）	必选

封面图片针对图书；图表、数据集多为期刊论文、会议论文、会议PPT汇报展示文档；图书、学位论文、各类报告中涉及部分章节提取问题；一般在资源引用中要注明出处，其余要求根据资源提供方具体要求而定。

表2 数字资源权益状况描述补充内容

资源附件	封面图片	使用需注明来源与版权归属
		无需注明任何信息
		无明确说明
	图示、表格	使用需注明来源与版权归属
		无需注明任何信息
		无明确说明
	数据集	使用需注明来源与版权归属
		无需注明任何信息
		禁止使用数据集
		无明确说明
	部分章节抽取	使用需注明来源与版权归属
		无需注明任何信息
无明确说明		

4 数字资源权益状况描述框架的应用建议

针对当前图书馆资源建设与服务中涉及的对于“权益信息响应服务”的需求，本文认为应基于上文中提到的数字资源权益状况描述框架，建立“从资源输入到资源输出”的权益信息响应服务工作流，包括资源权益信息登记、权益信息管理、资源请求信息登记、资源利用与服务权益合规性审核等内容。

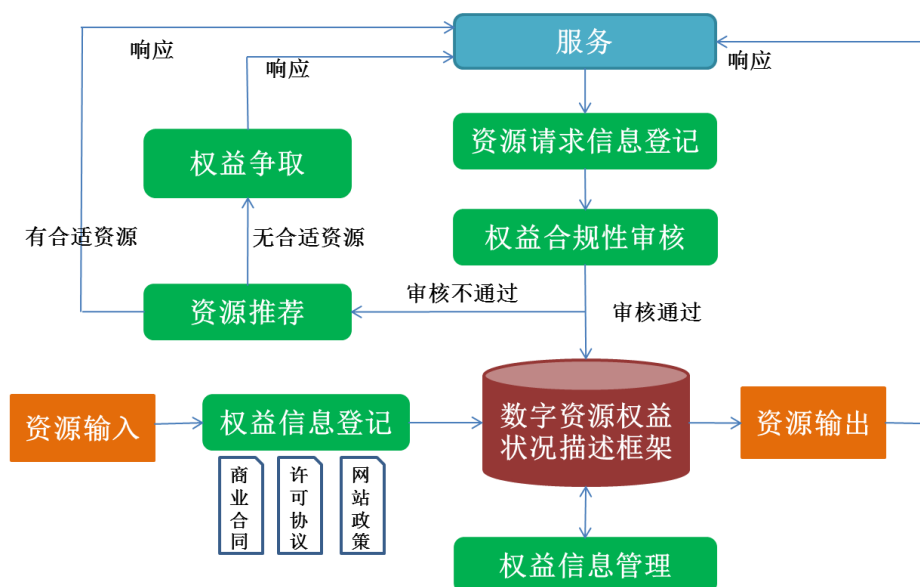


图1 数字资源权益状况描述框架响应服务的工作流

(1) 权益信息登记

按照本项目提出的知识资源权益描述框架，根据商业资源、开放资源和自建资源的合同、许可协议、网站政策，以及法律法规的规定选等，登记每个入库资源的权益信息。

(2) 权益信息管理

对登记入库的权益信息进行有序、结构化管理，包括：

- ① 按元数据、全文、附加数据等不同资源层级进行管理；
- ② 按版权属于作者、出版社、公有领域等进行管理；
- ③ 按使用权限，是否授予复制权、传播权等著作权 17 项权利进行管理；
- ④ 按获取权限，是否允许单篇本地下载，是否提供开放接口、是否允许机器抓取等进行批量获取进行管理；
- ⑤ 按存储权限，是否允许再集成进行本地化保存进行管理；
- ⑥ 按文本挖掘权限，是否允许文本挖掘、再利用进行管理；
- ⑦ 按使用目的，是否允许提供商业服务，或只能作为个人科研、教育使用进行管理。

(3) 资源请求信息登记

资源请求信息登记的目的，是为了在权益合规性审核中，将资源利用方式和目的与资源权益登记信息，即权益要求进行匹配。

需选择登记如下信息：

- ① 请求资源的基本信息，包括题名、唯一标识符、链接网址等；
- ② 请求内容：元数据、全文、附加数据
- ③ 请求方式：单篇下载、批量下载
- ④ 使用目的：商业、非商业
- ⑤ 服务类别：文献服务、情报分析、集成服务、数据挖掘服务
- ⑥ 使用范围：院内、院外、个人
- ⑦ 使用方式：发表、署名、修改、保护作品完整、汇编、改编、翻译、复制、信息网络传播、发行、本地保存、文本挖掘；
- ⑧ 资源请求责任人信息。

(4) 权益合规性审核

通过机器自动匹配加人工辅助的方式，将资源请求登记的信息同资源本身登记的权益信息进行比对。审核通过的资源，可以响应服务诉求、反馈资源；如果权益审核不通过，则通过大数据知识资源中心根据用户填写的资源需求和服务方式，为请求者推荐其它符合权益要求的资源；如果系统里没有其它合适的资源，则建议通过线下方式，针对审核不通过的权益项目，与该权益的相关方进行联系谈判、争取权益。

(5) 资源推荐

图书馆可以根据用户的服务需求，从权益的角度，获取到最符合用户需求的资源。如针对数据分析与挖掘利用服务，图书馆员应优先选择“授权许可”为开放许可，“版权状态”为公有领域，或者 CC 0/CC BY 等协议的资源；同时选择有

批量接口，或者允许机器爬取的资源。若用户已指定资源需求，或者所需要的资源中权益状况复杂，存在着服务需求与资源权益状况冲突情况，则触发“知识资源权益争取与合作机制”。

（6）权益争取

未得到满足的资源需求和服务需求，可以通过知识资源权益状况描述框架明确，哪些使用方式超出了资源授权、如何与版权人取得联系，获取需要向版权人或出版社争取的权益清单。图书馆员在获得上述信息后，应主动与相关权益方开展合作谈判，规避可能权益风险、最大限度保障服务顺利开展。

5 结语

做好图书馆数字资源权益管理是保障图书馆业务顺利开展的重要内容。图书馆应从机构知识产权战略的角度，对于所获取的商业资源、开放资源以及自建资源等做好权益信息的登记、所获得的权利项管理，并进一步明确促进信息的传播和利用过程中图书馆能争取的权利。本文制订的数字资源权益状况描述框架希望能为相关研究和实践提供一些参考和借鉴。

参考文献

- [1] CAROLI C, RRAPI E, SCIPIONE G, et al. ARROW: Accessible registries of rights information and orphan works towards Europeana[J]. D-Lib Magazine, 2012, 18 (1): .
- [2] 陈湛绮. 版权清理: 基于 MOOC 的图书馆版权管理新功能[J]. 图书馆学刊, 2015(7): 18-20.
- [3] 李华伟, 富平. 试论国家图书馆数字资源建设与服务业务流程中的版权管理[C]//国家图书馆科学讨论会, 2007.
- [4] 吕淑萍. 图书馆数字资源版权管理实践与案例[M]. 北京: 国家图书馆出版社, 2013.
- [5] 谢强, 石慧. 图书馆移动服务版权管理平台设计与实现[J]. 图书馆论坛, 2013, 33 (2): 102-106.
- [6] 章岑, 刘静羽, 黄金霞. 开放获取期刊集成服务权益问题研究[J]. 图书情报工作, 2016, 60 (10): 33-39, 94.
- [7] 刘祯娜. 《欧盟数据库指令》研究及对我国图书馆数据库法律保护的影响[J]. 新世纪图书馆, 2011 (11): 89-91.
- [8] 徐轩, 孙益武. 英国数据挖掘著作权例外条款研究及其启示[J]. 图书馆建设, 2015 (9): 10-14.
- [9] 陈笑. 论 Robots 协议下搜索引擎数据挖掘行为的法律责任[D]. 北京: 北京邮电大学, 2014.

作者简介

刘静羽, 女, 1989 年生, 学位, 职称, 研究方向: xxxxxxxx, E-mail: xxxxxx。

英文题名
英文作者
(英文机构)

英文摘要：
英文关键词：