

数字图书馆知识产权 保护研究进展

SHUZI TUSHUGUAN ZHISHI CHANQUAN
BAOHU YANJIU JINZHAN

 《图书情报工作》杂志社 编

数字图书馆知识产权 保护研究进展

《图书情报工作》杂志社 编

海洋出版社

2016 年 · 北京

图书在版编目 (CIP) 数据

数字图书馆知识产权保护研究进展/《图书情报工作》杂志社编. —北京:
海洋出版社, 2016. 3

(名家视点. 第7辑)

ISBN 978-7-5027-9354-8

I. ①数… II. ①图… III. ①数字图书馆-知识产权保护-研究进展
IV. ①D913.04

中国版本图书馆 CIP 数据核字 (2015) 第 308496 号

责任编辑: 杨海萍 张 欣

责任印制: 赵麟芬

海洋出版社 出版发行

<http://www.oceanpress.com.cn>

北京市海淀区大慧寺路8号 邮编: 100081

北京朝阳印刷厂有限责任公司印刷 新华书店北京发行所经销

2016年3月第1版 2016年3月第1次印刷

开本: 787 mm × 1092 mm 1/16 印张: 18.25

字数: 318千字 定价: 48.00元

发行部: 62132549 邮购部: 68038093 总编室: 62114335

海洋版图书印、装错误可随时退换

《名家视点丛书》编委会

主 任：初景利

委 员：易 飞 杜杏叶 徐 健 王传清
王善军 刘远颖 魏 蕊 胡 芳
袁贺菊 王 瑜 邹中才 贾 茹
刘 超

序

2016年新年伊始，由《图书情报工作》杂志社策划编辑的《名家视点：图书馆学情报学档案学理论与实践系列丛书》第7辑共5本，已由海洋出版社出版。这一辑丛书是编辑从近年来发表的论文中精选出来的。这是《图书情报工作》杂志社与海洋出版社联袂奉献给中国图书情报界的新年志喜，也是为国家“十三五”规划开局之年给图书情报从业人员的一份礼物。

在这份礼单中，首先是《图书馆发展战略规划与趋势》。在这本书中，我们收录了26篇文章，分专题篇、国外篇、国内篇三个部分，集中展示了国内学术界对图书馆发展战略规划与趋势的研究成果。创新发展，规划先行。面向“十三五”，图书馆需要在把握趋势、把握大势的基础上，确立新思维、制订新战略、采取新行动。“十三五”规划制订得好坏，直接影响每个图书馆今后5年的发展，而今后5年对图书馆的转型变革是至关重要的，是挑战，也是机遇。这一组文章基本上能展现国内外图书馆发展趋势和战略规划的特点和要点，相信对每个图书馆及管理者和图书馆员都具有重要的参考和借鉴价值，应成为制订“十三五”规划和指导“十三五”期间图书馆工作的重要案头书。

其次是《新环境下图书馆用户信息行为》，共收录26篇重要文章，分专题篇、网络篇、服务篇和综述篇4部分。用户信息行为是图书馆学情报学最重要的研究对象。图书情报服务做得好不好，往往是由图书情报机构对其服务对象（用户）的信息需求和信息行为的认知和分析深度所决定的。在当前变化的信息环境下，我们对用户的信息需求及其行为的跟踪、揭示和研究是非常不够的，这不仅

是由用户信息行为的复杂性所决定的，也是因为我们为用户行为的研究仍缺乏前瞻性的理论、科学的方法和有效的技术。本书所收录的文章将为我们进一步的研究提供新的起点、新的视角和新的结论，有助于我们对用户信息行为提供完整和深入的认识。

在今天图书情报机构提供的信息服务中，专利是不可忽视的。专利被认为是创新性研究或应用成果的代表，代表的是科技创新能力，对企业和各类机构而言是十分核心的信息资源和创新支撑。在《专利情报研究与服务》一书中，我们收录了28篇文章，分专题篇、方法篇、应用篇、评价篇4个部分，展现了国内专利领域的专家学者在专利的引文、工具、挖掘、服务、评价等多个方面有代表性的研究成果，表明国内图书情报界在应用专利推动国家的创新驱动发展战略中所开展的卓有成效的服务工作。总体而言，国内图书情报机构在对专利的重视程度上不够，在应用专利推动科技创新的实践力度上不够，在将专利信息资源转化为现实生产力的实际效果上不够。期待这些文章能对解决这些问题产生一定的推动作用。

网络舆情的研究随着新媒体环境的出现而愈发引起包括政府和相关机构的高度重视，也吸引了广大研究人员的积极参与。在名家视点第5辑中推出的《新媒体环境下的网络舆情研究和传播》一书受到读者好评，现已售罄。故这一辑将此书再版，增加了一些最新的稿件，使该书跟上新的形势。在不少图情机构，网络舆情的监控与分析，已经成为一项重要的情报研究和咨询服务。

最后一本书是《数字图书馆知识产权保护研究进展》。随着数字图书馆建设与发展，与数字图书馆相关的知识产权问题也显得愈加突出，往往是著作权法等相关法律未曾涉及的新的问题。知识产权问题能否解决好，关系到作者知识产权保护与数字图书馆可持续发展的平衡问题。二者不应是矛盾的，而是数字图书馆发展中必须直面、解决的问题。本书收录26篇文章，分策略篇、实践篇、综合篇，展示了我刊近年来发表的重要的相关研究成果，体现了作者们

在有关数字图书馆知识产权问题的认知、实践和策略，有助于启发我们更深入的思考，提出更加符合法理和现实环境的解决对策。

“十三五”已经到来，图书情报界需要重新定位，前瞻谋划，大胆探索，砥砺前行。就图书情报机构的转型而言，这是一个非常关键的5年。如何做好规划，做好布局，寻求新的突破，重塑图书情报新的职业形象，赢得应有的职业尊严和专业地位，不仅关系到这5年的发展，而且直接影响未来10年或者更长时间图书情报机构的生存。我们既要有危机感和忧患意识，也要提振信心，抓住机遇，看到未来发展的前景。“图书馆，不要自甘寂寞”（《人民日报》2015年12月22日第12版语），代表的是全社会对图书馆的期许，也是对图书馆人的鞭策。图书馆若想不被边缘化，拯救自己的只有图书馆员自己。

初景利

《图书情报工作》杂志社社长、主编，教授，博士生导师

目 次

策 略 篇

数据挖掘版权政策构建研究:研究型图书馆的立场	(3)
国外电子资源授权许可协议法律适用问题研究	(15)
图书馆合理使用的比较研究	
——以海峡两岸著作权法为对象	(24)
论我国 IR 存取模式及其版权政策体系之完善	(39)
我国可适用于图书馆的著作权例外规定及其适用障碍剖析	(55)
数字图书馆新媒体阅读服务著作权侵权风险分析与控制	(63)
网络信息资源共享的著作权利益补偿机制研究	(71)
个体读者借阅权保障策略分析	(84)
图书馆著作权作品获取权的限制与保障	(93)

实 践 篇

美国大学图书馆版权服务调查及启示	(103)
大学图书馆参与 MOOC 版权服务的实践及启示	
——以杜克大学图书馆为例	(113)
私立网络著作权授权规则对图书馆的实践价值及影响	(120)
一种用于版权保护与盗版追踪的文档图像多功能水印方案	(131)
图书馆行业 OpenAPI 利用的权限控制	(143)
新的国际版权立法对图书馆为视障者服务的影响	
——基于《马拉喀什条约》的思考	(152)

图书馆印刷品阅读障碍人士版权例外研究

——《马拉喀什条约》述评及对中国图书馆界的建议

..... (161)

图书馆孤儿作品利用的法定许可制度研究

——以著作权法第三次修改为背景 (172)

孤儿作品的版权困境及解决路径 (182)

综 合 篇

云计算环境下的图书馆与著作权保护 (192)

基于可拓理论的网络信息资源著作权侵权风险评估 (200)

三网融合背景下的图书馆功能扩张与著作权诉求转移研究 (218)

图书馆服务的新定位与分享著作权利益的新内涵 (227)

学术信息资源数据库商“独家授权策略”的垄断性分析 (235)

复制权例外对图书馆数字资源长期保存的影响剖析 (243)

国防工业合作创新中的国防知识产权分析 (251)

我国知识产权研究领域的共词网络分析 (263)

策 略 篇

数据挖掘版权政策构建研究： 研究型图书馆的立场^{*}

数据挖掘（data mining）是利用计算机程序进行数据处理，提取和组织潜在在有用信息的过程^[1]，是大数据时代分析海量数据的重要研究方法。研究表明，获取大量文献的许可成为研究者开展数据挖掘的重要成本之一^[2]，过高的许可成本和落后的版权制度制约着数据挖掘在科学研究中的发展。2011年5月，英国学者伊恩·哈格里夫斯（I. Hargreaves）提出^[3]：法律不应限制新技术和工具在研究尤其是非商业性研究中的使用，而数据挖掘是一种具有代表性的新型研究工具，建议英国政府应当革新其版权法框架为非商业性研究中的数据挖掘提供版权例外。2014年6月，英国在其新修订的著作权法中引入了非商业性研究中的数据挖掘例外。在世界知识产权组织（WIPO）的舞台上，关于为数据挖掘提供版权例外的讨论亦非常热烈。在实践层面，出版商爱思唯尔（Elsevier）等更新其数据挖掘政策，并通过与研究型图书馆签订的采购合同对图书馆及其用户产生影响。我国研究型图书馆是高等院校和研究机构的文献信息中心，作为内容的创造者和使用者，应明确在数据挖掘相关问题上的立场，厘清数据挖掘过程中数据获取、复制及其结果利用等环节涉及的法律问题，借鉴相关国家的做法，制定完善的数据挖掘版权政策。

1 研究型图书馆构建数据挖掘版权政策的需求分析

高校和研究机构是学术内容的创造者，也是文献和数据的使用者。研究型图书馆作为高校和研究机构的文献信息中心，承担着保存文献、传承文明、支持教学科研的职责。综合考虑研究型图书馆在高校和研究机构中的角色，有助于厘清研究型图书馆在数据挖掘问题上的立场。

^{*} 本文系教育部人文社会科学研究青年基金项目“基于《马拉喀什条约》视角的图书馆无障碍服务版权制度研究”（项目编号：14YJC870023）和上海图书馆学会一般项目“研究型图书馆电子图书阅读推广载体研究”（项目编号：14STX008C）研究成果之一。

1.1 高校和科研机构的数据挖掘版权政策需求分析

研究型图书馆是高校或科研机构的一部分,构建其数据挖掘版权政策需充分考虑高校及科研机构所处的版权政策环境。从内容创造者和内容使用者的角度,高校和科研机构完善其数据挖掘版权政策有其必要性。

作为科学内容的重要生产方,高校和科研机构每年产出大量的学术科研成果:仅2013年,我国各类高校共发表学术论文797 104篇,出版科技著作12 060部,申请各类专利106 714件,专利授权数达68 971件^[4]。为了充分保障科研工作者的精神权益和经济利益,大学和科研机构应完善数字环境下的版权政策,遏制盗版和侵权,鼓励科研创新。

数据和科学文献是开展科学研究的基础。从内容使用者的角度,高校和研究机构日渐关注本机构研究人员是否可以在科学研究中很好地利用数据开展数据挖掘,尤其是在医学、化学、计算机等学科领域深入应用数据挖掘。

1.2 研究型图书馆的数据挖掘版权政策需求分析

大数据环境下,研究型图书馆需完善其数据挖掘版权政策以应对其面临的诸多新挑战。研究型图书馆作为机构的文献信息中心,收藏和订购了大量的纸本和数字资源,为高校和科研机构提供重要的文献保障。尤其是许多研究型图书馆承担着运营本校机构知识库或收集和管理本机构科研产出的职责;例如中国科学院文献情报中心运营着中国科学院机构知识库^[5],复旦大学图书馆推出了复旦大学社会科学数据平台^[6]。

一方面,研究型图书馆数据挖掘版权政策应当注重知识产权保护。研究型图书馆及其所属的高校和科研机构承担着许多科学研究任务以及科研支持工作,产生大量的数据和成果,其中许多内容和文献具有自主知识产权。另一方面,研究型图书馆数据挖掘版权政策应当注重公共利益的考量。科学研究是社会创造知识、支持创新和促进发展的重要手段,特别是各国投入大量公共财政资助科学研究并产生众多的科研成果。这些知识资源在全社会的开放获取将促进知识的传播利用,并推进科学事业的全球化进程,让知识普惠社会发展。这要求研究型图书馆有效平衡版权权利人和公众的利益,积极推进开放获取,在保护知识产权的前提下有效促进知识传播和信息共享。

因此,作为文献的收藏和传播者,作为机构库和数据平台的运营方,研究型图书馆应当明确相关科学数据或多类型作品的版权归属和再利用政策,明确其用户如何利用大量的数据库资源和开放资源,如何解决数据挖掘涉及的版权问题,并重点明确基于非商业性研究目的数据挖掘涉及的主要法律问题和制度障碍,构建完善的数据挖掘政策,以支持研究人员更好地利用图书

馆藏资源和开放资源进行科研创新。

2 版权法视角下的数据挖掘法律问题分析

在数据挖掘过程中,获取数据、分析数据和处理数据挖掘结果等步骤涉及诸多版权问题。

2.1 数据获取(复制)中涉及的版权问题

不同于图书馆传统收藏的纸本文献,数据挖掘中涉及大量不同类型的电子数据的获取和复制。这些数据既包括对印刷型文献的数字化后产生的数字内容,也包括大量以数字形式存在的内容和数据,如大量的数字文献、科学数据、用户信息、计算机后台数据等。

研究型图书馆及其用户开展数据挖掘时其内容来源多样,数据量大。例如,图书馆对读者信息数据具有所有权,而大量的数据库文献与数据则通过订购获得合法获取权,另有许多数据来自于网络。所以,在获取不同数据源时,首先需要明确其权利归属或授权类型,避免侵权风险。

2.2 数据挖掘分析方法所涉及的版权问题

在数据挖掘过程中,需要对源数据进行分析 and 重组,当使用某种程序或方法进行数据分析时,可能复制源数据库的完整内容或核心内容。如果数据库所有人禁止或限制特定种类的挖掘程序或方法,使用者是否有权对源数据进行某种形式的破解?

在对所订购数据库等非自有知识产权内容进行数据挖掘时,研究型图书馆可能对源内容实施实质性复制或数据库设置的技术措施进行破解;这些行为存在一定侵犯版权的风险。

另外,各国十分重视数据挖掘专利的申请与保护,仅美国就有数据挖掘方面的专利 13 841 项^[7],我国至少有数据挖掘相关发明专利 395 项^[8]。因此,在开展数据挖掘之前,应当进行专利检索和调查,确保算法和实施过程不侵犯相关专利权。

2.3 数据挖掘结果处理所涉及的版权问题

在数据挖掘中,研究者往往设定数据的阈值或设定对象控制的级别,从而筛选出符合需求的数据并进行分析处理。在这一过程中,阈值的设置非常关键,直接影响结果的呈现。如果阈值设置不当,可能会因无法预见、识别、控制数据挖掘对象或使用范围是否经过合法授权,导致挖掘出的数据信息存在外泄的可能,从而损害用户或他人利益^[9]。因此,数据挖掘中,应该关注结果可能对版权权利人可能存在的风险,设置合理的结果使用方式和范围,

禁止对源数据和挖掘结果进行滥用。

3 数据挖掘版权政策构建模式研究

研究型图书馆构建数据挖掘版权政策时,与宏观的国家政策及微观的图书馆业务政策息息相关。在国家层面,版权法等知识产权法律法规对研究型图书馆的科研支持工作有着重要的规范作用。在图书馆业务中,随着数字时代的研究型图书馆与数据库出版商等内容提供方的合作逐渐深入,出版商的数据挖掘政策也成为影响研究型图书馆开展数据挖掘版权政策构建的重要因素。以下从国家和数据库商两个层面,讨论数据挖掘版权政策的两种模式。

3.1 意思自治:出版商倡导通过许可合同规范数据挖掘

意思自治原则,或称“自愿原则”,在我国《民法通则》第4条和《合同法》第4条中得到确认,是指民事主体不仅能够按照自己的意思表示来实施对其具有法律约束力的任何民事法律行为,而且其所实施的民事法律行为均按照他们的意思表示来产生法律效力,即使是在涉外版权转让或许可合同中,当事人亦可以约定具体适用哪一个国家的法律,这是当事人意思自治的体现^[10]。

在数据挖掘相关的版权政策方面,出版商主张遵循意思自治原则,通过许可合同建立有效的数据挖掘许可机制。近年来,国际科技与医学出版商协会(International Association of Scientific, Technical & Medical Publishers,简称STM)、欧洲出版商协会(European Publishers Council,简称EPC)以及出版商Elsevier等相继发布或更新数据挖掘政策,表明其在数据挖掘尤其是非商业性数据挖掘上的立场。这些政策将直接影响研究型图书馆及其用户对所订购资源的利用。

3.1.1 主张通过许可授权数据挖掘

为了构建一个充满活力的媒体和出版生态系统,STM、EPC等主张:许可是解决技术和版权问题的核心^[11],许可机制可以保证数据挖掘的可靠性、有效性和高效性^[12]。出版商允许用户对所订购的内容进行挖掘,并通过提高内容可挖掘性、开发数据挖掘平台和工具等提供多样化的数据挖掘解决方案^[13]。

另一方面,出版商认为,访问协议和许可协议等将全面覆盖商业性研究和非商业性研究,并囊括不适用版权例外的诸多情况。因此,即使存在数据挖掘的版权例外的情况,亦可以通过许可合同对商业性数据挖掘的权利行使、费用等进行全面的规范。许可模式依然是市场中解决数据挖掘这一问题的最

有效途径。

3.1.2 推出数据挖掘的示范许可条款

2012年3月15日,STM发布了一个适用于多种情况的数据挖掘的许可示范条款^[14],通过详细的许可条款规范订购用户的非商业性数据挖掘行为,同时确保数据库商或出版商的利益。STM主张将非商业性文本与数据挖掘的权利纳入许可示范合同,并对“文本与数据挖掘输出(Text and Data Mining Output)”进行定义,规范订阅者对订购的内容及出版商网站进行合法下载或数据挖掘时的权利,同时规定订购者的限制性行为。示范条款还包含数据库的系统安全、费用、违约、合同终止等相关内容。

2014年2月31日,出版商Elsevier更新数据挖掘政策^[15],允许科研人员基于研究目的对已订购的内容进行数据挖掘,并在订购合同中更新了相关条款。政策对数据挖掘的方式(通过Elsevier提供的应用程序接口(API)以连续并自动的方式提取、索引和/或处理订阅的内容)、数据挖掘结果的利用方式(数据挖掘结果应以知识共享(Creative Commons)中的“署名+非商业性使用(CC-BY-NC)”模式)等作出详细规定。同时对订阅产品的使用作出限制,明确:除非Elsevier明确许可,订阅者及其授权用户不得使用任何机器人、蜘蛛、爬虫或其他自动下载程序、算法或装置持续自动搜索、获得、摘录、深链接、索引订阅产品或扰乱订阅产品的运行,或实质性、系统性地复制、保留或再分发订阅产品。

3.2 法定例外:数据挖掘的版权例外国家立法

3.2.1 国家立法实践:以英国为例

2014年6月,英国修订其版权法,在第29条(研究、个人学习和基于非商业目的文本与数据分析)^[16]新增关于数据挖掘的版权例外条款,其基本内容为:有权获取作品的主体开展的基于非商业性目的的数据分析之复制行为,不构成版权侵权。同时,设置了限制条件,明确在转让、不符合目的要件、交易等情况下,对作品的复制构成版权侵权。英国的非商业性数据挖掘的版权例外制度同时明确与合同的关系,规定“任何阻止或限制本款中所指的不侵犯版权之复制的合同条款不具有执行力”,为英国研究型图书馆开展非商业性数据挖掘版权例外提供了法律基础。这一例外条款是迎合新技术和工具在研究中的使用的需求、积极改革版权体系的表现,也将为英国的科研创新和经济竞争力的发展注入新活力。

3.2.2 国际立法协调:以WIPO相关议案为例

年来,在知识产权组织平台上,数据挖掘纳入了“图书馆和档案馆

相关的例外与限制”和“关于教育、教学和研究机构以及其他残疾人的例外与限制”两大议题的讨论范畴，国际图书馆协会联合会（IFLA）、非洲集团代表等呼吁为数据挖掘的版权例外设立统一的国际立法标准，推动各国对该议题的立法。

IFLA 认为，为了充分利用海量数据并共享数据挖掘结果，确保图书馆在数字环境下提高信息服务能力，充分履行职责，应为图书馆数据挖掘提供版权例外或限制；而许可合同并不是解决数据挖掘相关问题的合理途径^[17]，在其发布的《关于图书馆和档案馆的版权例外与限制的条约建议稿》4.4 版本^[18]中，就数据挖掘作出了专门规定（第 20 条）。

而非洲集团建议将数据挖掘纳入“研究”的范畴，指出“基于非营利科研目的，通过搜索引擎、自动化知识发现工具或其他任何迄今为止了解或今后发现的数字手段，对任何合法获取的受版权保护的作品进行复制和再使用，不对版权及相关专有权构成侵权，但应在合理的可行范围内标明出处”^[19]。这一条款建议为数据挖掘等新技术在非商业性科学研究中的运用提供了版权例外，体现版权法框架应适应新技术发展的要求，同时对标明出处、技术保护措施等作出规定，具备一定的开放性和前瞻性。

3.3 数据挖掘版权政策构建模式的比较分析

3.3.1 法定例外模式的优势及其实施

以英国的立法实践为例，英国关于非商业性研究的数据挖掘的版权例外制度有效平衡了权利人和公众的利益，一方面明确了有权获取的主体有权对数据进行基于非商业性研究目的的数据挖掘，而不能限制其使用何种计算机程序、方法或算法，从而简化了科研人员从事非商业性数据挖掘的许可成本，将有效提高科研的效率。同时通过对数据挖掘行为的充分限制，保证例外规则在“三步测试法”的合理框架内运行，确保权利人的正当利益不受侵害。这一模式的制度设计充分考虑新技术和工具在科研中的需求，成为积极革新版权框架的例证。

在例外条款的实施方面，英国数据挖掘版权例外制度特别规定了与合同的关系，指出内容提供者不应不合理地限制研究者对其合法获取的文本和材料进行复制以便开展数据挖掘的权利，明确为了维护自身利益而规避例外条款的合同应视为无效，为例外条款的有效实施奠定基础。但在实际中，不同利益方对例外条款的解读与实施的差异等原因导致版权例外制度的完全实施存在一定困难，研究型图书馆往往处于相对被动的劣势地位。

例如，Elsevier 更新的数据挖掘政策中规定必须通过其提供的 API 获取已