

· 论坛 / PERSPECTIVE ·

化学专业数据库知识产权保护的探索

徐挺军, 陈维明

中国科学院上海有机化学研究所, 上海 200032

摘要: 为了探讨科学数据库知识产权保护与数据共享问题, 从化学专业数据库角度出发, 研究科学数据库知识产权保护所需的条件, 即“额头出汗”原则和独创性原则。结合自身特色阐述了化学专业数据库为了让数据库得到版权保护, 并能创造更好的知识产权保护环境, 而在知识产权保护方面的采取的应对策略和措施。介绍了化学专业数据库在转变服务模式、增加数据创新等方面的探索, 尝试解决科学数据库知识产权保护和数据共享之间的矛盾, 以更好的适应知识产权时代。

关键词: 化学专业数据库; 数据库版权; 知识产权; 科学数据库; 数据共享

doi: 10.11871/j.issn.1674-9480.2018.02.002

Exploration of Chemical Database in Intellectual Property Protection

Xu Tingjun, Chen Weiming

Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai 200032, China

Abstract: In order to explore the problem of intellectual property protection and data sharing in scientific databases, from the perspective of chemical database, the conditions of intellectual property protection in scientific databases had been studied, those are “the sweat of the brow” principle and originality principle. Combined with its own characteristics, this paper expounded the strategies and measures taken to protect the intellectual property of the chemical database, and to create a better environment for intellectual property protection. And introduced the exploration of chemical database in transforming service mode and increasing data innovation, try to solve the contradiction between intellectual property protection and data sharing in scientific databases, so as to better adapt to the era of intellectual property.

Keywords: chemical database; database copyright; intellectual property; scientific databases; data sharing

基金项目: 中国科学院信息化专项科学大数据工程 (XXH135); 上海市化学化工数据共享服务平台 (18DZ2294000)

引言

“十一五”至“十二五”规划期间,科学数据库根据学科特色对数据库资源进行了整合,并且创造性的发展了资源共享网络服务模式,已经建立起一套资源充足、标准规范、服务完善的信息共享系统,成为国内科研领域基础科学数据的标杆^[1]。化学专业数据库作为科学数据库项目最早参与的数据集之一,建立了以基础化学数据为核心的专业化学数据库群,逐步形成了具有自己特色的数据资源体系和服务能力,成为国内最主要的公益性化学数据服务中心。现阶段科学数据的共享方案已基本完善,数据共享平台的建设也初具规模,科学界对科学数据共享的理念已经逐步达成共识。化学专业数据库通过化合物唯一登录号实现了跨平台的化学数据整合^[2],同时化学专业数据库网站也通过数据服务平台进行数据共享服务。然而随着时代的发展,这种形式的数据共享,开始慢慢显现各种问题:一些比较重要的数据库数据不断遭到恶意下载;数据库服务网站遭受恶意攻击,从而大大折损化学专业数据库整体的服务质量;一些中小企业希望拿到某些完整的数据信息,然而数据库所有者难以评估这些数据的交易价值;而数据库所有者担心版权、信息安全等问题,无法完全开放下载数据;数据库后期的维护更新耗时耗力,单单以申请科学院信息化项目的形式无法维系后续的工作。这些问题究其根本还是科学数据库知识产权保护和数据共享之间对立统一的矛盾关系导致的:没有知识产权保护就没有数据创新的动力,但是过度的限制和保护又会妨碍社会公众对科学数据的合理使用^[3]。因此,无论从保护化学专业数据库本身还是其未来发展来考虑,探讨数据库知识产权及其相关问题是十分必要的。

针对科学数据库共享的现状,围绕科学数据的共享授权与应用,学者提出了科学数据知识产权保护和共享平衡的数据工作政策。通过构建用于授权控制的科学数据门户,达到分散型科学数据共享设施的统一用户授权与管理,从而实现更加安全有效的科学数据、元数据共享服务^[4]。数据出版近年来成为出版界和数据共享界共同提出和积极探索的新概念,有望从

机制上改变目前科学数据库知识产权保护和数据共享之间存在的根本问题。科学数据出版是深化数据共享的重要手段,既能激励数据生产者进一步开发维护更新和共享数据,又能很好地保护数据的知识产权^[5]。

中国科学院科学数据库已进入“十三五”阶段,化学专业数据库在该阶段的总体目标为显著提高数据资源共享水平和数据整合能力,包括进一步提升数据量和扩大数据覆盖范围,建立建全新的数据应用和特色数据库的开发。在此背景下,如何把握数据库的知识产权问题成为化学专业数据库未来发展的关键。

1 科学数据库知识产权保护的依据和条件

《世界知识产权组织版权条约》规定:“数据或其他资料汇编,不论用任何形式,只要由于其内容的选择或编排构成智力创作,其本身即受保护。”可见数据库作为信息收集、加工处理、以及存储的计算机形式载体,其相关数据信息受到知识产权法律保护这一点,已得到广泛共识。我国现行《著作权法》并没有对数据库版权问题进行特别的规定,但在其第14条(汇编作品的著作权归属)中规定:“汇编若干作品、作品的片段或者不构成作品的数据或者其他材料,对其内容的选择或者编排体现独创性的作品,为汇编作品,其著作权由汇编人享有,但行使著作权时,不得侵犯原作品的著作权。”而《伯尼尔公约中》对数据库定义有如下描述:“所有的信息(数据、事实等)的编纂物,不论其是以印刷式、计算机存储单元形式还是其他形式存在,都应视为数据库。”根据以上内容,可以认为科学数据库在我国作为一种汇编作品,受到版权保护。

根据以上法律法规条文,数据库可以受到版权法保护,但不是所有数据库或者数据产品都有知识产权,这取决于该数据库或者数据产品是否能称之为“作品”。而根据知识产权保护法相关解释,某事物成为“作品”必须有两个前提:即“额头出汗”原则和独创性原则。“额头出汗”原则认为,作者在完成作品(资料收集、选摘、编排、创作等)过程中必须投入一定的时间、精力和劳动,其作品才能作为受版权

法保护的作品。而一些简单的数据库, 仅仅复制、链接或者汇集的某些数据, 并不能成为作品受版权法保护。科学数据库通过几十年的持续建设, 其数据信息累积量已达 655TB。“十二五”期间, 共发表科学数据库相关论文 751 篇, 申请软件著作权 55 项、专利 30 项^[6]。其数据内容是各个研究机构在长期的科学研究实践中, 通过观测、考察、试验、计算等多种途径产生和累积的大量具有重要科学价值和使用意义的科学数据和资料。科学数据库的研制和建设过程是知识积累、加工、利用和传播的过程, 显然科学数据库符合“额头出汗”原则。

化学专业数据库的独创性: 化学专业数据库在数据资源建设方面, 对化学领域的数据库资源进行了整合, 自建数据库达到了 28 个, 总记录数超过了 907

万余条, 形成了覆盖化合物结构、化合物性质、研究文献等各种类型的专业化学数据集合, 化学专业数据库数据集资源目录见图 1 所示。而其数据共享服务方面, 化学专业数据库在近三年来, 为 8 千余家单位的 2 万多用户提供了 132 万次检索服务。化学专业数据集在其规模、内容、形式上, 任何其他公益性化学数据库都无法比拟, 为国内以至世界独有。

2 化学专业数据库知识产权保护的应对策略和措施

2.1 确保合理使用

《著作权法》在保护作品的著作权人和传播者利益的同时, 还兼顾了社会公共利益, 防止权利被滥



图1 化学专业数据库数据资源目录

Fig. 1 Directory of chemistry professional database data resource

用, 阻碍和束缚科学技术的进步和文化的繁荣。在特定条件下, 法律允许他人自由使用享有著作权作品而不必正的著作权人的同意, 也不必向著作权人支付报酬的制度。其限制范围包括: 个人学习研究、适当的转载引用以及用于课堂教学或科学研究等等。

建设科学数据库的出发点和立足点就是服务于国民社会经济发展, 因此, 化学专业数据库在探索版权保护的同时, 必须首先要确保数据库信息能被有效合理的使用。结合数据库这个特殊“作品”, 以及化学专业数据库自身特性, 版权法中的合理使用在这里可以作为两个方面的应对策略和措施: 一方面化学专业数据库加强了自身的数据共享服务建设, 确保广大普通的科研用户, 能够方便快捷地访问数据库页面, 简单高效地获取数据和相关信息; 化学专业数据库还加强了数据安全策略, 做好灾备镜像, 保证网页服务的时间持续性稳定性的同时, 还采取各种技术措施, 防止遭到恶意攻击破坏而导致的数据服务中断; 数据库开发人员也不断维护和更新数据内容, 保证数据质量以及数据价值; 另外一方面, 化学专业数据库加强了推广和宣传, 提升在学校师生、科研单位研究人员、中小企业用户以及普通民众中的影响力, 积极参与公众科普教育活动, 扩大“合理使用”的范围。

2.2 增强法律保护意识

从法律层面来说, 科学数据库体系作为一个整体, 适用于版权法相关规定受到法律保护。但是根据科学数据库的特点, 各个学科、领域的数据库物理上相对独立, 虽然同在科学数据库体系这个大平台进行数据共享和服务, 但是数据库的开发和其运维人员分散在各个研究所。对于科学数据库体系中分独立个体的化学专业数据库来说, 由于技术能力的局限性, 其数据安全无法得到完全的保证。更为重要的是, 化学专业数据库中数据信息的“独创性”和“额头出汗”原则难以自我鉴定, 在受到恶意下载甚至攻击破坏时, 版权法是否能保护其权益受到质疑。

因此, 数据库所有者必须要增强版权保护意识。数据库访问控制、用户管理、页面版权申明等措施, 都是有效的数据库安全和预防策略。当遭到其他单位

或者个人大规模批量恶意下载数据, 或将数据另作发表、使用, 不当获益时, 单靠《著作权法》可能很难追求其责任, 无法保护数据库所有者的合法权益。在我国首例数据信息诉讼案——“阳光诉霸才案”中, 法院认为阳光公司并不能有效的证明其“SIC 实时金融”数据库的“独创性”特点, 因此无法作为《著作权法》中的“作品”受到版权法保护。但是在数据的收集加工处理等过程中, 付出了一定的投资和工作, 其数据库可以看成是商品(产品)正当利益应受到相应保护, 最终法院以“不正当竞争”为由判令阳光公司胜诉^[7]。由此可以看出, 在某些特定案例中, 数据库所有者也应尝试《反不正当竞争法》, 来对其数据库信息权益进行保护。

另外, 在化学专业数据库实践过程中, 以项目合作的方式进行的数据共享、数据转让、数据使用等是常见的运作。在这种情形下, 数据库所有者和项目合作方签订相应的数据库版权合同条款, 是一种有效的保证数据信息权益的手段。当双方签订了合同条款后, 各自在条款内容限定范围内运作, 可以免去往后相关版权的纷争, 作为版权保护的补充, 《合同法》也是数据库所有者相关权益的另一重保障。

3 化学专业数据库知识产权保护的探索

3.1 数据创新

无论从科学数据库知识产权保护还是从其自身发展的角度看, 加强数据创新是科学数据库必须前进的方向。数据创新有两方面涵义: 其一, 首先要增强数据库数据的完整性、正确性以及权威性, 数据信息服务得到用户的认可, 科学数据库才有存在的价值和知识产权保护的必要; 其二是要加强数据库的“独创性”, 这不仅仅是使数据库成为“作品”受到版权保护, 更是科学数据库价值提升的关键。例如: 美国化学文摘社(CAS)出版的化学资料电子数据库SciFinder®和STN®, 已经成为学术界和工业界最权威的数据信息来源, 是化学、生物以及生命科学等研究领域不可或缺的和研究工具^[8]。它独占世界商业科研数据库数据信息服务市场最大份额本质是其

“独创性”，因其资料量巨大、学科领域覆盖全面、数据信息权威精确，并不断增加数据内容和更新数据信息，是其他类似数据库体系所无法比拟的。

化学专业数据库面对国内化学相关领域对化学数据服务不断增长的需求，和国外商业数据库价格不断上涨带来的挑战，需要进一步扩展现有化学数据资源的数量和规模，通过化合物基础数据库实现化学数据的整合以更好满足应用需求：继续增加化合物基础数据库的数量，争取达到千万级别，满足化学数据整合和化学行业数据信息的需求；通过数据采集和处理，研究材料数据的结构和特点，建立专业材料数据库，从而突破材料领域和化学学科交叉的发展瓶颈；建设化学物质生物活性数据库，包括生物活性实验数据和预测数据，从化学角度提供生物医药领域所需的数据和服务。

3.2 服务模式的转变

最原始的数据库服务以数据存储媒介光盘的形式发行，例如美国化学文摘社的 CAonCD®、中国化学文献数据库 CCBd-CD® 等。目前的科学数据库服务和共享模式以服务网站数据检索查询为主，传统的数据服务模式内容简略单一，服务质量评价一般，数据的使用率也较低，也给某些有意的数据侵权者提供了方便。如果侵权者能够轻易地获取科学数据库中的重要数据项，直接进行分发或者售卖，科学数据共享平台也就没有了存在的价值。在大数据时代，单纯的向用户提供简单的数据检索已远远不够，化学专业数据库应以数据为核心，从纯数据生产升级为科学决策，从简单数据查询升级为策略咨询，依靠大数据来支撑科研活动。人工智能是当今社会科技发展的新方向，化学专业数据库基于自身丰富的化学物质基础数据和化学反应数据资源，采取机器学习的策略研究和预测新的反应路径或者新的合成方法成为可能^[9]。未来，化学专业数据库可以依靠多年来积累的化学数据信息，将基础数据变为大数据预测的基石，让数据库所含“知识产权”更有价值，并能确立新的发展方向。

3.3 从社会效益到经济效益

数据库版权保护的本质之一，就是保护数据库所

有者的经济权益。科学数据库的宗旨和特性决定它并不会直接产生经济效益，但是如何让科学数据库产生社会效益甚至间接的经济效益，是一些无法单纯以科学院信息化项目维持发展的数据库的出路。当然，拿科学数据作为“商品”直接出售(卖数据)并不具有可持续发展前景，对数据库普通用户的收费也不符合“合理使用”原则，单个分散的数据库也不具备收费的可操作性。更符合可持续发展的策略，应该是利用科学数据，为科研工作提供数据支持、为市场商品提供数据信息、为企业生产活动提供数据保障，从而让科学数据库产生社会效益和间接的经济效益。例如，国药集团的中国试剂网，是国内最大的实验室产品和服务提供商，其试剂产品背后，更是一系列基础数据的支持，从试剂的物化性质、实验室数据、使用分析标准，到各种参考文献等等，一套完整的数据链才能支撑起一个有效的服务平台。

化学专业数据库除了继续参与科学数据库建设相关项目外，还应借助中科院科学数据大平台的影响力，利用现有的数据库建库技术和资源，充分扩展数据库的业务范围。加强与相关领域科研单位合作，研发更加专业的、更有针对性的化学数据库。加大支撑中小企业特别是初创企业的力度，以更加灵活的数据服务方式促进企业发展。对于数据价值衡量方式问题，化学专业数据库不应局限于数据采集处理耗费的人力物力成本，更应从知识产权保护的角度出发，将现有的数据加工处理技术和数据库建库技术申请为专利或者软件著作权，将技术和经验以版权认证的形式转换成经济价值。

4 总结

随着信息时代人们对知识产权的越来越重视，化学专业数据库也不得不面临新的挑战。作为国内最大的化学领域基础数据共享服务平台，化学专业数据库如何既能更好的服务科研工作者支持国民经济，又能建立起完善的知识产权保护体系，从而取得更好的社会效益和一定的经济效益，是数据库建设单位必须着重考虑的问题。对化学专业数据库知识产权保护的探

讨, 是让我们以知识产权的角度, 考虑科学数据库前进的方向, 利用数据创新以及数据服务模式创新去迎接新的挑战, 让科学数据库真正成为国民科普教育、经济社会文化发展的基石。

参考文献

- [1] 中国科学院计算机网络信息中心. 中国科学院科学数据库的建设与发展[J]. 中国基础科学, 2002(4): 50-55.
- [2] 陈维明, 朱翠娣, 赵英莉, 等. 论数据唯一标识与化学数据的集成[C]. 第九届科学数据库与信息技术学术研讨会, 广西桂林, 2008.
- [3] 朱雪忠, 徐先东. 浅析我国科学数据共享与知识产权保护的冲突与协调[J]. 管理学报, 2007, 4(4):477.
- [4] 刘润达, 孙九林, 廖顺宝. 科学数据共享中数据授权问题初探[J]. 情报杂志, 2010, 29(12):15-18.
- [5] 吴立宗, 王亮绪, 南卓铜, 等. 科学数据出版现状及其体系框架[J]. 遥感技术与应用, 2013, 28(3):383-390.
- [6] 黎建辉, 周园春, 胡良霖, 等. 专题:科学数据与创新应

用——中国科学院科学数据云建设与服务[J]. 大数据, 2016(6):1-13.

- [7] 袁晓东, 李晓桃. 电子数据库法律保护—兼评我国首例数据信息不正当商业竞争案[J]. 法学, 2000(5): 56-58.
- [8] 彭巍. 美国《化学文摘》带给我国化学文摘型数据库建设的启示[D]. 东北师范大学, 2007.
- [9] Mhs S, Preuss M, Waller M P. Planning chemical syntheses with deep neural networks and symbolic AI[J]. Nature, 2018, 555(7698):604.

收稿日期: 2018年2月12日

徐挺军: 中国科学院上海有机化学研究所, 工程师, 硕士, 主要研究方向为化学数据库。

E-mail: xutingjun@sioc.ac.cn

陈维明: 中国科学院上海有机化学研究所, 课题组长, 博士, 主要研究方向为化学信息学。

E-mail: chenwm@sioc.ac.cn