

语言经济知识产权治理结构研究

董 涛

(北京第二外国语学院语言资源高精尖创新中心, 北京 100024)

摘 要:现代社会,语言不仅是人类重要的文化资源,同时也是一种稀缺的经济资源。语言资源能够带来巨大的红利,与语言相关的经济活动不断增多,逐步形成经济增长的一个新的因素,并催生语言产业这一新的经济形态。语言经济是以语言文字为内容、材料,或以语言为加工、处理对象,生产各种语言产品或提供语言服务的经济形态。知识产权作为保护语言成果最重要的产权制度,对于语言经济的产业结构、产品形态及竞争秩序都产生着巨大的影响,不仅规范着语言产品交易的市场秩序,型塑着语言经济的结构与方向,而且在三重意义上拓展着语言资源的价值范畴。深入研究语言经济的运行规律,发现语言资源、产品与成果的内在特征,构建一个良好的知识产权治理结构,以引导与推动语言经济快速、有序地发展,具有十分重要的意义。

关键词:语言;语言经济;知识产权;治理结构

中图分类号:DF523

文献标识码:A

文章编号:1005-0566(2022)09-0016-11

IP governance structure in the language economy

DONG Tao

(Innovation Center of Language Resources, Beijing International Studies University, Beijing 100024, China)

Abstract: In Modern Society, language is not only an important kind of cultural resource, but also a scarce kind of economic asset. Language can bring huge profit. With the increased economic activities related to language, there comes a new kind of economic form of language. The economy of language includes mainly those industries manufacturing products or providing services with language as major factors. As an important legal mechanism to protect the language product, the IP regime affects industrial structure, product mode and competition order largely in the language economy. IP regime not only regulates market order of language products, shapes the structure and direction of language economy, but more importantly, it will expand the value of language resources in three dimensions. The deep research of internal rule of language resources, products and fruits can make significant contribution to the establishment of good IP governance of language economy.

Key words: language; language economy; intellectual property; governance structure

一、语言经济中的知识产权治理

语言经济是以语言文字为内容、材料,或以语言为加工、处理对象,生产各种语言产品或提供语言服务的经济形态。语言经济不破坏自然环境,

不产生工业污染,是一种绿色环保的经济形式,契合了高质量发展的时代诉求^[1]。语言产品绝大多数是智力劳动成果的知识产品,是知识产权保护的重要客体。知识产权对于语言经济的产业结

收稿日期:2021-09-06 修回日期:2022-08-28

基金项目:国家语言资源高精尖创新中心项目“语言资源知识产权法律保障研究”。

作者简介:董涛(1971—),男,四川乐山人,北京第二外国语学院语言资源高精尖创新中心教授,法学博士,研究方向为知识产权法律、政策与战略。

构、产品形态及竞争秩序都产生着巨大的影响。语言经济中的知识产权治理就是通过确认语言相关产品与服务产权归属的方式,为语言市场与语言产品与服务的交易带来规范的秩序,从而推动语言经济良性、快速发展。语言经济中的知识产权治理结构具有以下几方面的特征。

(一)治理基础:多级架构的规则供给

语言经济业态,包括语言教育与培训等服务,语言产品的生产,语言产业的发展与语言市场的构建等,会受一国语言规划、语言教育、民族政策、出版发行、新闻管制等法律、法规的影响,表现为一种多层级的规则框架体系。截止到2016年,我国与语言文字有关的法律法规近2200项,形成包括宪法、法律、法规、规章和规范性文件在内的5个层级的法律法规体系^[2]。仅2000年之后颁布的各种语言法规就有上百项。2000年的《国家语言文字法》标志着我国语言文字事业走上法治化轨道。在这个多层级的规则框架下,我国建立起了政府主导、语委统筹、部门支持、社会参与的治理体系。语言的知识产权治理是嵌入这个大的治理规则框架体系中的一部分,包括专利法及其实施细则、商标法及其实施细则、著作权法及其实施细则3部主干法为基础,以及各部委与地方颁发的关于创新激励、管理、运用等法规为补充的知识产权规则体系。

(二)治理对象:权利法定的准则主义

知识产权奉行的既是一种权利法定的原则,即除法律明确规定外,任何人不能自行创制权利的类型,决定权利的范围,又是一种准则主义原则,即任何智力成果合于条件时,都能得到法律的承认。语言产品只要符合知识产权法律规定的条件,就能够进入这一治理体系,获得相应的保护。细化起来,主要包括以下几个方面:从客体类型来看,不是所有的语言产品、服务、技术与符号都能获得知识产权保护,只有那些满足这一条件的才能纳入到保护框架之中。否则将无法受到保护。从取得条件看,知识产权的取得也必须符合法律

预先设定的条件。如是语言作品,必须符合版权法的“思想/表达”二分法要求,获得保护的只能是独创性的语言文字组合,不能延及背后的思想。如是语言技术,必须符合专利实质性条件,即新颖性、创造性、实用性的“三性”。如是标识符号,则必须具备商标法上的识别性。从权利权能看,知识产权法从“长、宽、高”3个维度,即权利存在的期限、权利行使的内容与形式、权利覆盖范围的大小、权利的地域范围等方面,对权利的内部结构进行了限定,使知识产权的权利范围与其对社会的贡献相匹配。

(三)治理形式:“授权/救济”的二分结构

知识产权领域一般采用了一种由行政机关前置负责审查授权、司法机关负责纠纷诉讼的“授权/救济”二分结构治理形式。语言技术专利与语言文字商标都采用的是这种治理模式。不过,当今世界各国对版权作品则采取了另一种模式,即不设前置授权程序,在遭遇侵权诉讼时,由法官集中确定作品是否符合版权规定的条件以及是否存在侵权问题于一身。知识产权制度通过这种事前确权的方式,向社会公众告知知识产权的权利边界,既防止第三人未经许可随意侵入创新者的自由空间,也可使得社会公众预知知识产权布局情况,发现其中可以自由行动的空间。同时,还以事后救济的方式惩戒侵权行为者,恢复受到侵害的创新秩序。知识产权通过这种治理模式可以为语言产品与服务的供给者提供稳定且可预期的市场竞争秩序。

(四)治理主体:政府主导的共治空间

在语言知识产权治理结构中,涉及的主体较多,除负责专利、商标、版权、地理标志等知识产权事务的管理机构与负责救济的司法机构外,还包括其他如科技、产业、教育、财政等中央与地方政府机构,以及负责语言规划、新闻出版、教育机构等的参与。语言产品、服务、技术,大多会在网络之中流动,因此,网络平台也在语言经济知识产权治理结构中起着越来越大的作用^①。大型集市管

① 如阿里巴巴、腾讯等构建的知识产权维权机制,每年要处理成千上万在其平台上销售产品涉及的知识产权纠纷。

理者、会展组织者对其所管理的营业场所、举办的活动中知识产权保护都承担着责任。广告发布者、邮政、快递经营者等主体,则通过审查、收寄检视等制度对知识产权侵害行为进行防范。可以说,在知识产权领域中,正在形成“政府主导—平台自治—舆论监督—自力救济”为主的多元共治空间。

据此,语言经济中知识产权治理结构可以看做是围绕特定客体(语言作品、语言技术、语言符号)的生产(版权中的“创作”、专利中的“发明”、商标的“设计”)、分配(版权中以“作者—版权人”、专利中以“发明人—专利申请人—专利权人”、商标中以“设计人—商标申请人—商标权人”关系展开)、消费(“权利人—许可人—侵权人”)关系展开的规范与管理活动。语言经济中知识产权治理主体除了包含语言经济一般治理主体之外,还应包含知识产权行业的各治理主体。语言经济中知识产权治理手段则包括各种规范围绕语言产品、技术、符号产生的利益归属关系的法律法规、政策文件、计划规划、禁令命令等复合措施。知识产权治理结构从以下几个方面影响着语言经济的发展。

1. 确认语言成果产权归属,规范市场交易秩序

文字作品、口述作品,以及相声、曲艺等作品,可以获得版权的保护,计算机软件、字体字库、影视配音字幕等也可纳入到版权的客体范畴。与语言相关的技术,可以申请专利权保护。一些具有区别性的语言文字符号,可以注册为商标。其他一些法律法规对特定的语言文字符号也进行保护,如《地名管理条例》,特殊标志的专门条例以及居民姓名与企业名称的登记管理条例等。《反不正当竞争法》对一些使用他人标志性文字符号标记的行为足以产生混淆的行为进行了保护和规范^②。通过为语言成果提供产权归属的方式,知识产权起到定纷止争,将市场经营者从仿冒、盗版、

侵权等行为中获得收益的消极动力转化为从技术创新和诚实经营中获得收益的积极动力,规范并净化着市场的交易秩序。

2. 创设语言产品形态,型塑市场竞争结构

语言产品的客体信息摆脱了有形物质形体的束缚,因此其上的财产权可以以一种切片的方式,创设不同的权利形态,形成“权利束”^[3]。知识产权是典型的以权利形成产品的领域,几乎每一权利形态都可形成一种产品形态。知识产权的内部制度设计,影响和决定着语言产品的生产频率、结构与方向。以版权作品为例,以语言文字为起点,利用版权的不同权能可以构建起不同类型的产品形态,从纸面的平面印刷形态延展到不同国家语言的版本;通过对三维形象角色进行保护延展到影视产业和衍生周边产品;通过对有声读物的保护形成有声读物产品,冲击“耳朵经济”。语言产品借助知识产权的影响力,产生着“种子功能”,在动漫、手游、话剧、玩具,甚至主题公园之间转化,衍生出包括文学、动漫、影视等领域的产品,形成丰富的产业链条,构建起一个立体化的产品生态体系。

3. 提供语言新需求,开辟语言新市场

随着经济发展,越来越多以知识产权为核心的语言产业涌现出来。以品牌命名产业为例,这个产业几乎就是依赖于商标保护而产生的。一些著名品牌如 Acer(弘基)、Lucent(朗讯)等,都是命名产业的杰作。据调查,1999 年全美所有品牌命名总量约 600 万次,年产值总计高达 150 亿美元。美国《财富》杂志认为命名业是美国增长最快的新产业之一,相关业务延伸到了国际市场^[4]。就现代语言信息处理产业来看,几乎密布着计算机软件与信息处理技术,属于完全依赖于专利与版权保护的专利密集型与版权密集型产业。此外,由于专利申请需要提供标准格式的技术背景说明书与权利要求书,需聘请专门的代理人员进行撰写。

^② 反不正当竞争法第 6 条规定:“经营者不得实施下列混淆行为,引人误认为是他人商品或者与他人存在特定联系:(一)擅自使用与他人有一定影响的商品名称、包装、装潢等相同或者近似的标识;(二)擅自使用他人有一定影响的企业名称(包括简称、字号等)、社会组织名称(包括简称等)、姓名(包括笔名、艺名、译名等)。”参见《中华人民共和国反不正当竞争法》第 6 条。

仅2018年一年,全球专利申请总量高达670万件^[5]。这构成一个庞大的语言撰写与翻译市场。知识产权提供了语言新需求,开辟了一个庞大的语言新市场。

二、表达形式:语言作品中的版权治理

语言是思想的表达形式,天然地与版权结合在一起。版权法是一个通过围绕“作品”概念展开运行的规则体系,但只有符合版权法要求的语言作品,才能获得版权的保护,进入版权治理的范畴。版权制度通过使用一种切片式的方式将语言成果进行切割,使得语言文字资源从按照语法规则排列堆砌的冰冷文字转化为一个个类型多样的产品(文字作品),甚至深入人心的鲜活灵魂(形象化权),并进而演变为控制不同价值链条的产业生态体系。

(一)治理客体的选择性:语言作品“思想/表达”二分原则

版权的基础理念是以法权的形式为“文学、艺术和科学领域内具有独创性并能以一定形式表现的智力成果”提供保护,通过这种方式,激励人们进行更多的差异化与独创性的个性表达。由于智力成果本身仅在人的头脑之中,无法让他人感知、复制和传播。为了使其享有财产权成为可能,必须寻找、构造或拟制一种作为财产权利支配的对象。这一任务落在了“作品”这个概念之上。福柯^[6]认为,作品就是专有名词符号表示出来的文本。海德格尔^[7]则认为,作品不仅是一个物,而还有别的东西附着于这个物之上。在美国宪法中,“作品”一词原文用的是“writings”,本意是指用文字写作的东西,尤其指手稿。在美国1973年的“戈德斯坦”案中,最高法院对宪法中的“writings”作了广泛的界定:“‘writings’一词可以局限于手稿或印刷品,但它可以被解释为包括了所有的,体现了创造性的智力劳动或美学劳动成果的表达。^[8]”在这里,“writings”已经转变为所有思想观念的表达形式了^[9]。

不过,版权保护的仅是作品之表达,而不能延及此种表达所传递的思想观念^③。这意味着作者不能将自己作品中所体现的思想观念据为己有。对于同样的思想观念,他人可以自由使用,也可以进行独创性的再表达。这就是版权法上的“思想/表达”二分法原则。通过这一方式,版权法构建起一个大的治理框架,即将作者自己创造的作品与前人积累的公有知识区分开来,明确了哪些可为版权人独占,哪些则为他人可自由使用的范围^[10]。

根据版权法,有几类是几乎完全以语言文字为基础的作品。

(1)文字作品,主要指以书面语言形式作为表达工具的作品,包含小说、诗歌、散文、论文等以文字形式表现的作品。此外,还包括计算机软件。

(2)口述作品,指作者以口头语言表述形成的作品,如即兴演讲、老师授课、法庭辩论等。

(3)音乐作品,指歌曲、交响乐等能够演唱或者演奏的带词或者不带词的作品。音乐作品可分为两部分,曲和词。其中文字作品主要指歌词部分。

(4)戏剧作品,指话剧、歌剧、地方戏等供舞台演出的作品。其中文字作品主要指剧本。

(5)曲艺作品,指相声、快书、评书等以说唱为主要形式表演的作品。

(6)书法作品,指一种带有美感的、传统的中国文字表现形式。在版权法中,书法作品是作为美术作品加以保护的。

(7)演绎与汇编作品。演绎作品,是对已有的作品进行改编、翻译、注释、整理并由此产生的作品。汇编作品,指的是将多个作品或作品片段汇集在一起的作品。语言产业中的数字图书馆、语料库、语言博物馆等,基本上都涉及汇编问题。

(二)切片式治理手段:语言作品形态多维化与立体化

语言符号具有信息传递的共享性特征,摆脱

③ 版权法上的思想,指的是想法、概念、原则、客观事实、创意、发明和发现、程序、工艺和方法等。版权法中的表达,则是指对于上述思想观念的各种形式或方式的表达,如文字的、音符的、数字的、线条的、色彩的、造型的、形体动作的表述或传达等。WTO“与贸易有关的知识产权协议”(TRIPs协议)第9条第2款规定,版权保护仅延及表达,而不延及思想、工艺、操作方法或数学概念之类。

了有形物质形体的束缚,使得语言作品中的版权因此可以一种切片的方式来治理,创设不同的权利形态,以“权利束”形式形成多维化产品形态与产业链条。《著作权法》第 10 条第(1)―(4)项是著作权中的人身权,第(5)―(17)项是著作权中的财产权。通过这种密集的权项设计,以最初的语言文字作品为起点,可以切割、衍生出不同的语言产品。如通过复制权,可以形成语言产品不同形态的产品,如录音录像制品、拓印制品、数字化产品、字体字库作品、形象化角色与周边产品等;借助翻译权,可以将语言作品区分成不同语种、国别与地域的产品;通过改编权,可以将语言作品转化为戏剧、影视剧本等产品;通过汇编权,覆盖着数据库产品、语料库产品、语言作品展览馆与博物馆产品、各种古籍点校产品等;通过表演权、表演者权、广播权、放映权等,可以形成各种表演作品,如相声、小品、曲艺、评书等产品。

语言产品也借助版权的不同权项,控制着语言产业链条的长短。例如,复制权和发行权,可以控制作品从纸面平面印刷形态到有声读物产品,控制语言产品的印刷、出版、新闻与发行等渠道;广播权、网络信息传播权,可以控制语言作品有线、无线及网络在线广播产业;通过改编权与摄制权,语言产品可以延伸及改编产品、三维形象角色以及各色玩偶玩具、形象角色等周边衍生产品,控制编剧产业与影视制作等产业;通过表演权与表演者权,语言作品可以控制相声、小品、曲艺、评书等表演产业,以及配音、播音等产业;文字书法等作品可以通过纳入美术作品范畴,以复制权的形式控制计算机字体字库产业、广播传媒产业以及影视摄制等产业;通过教材、教辅、试题试卷等语言文字作品的版权,可以控制语言能力培训、测试

以及语言能力康复等产业。通过这些不同的权能与权项,语言作品产生着“种子功能”,在动漫、手游、话剧、玩具,甚至主题公园之间转化,衍生出包括文学、动漫、影视等不同领域的产品,形成越来越长的丰富多样的产业链条,构建起立体化的语言产品生态体系。

(三)司法主导治理:语言作品权利的稳定预期

与专利和商标领域不同的是,在版权领域,采用的是自动获得的原则,即尽管存在向版权局进行登记的版权登记制度,但是作品从创作之日起即可自行获得,不需要经过一个前置的行政审批过程。在遭遇侵权诉讼时,由法官先确定作品是否符合版权法条件以及权利归属,然后再确定是否存在侵权与否的问题。可以说,在语言作品版权领域中,构建起来的实际上是一种司法裁判为主导的治理机制。美国尼科斯(Nichols)案是一个最初用于文字作品的案例,汉德法官发展出来的判断侵权的“抽象层次法”^④。后来的“阿尔泰”案将其扩展到计算机程序侵权判定,进一步归纳为一种“抽象—过滤—比较”三步走方法^⑤。因此,在现代版权制度中,法官大多通过这样一种“抽象—比较”的方式来实现治理目标。由于将两项任务集于一身,在这些判定的不同环节里,存在较大的主观歧义,难以取得统一的认识。尤其是在语言文字作品中,这种判定更容易受到对文意理解的主观因素的干扰,从而导致权利缺乏稳定性预期。

在 2018 年北京菲林律师事务所诉百度网讯纠纷案中,法院判定计算机软件生成的涉案文章内容不构成作品^⑥。但在同年的另一个案例中,审理法院却认为机器生成文章无论从外在表现形式与生成过程看,都源于创作者个性化的选择与安排,

④ 抽象(abstraction)就是将被侵权作品的独创性部分一步步分解成各个抽象层次的组成部分,包括情节、主题、对话、氛围、节奏、场景、顺序、角色等。过滤(filtration)就是将受保护的表达从不受保护的部分中分离处理,将最终可获版权保护的核心内容保留下来。比较(comparison),就是将受版权保护的核心内容与被控侵权作品进行比较。这包括两种情形:一是“字面相似性”(literal similarity),如被告原封不动地复制了原告的作品,或在文字、音乐、戏剧等作品中逐字逐句地抄袭了原告作品。另一是“非字面相似性”(nonliteral similarity),被告不是原封不动地抄袭原告作品,而是对原告作品改动性使用。Nichols v. Universal Pictures Corp., 45 F.2d 119, 7 USPQ 84(2d Cir. 1930).

⑤ Computer Association International, Inc. v. Altai, Inc., 982 F.2d 693(2d Cir. 1992).

⑥ 北京菲林律师事务所诉北京百度网讯科技有限公司著作权侵权纠纷一案民事判决书,京 0491 民初 239 号(2018)。

属于我国著作权法所保护的文字作品^⑦。在“温瑞安诉北京玩蟹公司”案^⑧与“武侠 Q 传游戏”案^⑨中,法院均主张,整体情节、主要人物、主要的线索脉络等构成作品的基本框架,这些主体框架之上进行的与原语言文字表达相同方式的再创作基本上都可以归属到改编权的范畴。但在“胭脂扣”一案件中,尽管小说《胭脂扣》与剧本《胭脂盒》从情节、角色、主线等来看,都大体相似^⑩,但主审法院却认为,两部作品存在显著区别,构成不同的故事,不构成侵犯改编权。在我国版权司法实践中,还有很多类似涉及语言文字作品的案例。案件事实基本相同,得出的结论却截然不同。这些都使得语言作品的版权保护陷入了严峻的稳定性危机。2011 年,全国版权登记的数量仅为 10 万件,2021 年,这一数字急剧上升到 600 多万件。这些都表明了权利的稳定性预期是语言作品版权司法治理中急需解决的头等问题。

三、科技赋能:语言技术中的专利治理

专利与语言文字有着不解之缘。从 19 世纪后半期起,专利就演变成一种纯粹的“语言描述”的权利。专利的客体范围、创造性高度、权利大小等都是依靠“语言描述”来确定。专利权侵权的判定也基本上围绕语言文字的解释来展开。在一定程度上可以说,专利的事业就是语言的事业。语言技术中的专利治理表现为以专门化的审查基准筛选可以获得专利的语言技术,并以法律诉讼作为保障,调整权利要求范围大小、保护强度高、保护链条长短,以此构建起语言技术创新领域中的竞争秩序,引导专利布局,控制语言技术更迭的方向、速度与频率,进而影响语言产品的形态与语言产业的形态。通过对专利权利要求中的文字表达与语言技巧的解读,语言资源的范畴借此得到再次拓展,开始由语言文字转化为一个个征服自然

的技术方案,并借助附着其上的专利权改变着社会的结构与产业结构。

(一)专利治理门槛:语言技术专利审查基准

语言技术是指运用或处理语言的技术,除软性的语言软件、语言教学法等外,还包括硬性的语言设备或语言工具。专利审查是语言技术进入专利治理体系的“入口端”,审查基准作为这个端口的过滤装置,关系到哪些语言技术以及多少语言技术可以进入这一治理格局。语言领域中的技术具有自己特殊性,主要包括语言信号传递、信息处理、语音识别与转化等,基本上是依靠使用计算机进行数据存贮、信息处理的技术。同时,这个领域中的技术很多时候表现出一种类似于智力活动的规则和方法,在申请专利时可能被认定为我国专利法第 25 条第 1 款第(二)项规定的对智力活动规则和方法不授予专利权的情形。因此,对语言技术,首先需要判定哪些属于专利权保护的客体,其次应当判定其所具有的创造性。为此,我国专利局《专利审查指南》(2020 年)第二部分第九章对关于涉及计算机程序的发明专利申请审查作出了若干规定,其中对语言处理计算机程序、汉字编码方法及计算机汉字输入方法等的审查进行了专门规定。

新审查指南指出,涉及人工智能、“互联网+”、大数据等发明专利申请,权利要求中往往包含算法、商业规则和方法等智力活动的规则和方法特征,在专利审查中,不应当简单割裂技术特征与算法特征或商业规则和方法特征,而应将权利要求记载的所有内容作为一个整体考虑。如果直接忽略这些特征或者将其与技术特征机械割裂,则无法客观评价发明的实质贡献,不利于保护真正的发明创造。因此,如语言技术专利申请中权利要求涉及的是抽象算法或单纯的商业规则和方法,

⑦ 深圳市腾讯计算机系统有限公司与上海盈讯科技有限公司著作权权属、侵权纠纷、商业贿赂不正当竞争纠纷案,粤 0305 民初 14010 号(2019)。

⑧ 温瑞安与北京玩蟹科技有限公司侵害作品改编权纠纷一审民事判决书,海民(知)初字第 32202 号(2015)。

⑨ 明河社出版有限公司、完美世界(北京)软件有限公司与北京火谷网络科技有限公司、昆仑乐享网络技术有限公司、昆仑万维科技股份有限公司侵害改编权及不正当竞争纠纷案,京民终 226 号(2018)。

⑩ 上海华某文化艺术有限公司与上海某剧院、罗怀某等侵害作品改编权纠纷二审民事判决书,沪一中民五(知)终字第 112 号(2012)。

且不包含任何技术特征,则该项权利要求属于智力活动的规则和方法,不应被授予专利权。但如果权利要求中包含技术特征,该权利要求就整体而言并不是一种智力活动的规则和方法,则不应依据专利法第 25 条第 1 款第(二)项排除其专利权的可能性。根据新《专利审查指南》(2020 年)的这些规定,涉及语言处理的技术属于可以获得专利保护的客体^①。

(二)专利治理高度:语言专利创造性的判定

语言技术领域的专利申请,多依附于计算机程序实现,既可以保护技术本身的方法,也可以保护实现其方法的装置及系统。获得专利的要求是要满足专利的实质性条件,即“三性”,新颖性、创造性与实用性。在这三性当中,创造性的判定是最复杂也是最关键的问题,只有满足达到一定创造性高度的技术才有可能获得专利。专利创造性的判断通常使用一种“三步法”的判断方法,包括以下 3 个步骤:①确定最接近的现有技术;②确定发明的区别特征和发明实际解决的技术问题;③判断要求保护的发明对本领域的技术人员来说是否显而易见。

对于语言技术专利申请的创造性判断,也适用这一判断方法。语言专利创造性的判定,常因语言技术的特殊性而表现出一些独特的特点。语言技术专利申请常见的特点是用功能性限定的方法权利要求,方法权利要求经常包括多个步骤,包含的技术特征也比较多。此外,语言技术中,为得到较好的技术效果,往往需要设置阈值或数值范围来对结果进行限定,不同的阈值或不同的取值范围会对处理结果产生不同的效果。但是,对于语言领域的技术来说,很少能够寻找到与专利申请中的数值或数值范围部分完全相同或部分重叠的在先技术特征。所以,在对具有数值或数值范围限定的权利要求进行创造性判定的时候,要根据说明书中的内容和本领域的公知常识来判断,如果该专利要求的数值或数值范围确实取得了预料不到的技术效果,则可认为是具有创造性。如

不能,则不具有创造性^[11]。

(三)专利治理表征:语言产业中的专利格局

1. 语言产业中的专利布局

现代竞争中,专利布局不仅可以看作是创新主体对专利制度的反馈,更可以看作是产业发展的生命线。现代语言产业在计算机字库、语料库、输入法等领域中,都存在着激烈的专利争夺^[12]。以字库产业为例,从中国大陆、香港、台湾与日本中文字库的分布情况来看,90% 以上的中文数字键盘汉字输入法技术专利被外国厂商所垄断,每年仅此一项要支付数亿元的专利使用费^[13]。从自然语言处理技术全球专利分布态势来看,截止到 2018 年 8 月,据 Patentics 全球专利库中检索到的自然语言处理的 4 万件左右的专利申请总体趋势,自然语言处理技术从 1983 年左右起步。1985 年申请量开始快速增长,2000 年当年申请量达到 800 件左右。2008 年至今,深度学习开始作为自然语言处理研究的主要技术,自然语言处理专利迎来新一轮增长。从自然语言处理技术全球专利申请态势来看,知识图谱占比最高,占 26%;其次是词法分析、句法分析和语言模型,占比为 19%;语义分析的占比最低,为 17%。从自然语言处理应用技术全球专利申请态势来看,机器翻译占比最高,达 32%;其次为自动问答,占 28%;再次为信息抽取(25%)、情感分析(9%),自动文摘最低^[14]。

通过对美国 1999—2018 年间自然语言处理技术专利数据统计分析可以看出,美国自然语言处理专利申请呈波浪式上升趋势。2012—2016 年间,专利申请快速增长,超过年 3 000 件大关,此后一直保持高位,逐渐进入技术成熟期。从授权状况看,1999—2011 年间,专利授权量保持在 1 000 ~ 2 000 件范围间。随着网络技术与人工智能的发展,2012—2016 年间,专利授权量超过 2 000 件,增势迅猛。2017 年后进入“3 000 +”时代。从技术方向看,发明专利主要集中在 10 个 IPC 小类。其中,G06F(电数字数据处理)和 G10L(语音分析识别)是专利授权前两位,二者占到样本专利总数

^① 参见国家知识产权局专利局《专利审查指南》(2020)第 6.2 节审查示例 10。

62%左右,是创新热度最高、发展速度最快的领域。第二梯队的是H04M(电话通信)、H04L(数字信息的创制)等小类,授权专利数量从2 000至5 000件不等。第三梯队是H04W(无线通信网络)、G06Q(特殊目的的数据处理系统或方法)等小类,授权专利数量在1 000件左右。从诉讼情况来看,美国语言领域专利诉讼在2014年左右达到一个高峰后,形成转折点,之后回落。这些反映了语言产业主要竞争者开始合并,出现行业巨头独占的局面^[15]。

2. 语言产业中的专利讼争

语言产业中的专利诉讼争端一向比较激烈。专利诉讼表面上是关于技术的法律纠纷,背后代表的则是语言产业中巨大的市场分歧,通过专利的法律诉讼来控制产业链条,重构产业结构,调整产业布局。以汉字输入法中的专利争夺为例,随着移动互联时代的到来,输入法的战略意义开始由“输入工具”上升到了“入口生态”,成为互联网企业争夺的重地。从上世纪末的北京王码公司的五笔字型专利纠纷到近年的搜狗与百度输入法专利之战,都显示了汉字输入法中专利之争的惨烈性。今天国内的输入法市场格局基本上呈现出三大阵营:搜狗与百度输入法处于第一阵营;腾讯QQ输入法和讯飞输入法处于第二阵营;谷歌输入法等则处于第三阵营。由于行业几方阵营产品属于同类型技术竞争,在专利布局方面也相对厚实,因此围绕市场发动以专利战为主要形式的激烈争夺。2009年,搜狗公司起诉腾讯等公司破坏“搜狗拼音输入法”运营^⑫。这被称为“中国输入法第一案”。2014年,百度与搜狗之间展开了一场号称中国“互联网专利第一案”的手机输入法专利大战。

2015年,搜狗起诉百度侵犯其17件输入法专利并索赔2.6亿,其中包括1亿赔偿的“一种向应用程序输入艺术字/图形的方法及系统”专利^⑬。搜狗与腾讯、百度之间的输入法专利大战,对输入法市场格局的影响巨大。

影响较大的输入法专利纠纷除了上面提到的以外,其他还有不少。如早期的“五笔字型专利侵权纠纷案”^⑭、“微软郑码输入法专利纠纷案”^⑮,以及最近的“小i机器人与苹果语音助手Siri专利大战”等。从语言技术专利诉讼来看,一般都会采用双线作战的策略。一方面,原告方以专利为诉因,向法院提起诉讼,要求被控方停止侵权,赔偿损失;另一方面,被告方会以自己专利也被对方侵害为由提起反诉,同时向国家专利复审机构提起原告专利的无效请求。最典型的就是小i机器人与苹果的专利纠纷^⑯。从这些关于语言处理技术专利诉讼争夺大战中可以看出,随着网络技术的发展,作为沟通人机交流,推动人工智能从感知智能向认知智能演进关键的语言处理技术,成为互联网巨头进行市场竞争和商业角逐的重要阵地。谷歌、微软等国际技术巨头都卷入其中,进行抢滩。尤其是在搜狗与百度这场围绕手机输入法旷日持久的专利大战中,从搜狗发起猛烈攻势到百度全力反击,其中共涉及近30项专利,发生了近50次的诉讼较量。在今天,在语言产业当中,专利布局已经成为进入这一领域的基本门槛与重要储备。

四、形象认知:语言符号中的商标治理

商标,是指在商业上使用的种种语言文字、图形符号的统称。这些语言文字、图形符号的主要功能是将某一厂商所提供的商品或服务与其他厂

^⑫ 北京搜狗信息服务有限公司、北京搜狗科技发展有限公司与深圳市腾讯计算机系统有限公司、北京奥蓝德信息科技有限公司等不正当竞争纠纷一案民事判决书,北京市第二中级人民法院,二中民初字第12482号,2009。

^⑬ 参见百度在线网络技术(北京)有限公司、百度时代网络技术(北京)有限公司与北京搜狗信息服务有限公司、北京搜狗科技发展有限公司侵害发明专利权纠纷一案民事判决书,北京知识产权法院,京73民初字第1073号(2016)等案件判决书。

^⑭ 北京王码电脑总公司与中国东南技术贸易总公司专利权权属、侵权纠纷二审民事判决书,北京市高级人民法院,高经知终字第30号,1994。

^⑮ 北京中易中标电子信息技术有限公司与微软公司、微软(中国)有限公司专利权权属纠纷一案民事判决书,北京市第一中级人民法院,一中民初字第5363号,2007。

^⑯ 苹果电脑贸易(上海)有限公司与国家知识产权局专利复审委员会不履行法定职责、行政裁决二审行政判决书,北京市高级人民法院,高行(知)终字第2935号,2014。

商所提供的商品或服务区别开来。语言文字符号借助商标的保护,开始转化为代表质量、信用、地位与身份的象征,并规范着商品交易的秩序,改变着市场竞争的结构。商标制度将语言文字符号的重要性提升到了一个前所未有的高度,语言资源的范畴借助商标的保护再一次得到了拓展。从本质上可以说,对商标的治理就是围绕语言、文字与符号展开的对市场竞争秩序的治理。

(一)从混淆到淡化:语言文字商标治理的规范转向

随着现代商业广告运动的兴起,商标中的语言符号、形象标记由于简明易记,方便消费者认知和确认商品的来源,成为广告中的必要要素。大量成功广告的宣传下,消费者产生对一定商标商品的心理偏好,商标的知名度和消费者对商标的忠诚度都大幅提升。在日益竞争激烈的市场上,那些获得知名度的商品可获得比一般利润更高的超额利润,品牌的塑造上升到与产品品质的提升同样重要的地位。商标中的语言符号、形象标记已不再是一个个简单的字词,而是代表背后商家长期积累形成的较高声望与商誉,并以此塑造起人们对于商品来源与品质以及商家服务质量的总体性形象认知。这类语言文字与符号标记的商业价值得到了大幅的拓展,成为厂商至关重要的无形财产。现代市场经济因此也发展到以品牌为中心的符号经济,即以语言文字和符号标记的商标作为载体,统筹整合各类科技创新资源、经济、社会和文化等要素,牵引经济集约化发展,展开市场竞争的一种经济形态。

当消费者认可某一商标所标示的商品或服务时,就产生了商标所代表的商誉。这就是商标所有人的财产。为了赢得消费者的青睐,商标所有人必须在相关的商品或者服务商进行投资,提高商品或者服务的质量、性能等。现代商标保护的基础在于维护商标识别的显著性,也即商标中语言文字、符号标记将商标使用者的商品或服务与他人商品或服务进行区别开来的能力。因此,商标保护的基础并非完全关注于表明商品的来源,确保对消费者的混淆、错误或受欺诈的可能性,更

要兼顾那种将商标与特定商标之间紧密联系的品质保证与商品推销能力,即尽管在非相关商品上使用,也将使该商标在公众心目中的特定意义,包括品质保证和说服力度,慢慢减弱或消散^[16]。为了保持商标的显著性,商标权对语言符号、形象标记的保护开始从制止混淆转向防止淡化,即防止由一个具有显著识别性的商业标识转化为一个不具有显著识别性的通用符号标记。

(二)从冲突到协调:特殊类型语言文字商标诉讼

文字商标是商标的一种主要类型,包括以各种文字、字母、数字组成,不含图形成分的商标。文字商标便于称呼、广告宣传。单词和短语,都属于文字商标。但这些文字、符号商标可能会存在着一些在先的其他权利,导致与商标权的冲突。商标治理的一个主要目标就是寻求商标权与其他各类冲突的权利之间的相互协调。与语言文字有关的商标诉讼情形包括商标权与姓名权、字体权、节目名称、字号、域名等之间的冲突。我国《民法典》第 1012 条的规定,“自然人享有姓名权,有权依法决定、使用、变更或许可他人使用自己的姓名”。我国商标法没有对姓名是否可以注册为商标予以明确规定。不过从实践来看,姓名作为一种标记符号,在符合商标法规定的条件下,也可注册为商标。尤其是名人的姓名,往往被公众理解为其所代表形象的信息,当与特定商业活动联系在一起的时候,可能产生较好的商品销售效果。《民法典》第 1017 条规定,具有一定社会知名度,被他人使用足以造成公众混淆的笔名、艺名、网名、字号、姓名和名称的简称等,参照适用姓名权和名称权保护的有关规定。因此,在未经他人许可,将他人姓名,尤其是名人的姓名,抢注为商标使用的情形下,常会出现姓名权与商标权之间的冲突。近年来影响较大的“乔丹”球鞋案就是典型的案例。

字库字体可以作为版权作品进行保护。因此,不同字库字体的文字有可能作为美术作品成为商标法第 32 条上的在先权利,从而产生与商标权之间的冲突。近年来,电视节目名称与商标权

之间的纠纷也多有发生,最典型的案例就是《非诚勿扰》商标侵权案。在这一案例中,涉及到电视节目名称是否属于商标性使用的问题。字号与商标之间也存在冲突。字号作为重要的商业标识之一,其作用类似于商标。由于国内现行法律法规对字号的法律保护比较繁杂、混乱,现行民法、反不正当竞争法、企业名称登记管理规定等法律法规都有涉及,并采用企业名称“区域登记注册”与商标“全国统一注册、分级管理”并行的体制,导致出现企业字号与商标权之间的法律冲突^[17]。此外,随着电子商务进入高速发展阶段,域名也逐渐发展成为网络经营活动中的重要商业符号。这样的标识作用给传统商标权的保护带来了挑战,使得传统实体经济的商标与网络域名之间的冲突纠纷开始凸显^[18]。域名与商标之间的冲突主要表现在恶意抢注的行为,包括故意将他人的驰名商标、著名商标、商号等,抢先注册为域名。

(三)从广告到专名:语言符号的特殊标志管理

现代社会,商标逐渐成为商品广告和宣传的代言人。通过广告,能大力提升商标的品质保证和销售说服功能,在消费者心中树立某种印象和地位,使商标变成一个有意义的带有高附加价值的符码。不过,由于广告的社会传播效能,因此各国对于广告活动中的语言文字与措辞用语都有严格的规定。2015年,我国《广告法》对广告中的措辞用语,以及禁止使用的语言文字作了详细规定,同时对广告使用数据、统计资料、调查结果、文摘、引用语等引证内容的,应当真实、准确,并表明出处。《广告法》明确规定,广告中的语言用语不得做贬损性的宣传,贬低其他生产经营者的商品或者服务。此外,新广告法还对药品广告、保健品广告、农药兽药广告、交通工具中的广告、烟草制品广告、酒类广告、教育培训广告、招商广告、房地产广告、农作物种子广告等特殊广告的用语都进行了非常详尽的规范。

此外,在市场竞争活动中,还有许多与商标相似的其他语言文字与符号标记。这些语言文字与符号标记大多与商业经济有着联系,能够带来一

定的经济价值,有的甚至特别巨大,如奥运会会标等。但是这些语言文字与符号标记又不同于使用在商品上的商标。当这些语言文字与符号标记使用在商品上,或者与商标同附在一起时,常常容易与商标产生混淆,因此国家对于这一类特殊的语言文字与符号标记进行了专门的管理规定。

(1)对特殊标志使用进行规范的《特殊标志管理条例》。

(2)对奥林匹克标志进行保护的《奥林匹克标志保护条例》、对世博会标志进行保护的《世界博览会标志保护条例》、对进博会标志进行保护的《关于规范使用中国国际进口博览会标志的规定》等。

(3)对个人姓名使用进行规范的《户口登记条例》,以及征求意见但尚未正式颁布实施的《姓名登记条例》。

(4)对企业名称使用进行规范的《企业名称登记管理规定》,以及在征求意见中的《企业名称登记管理条例》。

(5)对地名使用进行规范的《地名管理条例》,以及在征求意见中的新修订的《地名管理条例》。此外,2019年国家知识产权局制定了《地理标志专用标志使用管理办法(试行)》,对地理标志专用标志的使用进行了统一和规范。

(6)对域名进行规范的《中国互联网络域名管理办法》《互联网域名管理办法》等。

五、结语

在现代社会,语言不仅是人类重要的文化资源,同时也是一种稀缺的经济资源。知识产权作为保护语言成果最重要的产权制度,不仅规范着语言产品交易的市场秩序,同时还塑造着语言经济的结构与方向,更重要的是,在三重意义上拓展着语言资源的价值范畴。为了确保语言经济的良性发展,我们要进一步完善语言产业中的知识产权法律体系、提高知识产权治理效能。要深入研究语言知识产权保护的规律,及时跟踪、研究语言文字产业中的新情况、新问题,根据实际需要和研究成果,进一步完善语言行业知识产权立法。要进一步加强监管,建立语言行业知识产权巡查制

度,构建线上线下两分,“获权—预警—监测—维权”四维的知识产权监管体系。要进一步优化服务,培育和发展语言行业市场化知识产权信息服务,提供语言行业优质的知识产权公共服务平台,满足社会不同层次对语言行业知识产权信息的需求。要进一步提升意识,既支持语言行业平台合法合理维权,又要以规则意识为支撑,不得滥用权利。国家主管部门需要厘定清晰的规则,有效化解语言平台与使用者之间的张力,促进语言经济的健康良性发展。

参考文献:

- [1]张卫国. 语言经济与语言经济学:差异与互补[J]. 学术月刊, 2015(3):93.
- [2]国家语言文字工作委员会组. 中国语言文字事业发展报告(2017)[R]. 北京:商务印书馆,2017:123.
- [3]DRATLER J. 知识产权许可(上)[M]. 王春燕,等译. 北京:清华大学出版社,2003:30.
- [4]赵世举. 语言与国家[M]. 北京:商务印书馆,2015:84.
- [5]WIPO. IP Facts and Figures[Z]. 2019.
- [6]福柯. 知识考古学[M]. 谢强,等译. 北京:生活·读书·新知三联书店,1997:23.
- [7]海德格尔. 林中路[M]. 孙周兴,译. 上海:上海译文出版社,2004:5.
- [8]GOLEDSTEIN V. California, 412 U. S. 546, 178 USPQ 129[Z]. 1973.
- [9]李明德. 美国知识产权法[M]. 2版. 北京:法律出版社,2014:230-231.
- [10]ROBERT P M, et al. Intellectual property in the new technological age [M]. 2nd ed. New York: Aspen Law & Business, 2000:379-380.
- [11]陈红红等. 浅谈语音领域的创造性审查[J]. 电声技术,2012(S1):45.
- [12]张伟. 语言文字:信息化时代的国家“硬实力”——专访国家语委副主任、教育部语言文字信息管理司司长李宇明[J]. 中国经济周刊, 2009(31):17.
- [13]贺宏志主编. 语言产业导论[M]. 北京:首都师范大学出版社,2012:4.
- [14]国家知识产权局学术委员会. 产业专利分析报告(第68册)——人工智能关键技术[M]. 北京:知识产权出版社,2019.
- [15]刘媛. 美国自然语言处理技术专利情报分析及启示——基于1999—2018年专利数据[J]. 科技管理研究, 2020(6):202-207.
- [16]FRANK I S. The rational basis of trademark protection [J]. 40 harvard law review, 1927(813):825.
- [17]谢冬伟. 商标权与企业名称的冲突及其解决[J]. 中华商标,2005(6):53.
- [18]徐璐. 互联网域名与商标的协调保护研究[J]. 武陵学刊,2020(3):69.

(本文责编:王延芳)