

国家知识产权文献及信息资料库建设 内容选择及建构思路探析

冯晓青¹, 赵秀姣²

(1.中国政法大学 无形资产管理中心,北京 100088; 2.北京交通大学 图书馆咨询部,北京 100044)

摘要:目前,国内法学文献信息数据库缺失,现有文献检索平台不足,是制约法学学术研究和相关实务的重要因素。针对我国知识产权文献及信息资料库的特点,从信息源选择、资料库平台建设、数据的收录加工过程等方面对国家知识产权文献及信息资料库的建设进行研究具有重要意义。

关键词:知识产权;文献及信息;资料库;检索平台

中图分类号:D913.4;G250.74

文献标识码:A

文章编号:1674-9014(2012)05-0050-06

目前国内法学文献信息数据库及检索平台,主要有两种形式:一种以专利检索为主,如国家知识产权局、万方专利、Socpat、百度专利搜索,为国内科研单位和企业专利用户提供了比较丰富的检索源;另一种以北大法意、北大法宝为代表,涵盖除了专利以外的法学文献资源,是法学科研究与学习的重要平台。以知识产权法为主的专业数据库检索平台在国内则尚未建立,因此,2010年度国家社会科学基金重大项目“国家知识产权文献及信息资料库建设研究”的目标是拟建立国内知识产权文献及信息资料库,为我国知识产权文献信息保障工作提供强有力的支撑,服务于我国创新型国家建设目标。为建立国内一流的国家知识产权文献信息平台,本文拟从信息源选择、资料库平台建设、数据的收录加工过程等方面对资料库的建设提出建议,从而推动该资料库建设的研究。

一 国家知识产权文献及信息资料库涉及的内容

信息内容是资料库的生命,我们在工作中,每天都会使用各种不同的数据库,在学科相近的数据库中进行选择时,通常会将信息量的大小作为选择的第一标准。国家知识产权局、中国知识产权网、中国

专利网、中国专利信息网、CNKI、万方专利检索、Socpat、Patentics和baidu专利搜索平台,不论是免费还是收费检索,都以中国境内专利文献的检索与分析为主。北大法宝提供国内的法律法规、司法解释和案例、仲裁裁决、裁判文书、中外条约、港澳台法律、法律文书和法学论文的检索。北大法意的文献包括法院案例、法律法规、合同文本、法学词典、WTO法律、法学论著和司法考试库。法学文献覆盖学科范围广泛,本文通过分析这些数据库信息量的优势和不足,结合我国知识产权法制建设现状和各类用户对知识产权文献及信息需求,总结出国家知识产权文献信息资料库的文献信息类型应包括以下内容。

(一) 知识产权法律法规及其他规范性文件

知识产权法律法规及其他规范性文件,主要包括专利、商标、著作权和其他知识产权的法律、行政法规、部门规章、地方性法规、地方规范性文件、立法与司法解释、军事法规规章、部门等专门系统的知识产权规范,以及我国参加的知识产权国际条约、团体与行业规范等。上述规范性文件特别是知识产权法律法规,是国家知识产权文献及信息资料库中最基本的文献及信息资料。应当说,目前现有的法律数据库和一些涉及法律和知识产权的网站,都有大量这

收稿日期:2012-07-13

基金项目:国家哲学社会科学基金重大项目“国家知识产权文献及信息资料库建设研究”(10&ZD133)。

作者简介:冯晓青,男,湖南长沙人,中国政法大学无形资产管理中心主任,教授,博士,博士生导师,研究方向为知识产权。

方面的资料与信息。但从用户需求、用户查找方便和内容挖掘深度来看,现行网站和相关资料库存在的问题是:第一,内容不够全面,一般只涉及其中最重要的知识产权法律法规,尤其是缺乏地方性法规和规章;第二,基本上以呈现信息为主,缺乏对法条背后的法理解释、制度沿革等深层次的专业加工和标引,非常缺乏专业增值服务;第三,缺乏立法过程、立法解释等重要文献及信息,不利于对知识产权法律规范的深度理解。

为克服上述困难,拟建设的国家知识产权文献及信息资料库在知识产权法律法规及其他规范性文件内容的构架方面,应采用以下模式:第一,按照上述关于知识产权法律法规和其他规范性文件分类标准,收纳全部现行有效文件和部分已经失效的文件,便于系统学习、研究和参考;第二,对其中重要条款、重要概念、重要制度和原则,分别进行深度专业标引,使用户可以方便地获取相关附加的专业信息;第三,保持适时更新状况,对新增的这方面文件,及时进行增补和深度专业标引。

(二)知识产权司法案例

知识产权司法案例涉及我国近些年来来的知识产权司法判例,具体可分为一审、二审、再审、调解等类别。随着我国企业和公民知识产权意识的加强以及知识产权法制建设的深入发展,我国各级人民法院审理的知识产权纠纷案件日益增多,仅以2010年为例,当年一审知识产权民事纠纷案件即有60000件之多。知识产权司法案例是特定历史时期发生的个案,而判决书本身则体现了人民法院知识产权法官对案件事实认定和适用法律的把握和理解,具有非常重要的学理研究价值和处理同类案件的指导意义。基于此,知识产权司法案例也是国家知识产权文献及信息资料库中极为重要的文献与信息。众多的这些案例,可以为知识产权法官、律师处理同类案件和提高自身专业素养提供极佳的研究素材和资料,也可以为知识产权教学研究人员和学生学习与研究知识产权实务提供良好的实务素材,当然也是相关案件当事人处理案件的重要参照。然而,从现有的法律数据库或相关网站涉及知识产权司法案例的情况看,情况不容乐观。主要问题如下:其一,信息不齐。受信息公开、透明度不够的制约,法律数据库或相关网站涉及知识产权司法案例数量有限,很多案例无法从公开的渠道查到;其二,信息单一,缺乏深度专业标引。很多信息源只是简单地重现判决书,没有延伸的相关信息。

为克服上述困难,拟建设的国家知识产权文献及信息资料库将着重于以下方面:一是知识产权司法案例收集的全面性、完整性,要尽量涵盖所有公开的知识产权司法案例。二是知识产权司法案例的专业加工与深度标引。拟提炼每个司法案例的内容提要、关键词与主题词,对每个案例中的关键概念和原理进行专门解释和延展性注释,同时通过技术手段,实现对同类案例的快速检索和分类。除此之外,建立每个案件的案例地图也将是一个创新性工作。三是及时新增最新案例,并对最新案例进行专业加工。通过这些手段旨在实现以下目的:用户能够方便地找到需要找到的案例;用户能够对选取的案例进行深度挖掘,更好地理解案例的精神和背后的法理;用户能够方便地进行同类案件的查找和学习,便于系统学习和研究。

(三)知识产权论著资料

知识产权论著资料涉及知识产权方面的专著,期刊论文、学位论文、会议论文等。这些文献及信息是从事知识产权研究,开展知识产权教育和人才培养的基本保障。目前,国内相关数据库和网站收录的这些知识产权论著资料比较分散,缺乏系统的整合,同时存在资料信息不齐的现象,拟建设的国家知识产权文献及信息资料库在内容上需要包含这些板块,并注意信息的及时更新。

知识产权类图书特别是专著,是内容比较成熟、资料比较系统的文献类型,国内有关知识产权专著众多,但目前这些文献零星地分散在各综合学科的图书数据库中,国内有超星电子图书数据库、方正电子图书数据库和书生电子图书数据库。关于国家知识产权文献及信息资料库建设应纳入总库,作为法学研究工作者、学者、专业学生的导读系统,并提供图书的全文在线浏览。当然,由于知识产权著作涉及相应的著作权问题,在将已出版的这些图书纳入总库中时必须先解决好这些作品使用的著作权问题。

期刊文章因为流通周期短,快速呈现专业领域学术科研成果,是除专门法律文书外数量最多的文献类型。需要关注国内期刊特别是核心期刊中涉及知识产权方面的文章,实现期刊文章的全文下载,并提供期刊的相关信息检索,如期刊基本情况介绍、板块分类、投稿方式、通信地址与级别等。当然,这里也涉及如何处理好相关的著作权问题。

学位论文的学术性较强,以国内主要科研院所、高等院校的知识产权专业硕士及博士学位论文

为主,是学生们多年研究的经验总结,在一定的程度上能反映出国内知识产权教学的成就,因而也是知识产权论著文献及信息收录的重要内容,必要时也可收录少量的优秀本科学士学位论文。

国内各种学术会议、专题会议上发表的论文和报告,多数以会议录的形式出现,是专业领域内最新研究成果报道的一种主要方式。法学领域的会议更是如此,在拟建设的国家知识产权文献及信息资料库中,可考虑收集入库单独归类。

(四) 知识产权科研项目及成果

知识产权科研项目是各类机构资助的关于知识产权的研究课题。近些年来,随着知识产权问题变得日益重要,我国各级政府部门和相关机构发布、批准了一大批涉及知识产权方面的研究课题。这些课题大致可以分为“纵向”、“横向”两类。当然,不同的单位对两者的划分标准并不统一。以中国政法大学为例,其《科研典》将国家级和省级自然科学基金、社会科学基金以及教育部人文社会科学规划项目和司法部人文社会科学规划项目^①,以及学校自身发布的年度人文社会科学规划项目纳入“纵向”课题,其他各类项目则纳入横向研究课题。无论是纵向课题还是横向课题,课题立项本身的信息对于指导后续课题立项和申报具有重要的指导意义。至于课题成果更是知识产权学术成果方面的重要组成部分。

从目前我国涉及知识产权科研项目的数据库和相关网站的情况看,存在的主要问题是:其一,信息严重缺乏,除了中国高校人文社会科学信息网、国家自然科学基金、国家社会科学基金等网站能够比较全面地找到相关课题信息外,其他信息很难找到,更不用说课题成果的公开。其二,大量课题成果以内部信息的形式存在,没有充分发挥资源共享、知识共享的特点,以致使国家和地方花费了很多资金资助的课题成果,实质上没有得到最佳社会效果。

为了克服上述问题,使我国数以千计的知识产权研究项目及其成果能够产生更大的作用,拟建设的国家知识产权文献及信息资料库将建立专题的知识产权科研项目板块,在对现行知识产权科研项目进行详细分类的基础上,尽量全面展示现有知识产权课题成果的信息和内容。其中,课题成果简介和精要将是这类文献及信息的实质内容。为此,需要各课题负责人予以配合,及时将其主持的各类课题成果的简介和精要(一般要求不低于 3000~5000 字)发送给资料库。这样,既可以扩大课题成果的影响,也

可以更好地发挥课题成果的社会效益。设想一下,未来的国家知识产权文献及信息资料库如果能够做到将不涉及保密的所有知识产权研究项目成果的简介和精要发布,读者将从中获取巨大的教益。

(五) 知识产权机构

知识产权机构涉及知识产权公共机构(代表公权力的知识产权行政、立法、司法机构)、知识产权教学研究机构(如事业性质的各高校知识产权院、系、所,科学研究所的知识产权中心、知识产权研究所)、知识产权服务机构(如企业性质的知识产权代理公司、事务所、服务中心)等。知识产权机构文献及信息旨在了解各类知识产权机构的职能和运作,充分利用其进行学习、研究、实务运作。以知识产权教学研究机构为例,其涉及的机构特色等的重要内容介绍、知识产权研究生培养方案、课程设置、导师信息等,能够为各知识产权人才培养单位开展知识产权教学科研活动提供启发,也可为学生们报考相关学校和单位提供决策依据。再以知识产权服务机构信息为例,它能够开展知识产权服务,促进知识产权运营提供较好的实用信息。从目前我国涉及知识产权内容的数据库和相关网站提供的信息看,尚缺乏完整、系统的知识产权机构文献及信息。为此,拟建设的国家知识产权文献及信息资料库将整合现行我国各类知识产权机构的信息,分门别类地提供相关文献和信息。

(六) 知识产权人物

知识产权人物文献及信息涉及知识产权领域的代表性人物的基本信息和从业经历、代表性成果与观点等。包括事业管理名人、专利审查与复审名人、商标注册与评审名人、版权登记与管理名人、知名法官、知名学者、知名代理人 and 律师、信息与社会服务名人及其他重要人物信息等。知识产权人物库事实上是我国知识产权人才库的重要内容。目前,国家正在大力加强知识产权人才工程建设,如“百千万知识产权人才工程”、国家知识产权专家库建设以及各省市的知识产权人才工程建设等。目前,我国现行涉及知识产权的数据库和相关网站也缺乏系统的知识产权人才库信息。为加强知识产权人才工程建设,充分发挥知识产权人才的作用,拟建设的国家知识产权文献及信息资料库将整合现有知识产权人才信息,建设富有自身特色的知识产权人物库。同时,为更好地宣传和发挥知识产权高级人才在推动我国国家知识产权战略实施中的重要作用,拟在知识产权人物库中专门建立知识产权高级专家库。

(七) 知识产权大事记

知识产权大事记反映了我国不同历史时期发生的知识产权方面的重大事件、重要活动。目前我国对知识产权大事的记载主要存在于《知识产权年鉴》和事件发生期间的媒体报道中,总体上缺乏系统收集这类文献及信息的资料库。为此,拟建设的国家知识产权文献及信息资料库将以19世纪以来知识产权领域内具有重大影响的大事记以年代或主题列序,涉及知识产权立法、司法、行政管理、教育与人才培养、政策与文化、相关的工作与会议、涉外知识产权合作、涉外知识产权纠纷等为内容。

(八) 知识产权确权信息及专利分析工具

知识产权确权类信息主要涉及专利申请、授权、无效等涉及专利确权的信息,商标注册申请、核准、异议、争议、撤销等涉及商标确权的信息,以及计算机软件著作权登记、集成电路布图设计登记等方面的信息。知识产权确权类信息为科研人员、技术开发工作者了解最新技术和国内国际发展态势提供了最直接的信息,特别是其中的专利信息,除了有效专利外,还涉及申请未授权、失效、转让、许可、担保和其他相关信息,它是我国科技人员科研立项、进行研究开发活动的关键信息,直接关系到我国创新能力建设。目前国家知识产权局和商标局分别建有公益性的检索数据库网站。例如,专利数据检索平台主要由国家知识产权局及下属部门或专利数据研究公司主办,提供我国现有专利的检索,全面且更新及时。能实现免费检索的平台有中国知识产权局、中国知识产权网、中国专利网、中国专利信息网和 baidu 专利搜索。商业性的数据库以 CNKI、万方专利检索为代表。

专利分析工具的产生缘于竞争激烈的技术市场。在检索专利文献的同时,能有效地为企业发展提供技术策略并能为评估竞争对手提供有利的情报。免费的专利分析软件以 Sopat 和 Patentics 为代表,商业的分析门户平台有东方灵盾(East Linden Doors,由北京东方灵盾科技有限公司与美国 MCAM 公司合作开发)。LexisNexis 公司的 Totalpatent(全球专利信息数据库)也是目前国外专利信息检索与分析的重要数据库。而 Dialog 推出的新一代的专利信息检索和分析服务平台“Innography 专利检索分析系统”值得借鉴,它通过独有的专利强度分析工具快速地挖掘出核心专利,帮助用户在海量的数据中迅速定位目标,整个检索分析过程都是可视化的,不需要输入复杂的检索指令,在节省检索时间的前

提下提高了阅读的质量。

由于知识产权确权类信息庞大、信息更新速度快,加之检索的要求高,未来的国家知识产权文献及信息资料库是否以及如何包含和体现知识产权确权文献及信息,是一个值得深入研究的问题。如果将其纳入国家知识产权文献及信息资料库,则意味着需要巨大的资金投入,并且需要处理好与现有数据库之间的关系,以免造成重复建设,如果只是提供信息,而不是提供文献本身,则相对容易,但也需要与现有数据库之间建立合作关系。

(九) 其他相关内容

在拟建设的国家知识产权文献及信息资料库中,还将建立知识产权主题词表,便于用户随时了解和掌握知识产权相关概念的内涵;编撰知识产权法学词汇的中文词典,收录国内知识产权法学专业词汇表,并列出具体的同义词、近义词和可替代词汇供作者查阅。介绍专业名称缘由,提高趣味性。多语种专业知识产权电子词典,以英语、法语、德语和日语为主要语种,提供最专业的知识产权法学词汇数据库,实现实时查询和在线翻译。同时,为了最大限度方便用户自行处理相关的知识产权问题,拟建设的国家知识产权文献及信息资料库还将设立知识产权自助板块,用户可以根据自己的需求直接查阅信息,获取相关的知识。

知识产权文献及信息包含丰富的内容,随着我国知识产权法制建设的深入开展,国家知识产权文献及信息资料库也需要与时俱进,适时充实新的内容。

二 国家知识产权文献及信息资料库建构的思路

(一) 资料库界面设计:友好易用

有了一个资源丰富的资料库并不代表用户一定乐于使用。我们经常看到很多数据库在介绍信息中列明了海量的数据量,但一打开平台,感觉并不佳。例如,有些界面凌乱,广告信息占据大部分界面;有些检索平台复杂,显然是针对专业情报检索人员的,普通大众很难快速适应;有些检索平台则过于简单,不便于短时间内提取相关度高的文献。现阶段的资料库检索技术十分有利于专业资料库资源的开发,各种检索方式的结合和真正意义上图片、全文检索是资料库革命性的改变。一个好的资料库检索平台一定要充分利用这些技术优势,方便资料库资源的提取。以下拟从普通检索人员的角度对既符合读者检索习惯,又符合知识产权法学专业

要求的友好平台建设提出如下要求。

1. 建立知识产权法专业信息分类体系。由于国外法律体系不一样,分类体系也不尽相同,如 West-law 采用 West 主题分类法(West Topics)与钥匙码检索分类系统(Key Number System),提供 400 多个领域的法律分类,覆盖的主题达到 290 万个,几乎覆盖全部法律实践部门。我国现有的知识产权法知识分类以“中国图书馆图书分类法”为基础,从各种不同机构对法学信息的分类来看较为粗糙,并没有专门针对知识产权的分类主题表。

遵循中图法的分类体系,结合知识产权法的学科特点,可确定知识产权本体,辅以汉语拼音字顺索引表,使文献信息简明集中。

2. 检索字段选择多样化。鉴于文献类型的多样化,信息在加工过程中应提炼出尽量多的字段以供读者选择,这样有利于信息的提取。CALIS(China Academic Library & Information System, 中国高等教育文献保障系统)针对不同的文献类型确定了元数据标准,可作为普通类文献如专著、期刊、学位论文、会议文献与年鉴的元数据。对于其它专门类的特色知识产权文献如判例、法律法规等文献,除了常用的标题、主题、摘要、责任者、时间、地点外,应增加如当事人、辩护律师、法官、法院名称、法律法规编号、专利的法律状态等特殊字段。

3. 专业检索、高级检索的优化。目前国内的数据库检索平台通常按检索难易度分为简单检索、高级检索与专业(专家)检索,逐次满足对数据提取精确度的不同需求,其中专家检索用来满足高端用户的需求,但使用过程相对复杂。而高级检索通常如同虚设,普通用户要么习惯用简单检索的方式使用高级检索,要么发现高级检索不能真正实现优于简单检索的结果。一个既能让普通用户和专业情报人员都适应的高级检索,在一定意义上能完全代替专业检索,可通过如下途径来实现。

首先,同一字段词汇之间逻辑组合方便易用。在大部分中文数据库的检索过程中,同一检索字段的逻辑“与”、“或”、“非”三种关系要实现同步一次性检索,通常只能通过专家检索的方式。而中国知识产权局网站的专利检索平台,对于同一字段的词汇之间的逻辑关系实现用“OR”、“AND”、“NOT”同步检索,非常方便适用,值得借鉴。

如在“专利名称”字段中可输入:轻轨 OR 城轨 OR 地铁 AND 交通 AND 运输

其次,字段数量设置灵活。通常情况下,4 至 5

个字段的组合能满足绝大部分用户的检索需求,如果用户对字段量有增加需求时,可设置增加减少按钮由用户自由选择。清华同方公司的中国知网(CNKI)主页高级检索设置了这样的增删按钮,如下图所示:

+	□	(全文)	<input type="text"/>
并含		(题名)	<input type="text"/>
或含		(主题)	<input type="text"/>
不含		(关键词)	<input type="text"/>

图 1 中国知网(CNKI)主页高级检索截图

再次,单库与跨库检索平台并存。既然有不同的文献类型,必然要用不同的检索界面,才能实现真正意义上的精确检索。但在同一个检索平台里,也必须要能实现多库统一检索但精确度相对较差的跨库检索平台,以满足不同级别读者的需求。

第四,检索结果输出多样化。读者在进行检索结果输出时,所提供的内容不仅要根据文献类型的特点,而且要根据法律类文书的特点提供多种字段的输出,选择范围涵盖所有的字段,读者可根据个人需要选择以全文、摘要信息或参考文献著录格式的形式输出。

第五,与文献管理软件无缝链接。除了普通的纯文本输出格式,科研工作者更期望文献的管理能得到软件的支持。目前,比较通用的文献管理软件,国外有 ProQuest 公司的 Refworks 和 Thomson 公司的 EndNote 软件,国内的有北京爱琴海软件公司的 NoteExpress。文献管理软件能方便读者对从不同数据库检索的数以百万计的电子文献进行管理,包括帮助检索并管理得到的文献摘要、全文,在撰写文献的过程中,可在正文中的指定位置方便地添加文中注释,然后按照不同的期刊、学位论文格式要求自动生成参考文献索引,能自由做笔记,并实现高效有序地管理笔记内容等功能。与这些软件供应商合作,不仅能实现与国际数据库的接轨,更能帮助读者在整个科研流程中高效利用电子资源。

(二)动态热门信息与定题分析服务

实现动态信息浏览,可在检索平台的主页设定非检索入口并提供如下内容:首先,主题为为核心的知识产权信息集中;其次,无序却最新最热门信息;再次,以点击量决定的优秀文献。

定题分析服务是高端情报服务。设定个性化账户以登录,使每一个合法用户能设定自己的个性化

界面,保存自己的检索历史和检索偏好,并提供定题信息推送服务,比如当读者设定好一个检索式后,一旦有新的文献出现,主动将该文献自动发送到读者的注册邮箱,减免读者的重复检索,更能帮助读者第一时间掌握最新动态,特别有利于特定用户根据案例跟踪进程变化。

(三)数据收录和加工的科学规范

数据的收录、加工过程是繁琐复杂的工作,为有利于计算机识别,实现电子、数字化,有利于数据重加工和整理,收录、加工过程的科学规范是头等大事。尤其在技术相对比较先进的今天,我们提供的检索结果要最大程度在利用现有先进技术的基础上满足不同用户的需求,例如,对一个好的外观设计、实用新型专利数据库而言,应充分利用现今图片检索的优势,便于用户实现用图片对图片的检索。对这部分文献及信息的收集要求就比较高,既往的文献格式通常很难实现对图片的计算机解析。资料库建设工作者和技术开发人员应形成良好的沟通渠道,在科学思想的指导下有序地进行数据收录和加工工作。

总之,国家知识产权文献及信息资料库建设是一个巨大的文献信息服务工程,在我国大力加强信息化建设,特别是国务院多个部委联合发布的《国

家知识产权事业“十二五”规划》将知识产权信息服务工程作为重大工程之一的背景下,加快我国知识产权文献及信息资料库建设具有不寻常的意义。此外,尽管是国内的知识产权文献数据库建设,也不能忽略国外知识产权文献及信息对中国用户的参考意义。目前我们熟知的国外的法学数据库有LexisNexis、Westlaw和HeinOnline等。很多高校图书馆已经购买了这些商业数据库,可直接在本地建立链接,方便读者同步阅读。也可根据实力单独建库,报道最新的国际动态、国际专利和最新科研成果,给中国的科研和企业用户提供全面的服务。

注 释:

① 国家司法部目前已终止发布这类年度研究项目。

参考文献:

- [1] 胡晓,魏雪梅.我国网络专利检索平台分析和评价[J].科技管理研究,2010(14):75-77,81.
- [2] 苏建华,邹育理.国内外法学数据库的比较研究——透视我国专业数据库建设存在的问题[J].现代情报,2008(5):79-81.
- [3] 刘刚,王燕荣.《中图法》法律文献分类体系刍议[J].政法论丛,1998(3):61-63.

(责任编辑:刘英玲)

(责任校对:张群喜)

National Copyright Literature and Information Database Construction: Choice of Content and Thought of Construction

FENG Xiao-qing¹ ZHAO Xiu-jiao²

(1.Intangible Capital Management Research Center, China University of Political Science and Law, Beijing 100088, China;

2.Advisory Department of Library, Beijing Jiaotong University, Beijing 100044, China)

Abstract: Nowadays, law literature database in China is lack of enough literature retrieval platforms. This paper, according to the feature of copyright literature and information database, studies the construction of copyright literature and information database from the aspects of choice of information source, construction of database platform and data including and processing, which is of significance.

Key words: copyright; literature and information; database; retrieval platform