

爬虫技术应用合法性引争议

保障数据安全亟待规制非法爬取数据行为

□ 本报记者 张维
□ 法制网见习记者 邢国涵

随着社会经济快速发展,数据的价值日益凸显,已然成为企业科技创新的必备要素。但企业通过技术手段获取数据时,数据抓取技术的应用行为是否合理合法,是一个值得深思的问题。

近年来,网络爬虫“爬取数据”成为热词,相关司法案例不断出现。据不完全统计,近些年涉及网络爬虫的司法案件达十余起,其中既包括民事案件,还包括刑事案件。这类案例甚至还有愈演愈烈之势。

在近日于上海举行的长三角数据合规论坛(第三期)暨数据爬虫的法律规制研讨会上,上海市人民检察院研究室副主任陈超然透露,检察机关正在积极推动企业合规改革试点工作,数据合规正是其中重点。“目前爬虫爬取数据案件非常普遍,当网络平台或者个人通过技术手段抓取别的平台数据时,这种行为是否合法,平台数据主体是谁,归属使用,值得深入研讨。”

杭州长三角大数据研究院副院长郭兵认为,数据爬虫作为中立性的技术,已在互联网产业领域得到广泛应用。需要注意的是,如果爬虫技术不当应用,会对其他竞争者的合法权益造成损害,甚至涉嫌违法或者犯罪,也将对产业的健康发展产生非常大的负面影响。

爬取数据涉嫌侵权

从技术角度看,爬虫是通过程序去模拟人类上网或者浏览网页、App行为,让其高效地在网上抓取爬虫制造者所需要的信息。

欧莱雅中国数字化负责人刘煜晨说,大多数网站拒绝爬虫访问,其中的原因既包括商业利益考量,也包括自身网站运营安全的考量。除了爬虫可能爬到网站不愿被爬取的数据以外,网站经营者往往还会担心爬虫干扰网站正常运营。

而非正规爬虫自动持续且高频地对被爬取方进行访问,服务器负载飙升,也会给服务器带来“难以承受”之重:应对经验不足的网站,尤其是中小网站可能会面临网站打不开,网页加载极其缓慢,有时甚至直接瘫痪的情况。

新浪集团诉讼总监张岳说,无论是爬虫还是实现其他目的的技术,就其本身而言,都是中立的,但爬虫技术的应用不是中立的,技术应用

核心阅读

在大数据时代的背景下,愈来愈多的市场主体投入巨资收集、整理和挖掘信息。如果任由网络爬虫任意使用他人通过巨大投入获取的数据资源,将不利于鼓励商业投入、产业创新和诚信经营,甚至可能直接违背了数据来源用户的意愿和知情权,最终势必损害健康的竞争机制。

都带有应用者的目的。这时候不应该评价技术原理,而是需要评价技术用来干什么,这一行为手段是否具有正当性。

提及网络爬虫,robots协议是绕不开的话题。robots协议(也称爬虫协议)的全称是“网络爬虫排除标准”,网站通过robots协议明确警示搜索引擎哪些页面可以爬取,哪些页面不能爬取。该协议也被行业称为搜索领域的“君子协定”。

刘煜晨说,当网络爬虫访问一个网站,robots协议像立在自己房间门口的一个牌子,告诉外来者谁可以过来,谁不可以过来。但是,这只是个君子协议,只能起到告示作用,起不到技术防范作用。

实践中,恶意爬虫爬取时不遵守网站的robots协议,并可能爬取到不该爬的数据,这种情形并非孤例。小红书法务负责人曾翔说,恶意爬虫案例经常发生在内容平台和电商平台。在内容上被爬取的更多是视频、图片、文字、网红互动数据,用户行为等,在电商领域则多为商家信息和商品信息。

“内容平台一般约定了相关内容知识产权归属发布者或者发布者和平台共同所有,这些爬虫没有签订协议就获得用户授权,涉嫌对知识产权人权利的侵犯。”曾翔说。

或应明确网站权利

这就涉及到数据的权属及是否能开放的问题。

上海市浦东新区人民法院知识产权庭法官徐弘认为,数据是内容产业的核心竞争资源,内容平台经过汇总分析处理后的数据往往具有极高的经济价值。

“如果要求内容平台经营者将其核心竞争资源向竞争对手无限开放,不仅有违‘互联互

通’精神的实质,也不利于优质内容的不断更迭和互联网产业的持续发展。”徐弘说。

恶意爬虫爬取数据案件频发的背后,是数据的价值增加,以数据为核心的市场竞争愈发激烈。

华东政法大学教授高富平说,进入大数据时代,数据价值再次凸显,现在的爬虫技术已经从原来的网页爬虫进入到底层数据的爬取。数据爬虫问题会变得越来越严重。

在大数据时代的背景下,愈来愈多的市场主体投入巨资收集、整理和挖掘信息,业内人士对此表示担忧:如果任由网络爬虫任意使用或利用他人通过巨大投入所获取的数据资源,将不利于鼓励商业投入、产业创新和诚信经营,甚至可能直接违背了数据来源用户的意愿和知情权,最终势必损害健康的竞争机制。

高富平认为,如果网站合法积累数据资源,那么这些数据资源就应该属于网站的资产。“允许数据生产者,控制者基于商业目的开放数据是有好处的,通过许可使用,交换交易等方式,可以让更多人享用数据服务。期待在未来确认数据所有合法生产者对数据的控制权、使用权。”

有序流转同等重要

目前,网站虽然可以指定相应策略或技术手段,防止爬虫抓取数据,但爬虫也有更多技术手段来反制这种爬取策略。

刘煜晨说,反爬和爬取的技术一直在迭代,在技术领域,没有爬不了的网站和App,只有愿不愿意爬和有多难爬的问题。

据了解,现实中恶意网络爬虫制造者抗辩时,往往将robots协议限制爬取与数据流转联系起来。徐弘认为,在“互联互通”的语境下,“有序”和“流转”同等重要,缺一不可,需排

除假借“互联互通”妨碍公平竞争,危害用户数据安全的行为。

“对于非搜索引擎爬虫的正当性判断,要考虑是否足以保障用户数据的安全性。包括身份数据,行为数据等在内的用户数据,从属性来讲不仅仅是经营者的竞争资源,同样具有用户的个人隐私属性,而此类数据的集合更涉及社会公共利益。”徐弘说。

据了解,近年来有关数据安全法律规范正在不断完善中。数据安全法作为数据安全的基本法,承载着解决我国数据安全核心制度框架的重要任务。此外,还有2019年通过的密码法,工信部拟出台《工业和信息领域数据安全管理办法(试行)》等,一些地方如深圳、上海等也在探索制定数据管理相关规范。



后疫情时代传统供应链亟待转型 专家建议

以内部供应链数字化对冲外部成本上升压力

□ 本报记者 张维

供应链的重要性在后疫情时代越发凸显。

去年以来,在包括逆全球化、贸易保护主义、新冠肺炎疫情等在内的多种因素影响下,全球供应链断供频发,风险陡增。提升产业链供应链的韧性和抗风险能力,已成为世界各国的重要命题。

中国同样要面对这一问题,但优势也很明显。在近日举行的2021全球智能物流峰会(GSSC)上,北京大学国家发展研究院名誉院长林毅夫指出:“我们拥有全世界供应链最完整的实体经济,在传统产业上有后来者优势,还可以利用引进消化吸收作为再创新的来源。”

林毅夫坚信,从2036年到2050年,中国应该还有每年6%的增长潜力,而且可以实现4%左右的增长。在新发展格局之下,需要企业家抓住中国后来者优势与换道超车的优势。

发展参差不齐

2020年,中国成为全球唯一GDP正增长的大国,在全球十大经济体中独树一帜。中欧国际工商学院京东运营及供应链管理学教授赵先德认为,其中一个重要原因就在于中国“在全球供应链有很好的掌控能力”。在疫情初期即使人们足不出户,仍然可以享受到货上门的服务,就是一个有力的佐证。“这种供应链能力使得我们的GDP能够增长。”

供应链的重要性也早为顶层设计所注意到。2017年10月发布的《国务院办公厅关于积极推进供应链创新与应用的指导意见》(以下简称《指导意见》),认为供应链具有三大重要意义:落实新发展理念的重要举措;供给侧结构性改革的重要抓手;引领全球化提升竞争力的重要载体。

《指导意见》提出,到2020年,形成一批适合我国国情的供应链发展新技术和新模式,基本形成覆盖我国重点产业的智慧供应链体系。供应链在促进降本增效、供需匹配和产业升级中的作用显著增强,成为供给侧结构性

改革的重要支撑。培育100家左右的全球供应链领先企业,重点产业的供应链竞争力进入世界前列,中国成为全球供应链创新与应用的重要中心。

在赵先德看来,目前供应链管理在不同公司中处于不同发展阶段,“我们根据过去这些年的研究发现,中国的供应链管理可以分成四个阶段,有一些企业还处于被动执行的第一阶段,有一些企业则已经进入了第四阶段”。

赵先德所说的“第一阶段”,是指改革开放初期,很多行业都是供不应求,在这种情况下企业选择了生产单一标准化的产品,通过规模效益降低成本,供应链不会参与其中。

到了第二阶段,客户需求多样化,负责供应链管理的人员开始参与竞争战略的制定,研究如何根据不同产品和细分市场的特征选择合适的供应链能力,并且通过供应链的网络流程以及资源的优化来建立起这样的能力,供应链从被动执行的角色成为企业竞争优势来源。

到了第三阶段,即在数字化时代,越来越多的用户强调体验,竞争的逻辑开始从产品主导逻辑转向服务主导逻辑。供应链在这一阶段要通过定制化的产品和服务组合来提高用户的体验,用户所判断的价值。这一阶段可称为“智慧和体验型的供应链”。

在第三个阶段的基础上,京东等一些供应链做得好的企业,基于供应链的能力打造生态圈,在生态圈里打造合作共赢的一个环境,通过合作机制的设计以及数字化工具的使用,吸引众多有不同资源和能力的组织参与进来,这一阶段也是第四阶段,可称为“供应链+生态圈”阶段。

再造生产模式

显然,数字化已经成为供应链当下发展避不开的关键词。

中国贸促会研究院副院长赵萍认为,数字化供应链有三个问题特别值得关注:第一,数字化供应链的核心技术必然是基于数字技术和智能化的技术,所以数字化与智能化是

数字供应链发展的核心技术支持。

第二,数字化供应链并不是引入一个数字供应链的系统就够了,上下游的应用场景配合至关重要。首先,从服务主体来说,如果企业自身没有实现数字化转型,数字化供应链很难落到实处。企业要从思维方式到具体的操作系统再到执行层面,都进行数字化转型。其次,从服务对象来说,包括商品库存、销售相关数据等能否对应用到每个SKU(Stock Keeping Unit,是库存保有单位)都要实现数字化,这是数字化供应链发展的一个重要前提。

第三,数字化供应链不仅仅是一个链条,它还为整个实体经济带来生产模式的再造。数字化供应链基于真实的海量数据研究需求,将原有大规模生产销售模式改造成以需求为中心的模式,以实现对市场更高柔性且更快速度的反应。

很多企业对于数字化供应链都在进行尝试和探索,与会的多个企业代表分享了他们的观点与经验。沃尔沃汽车亚太区执行副总裁兼法务副总裁胡彦认为,数字化是企业构建智能服务新生态的关键,是企业发展的必由之路。

作为传统制造业企业的安利,过去面对的是营销伙伴、经销商,“我们不知道我们的产品真正的消费者是谁?但是大数据打通了这个壁垒,让企业能够知道我的产品最后被谁消费了。”安利(中国)供应链副总裁陈森林说。

赢彻科技创始人和CEO马喆人对通过自动驾驶技术产生的货运机器人则表现出浓厚的兴趣,吸引众多不同资源和能力的组织参与进来,这一阶段也是第四阶段,可称为“供应链+生态圈”阶段。

俊富集团创始人和董事长赵民忠说,自己企业生产的是在个人护理领域的快速消费品,消费者关心的是品质好,成本低,交货快,选择多,这就需要做生产供应链的改造,“从供应链一体出发才能把这个问题解决掉”。京东就拿出的一体化供应链的方案与

实践。“随着物流与供应链产业逐渐进入以数字化引领的新时期,物流与供应链创新在构建现代经济体系中的作用日渐凸显。在这一特殊背景下,一体化供应链物流将成为企业高质量发展的增长点,数字化转型的着力点、端到端效率的新引擎,创造长久而深远的价值。”京东物流CEO余睿说,作为新兴的物流服务模式,一体化供应链物流服务已成为供应链“补链”“强链”的重要途径,既可以提供覆盖贯穿供应链战略到执行的解决方案,又可以提供从方案到网络到运营的一体化落地支撑。

对冲外部压力

的确,伴随消费、生产、流通的悄然变化,传统供应链亟待转型。

据余睿分析,新一代消费者个性化不断增强,倒逼品牌商快速响应需求。另外,伴随着消费端变化,流通渠道不断迭代,全渠道多场景销售发展加速,产品销售波动性大幅增加,这对品牌商的传统分销模式和库存管理方式造成了巨大的挑战。反馈到生产端,精益生产、柔性生产是一大转型难题,如何更好地平衡成本与效率,在竞争中保有优势,是品牌商、制造商所面临的挑战。

赵萍在调研供应链中发现,过去一年企业共同面对的挑战有三个:一是供应链能否稳定,有没有保证,这是对供应链最低的要求。二是很多企业愿意建设自己的供应链以实现数字化转型,但不知道与谁合作,找到一个能够提供从供应链设计、规划,执行的“一站式”服务的合作者并不容易。三是在企业接受一体化供应链物流服务的同时,如何在执行层面和这一系统有效对接。

赵萍特别提醒,外在的环境是我们所改变不了的,但是如何通过内部供应链的优化、数字化来有效降低成本,提高效率,以对冲外部成本上升的压力是很重要的。

“面向未来,企业之间的竞争取决于供应链的竞争。”赵萍说,这是因为供应链能有效帮助企业更好地应对市场的不确定性,挖掘内部降低成本的潜力,从而使企业在应对市场的千变万化中能立于不败之地。

学术谏言

□ 王春晖

于11月1日实施的《中华人民共和国个人信息保护法》,在国家层面建立了个人信息保护制度。为加强个人信息保护宣传教育,推动形成政府、企业、相关社会组织、公众共同参与个人信息保护的良好环境,确保个人信息保护法的各项要求和规定在社会生活中落到实处,建议在学习和贯彻实施个人信息保护法中重点把握好以下十大核心要点。

第一,“规范个人信息处理活动”是个人信息保护法的核心。个人信息保护法第一条规定:“为了保护个人信息权益,规范个人信息处理活动,促进个人信息合理利用,根据宪法,制定本法。”由此可见,就个人信息保护法的实质功能而言,它是一部个人信息处理活动行为规范法,“规范个人信息处理活动”处于整个个人信息保护法的核心地位,只有夯实这一关键环节,才能确保实现保护个人信息权益和促进个人信息合理利用之目的。

第二,个人信息的内涵与匿名化。个人信息保护法第四条第一款明确了“个人信息”的定义:“个人信息是以电子或者其他方式记录的与已识别或者可识别的自然人有关的各种信息,不包括匿名化处理后”的信息。”这意味着“个人信息”经匿名化处理后即不属于个人信息,也就无需适用个人信息保护法的相关规定,这一规定体现了个人信息保护法对个人信息“保护”和“利用”并重的立法精神。

第三,个人信息保护法的域外效力。根据我国个人信息保护法第三条第二款规定,在中华人民共和国境外处理中华人民共和国境内自然人个人信息的活动,有下列情形之一的,也适用本法:(一)以境内自然人提供产品或者服务为目的;(二)分析、评估境内自然人的行为;(三)法律、行政法规规定的其他方式。

第四,个人信息处理的核心原则。个人信息保护法主要确立以下五项重要原则:一是遵循合法、正当、必要和诚信原则;二是采取对个人权益影响最小的方式,限于实现处理目的的最小范围原则;三是处理个人信息应当遵循公开、透明原则;四是处理个人信息应当保证个人信息质量原则;五是采取必要措施确保个人信息安全原则等。需要注意的是,个人信息保护法总则部分第六条确立的两个“最小原则”,即“采取对个人权益影响最小的方式”“限于实现处理目的的最小范围”,是个人信息处理应遵循的核心原则,尤其是后者,是禁止“过度收集个人信息”的关键要点。

第五,以“告知—知情—同意”为核心的个人信息处理规则。个人信息保护法构建了以“告知—知情—同意”为核心的个人信息处理规则体系。个人信息处理者“告知”的目的是为了确保被告知者的充分“知情”,只有被告知者在充分知情的前提下才能自愿、明确地作出决定。为此,个人信息保护法第十四条明确规定:“基于个人同意处理个人信息的,该同意应当由个人在充分知情的前提下自愿、明确作出。”

第六,敏感个人信息的认定与保护规则。个人信息保护法对自然人的隐私信息未作出专门规定,而是将个人信息分为敏感信息和非敏感信息,并专节设置了“敏感个人信息的处理规则”,并采用了“个人信息被泄露或者非法使用+危害后果+列举重要敏感个人信息”的立法技术,明确“敏感个人信息一旦泄露或者非法使用,容易导致自然人的人格尊严受到侵害或者人身、财产安全受到危害的个人信息,包括生物识别、宗教信仰、特定身份、医疗健康、金融账户、行踪轨迹等信息,以及不满十四周岁未成年人的个人信息”。

第七,严禁“大数据杀熟”以及“用户画像”等涉及不当自动化决策的做法。针对“大数据杀熟”“用户画像”和“算法推荐”等涉及个人信息自动化决策的热点问题,个人信息保护法予以明确规范:首先,个人信息处理者利用个人信息进行自动化决策,应当保证决策的透明度和结果公平、公正,不得对个人信息在交易价格等交易条件上实行不合理的差别待遇;其次,个人信息处理者通过自动化决策方式向个人进行信息推送、商业营销,应当同时提供不针对其个人特征的选项或者向个人提供便捷的拒绝方式;再次,个人信息处理者通过自动化决策方式作出对个人权益有重大影响的决定时,个人有权要求个人信息处理者予以说明,并有权拒绝个人信息处理者仅通过自动化决策的方式作出决定。

第八,个人信息跨境提供规则。个人信息保护法明确了个人信息处理者向境外提供个人信息应当具备的基本条件,如关键信息基础设施运营者和处理个人信息达到国家网信部门规定数量的处理者,确需向境外提供个人信息的,应当通过国家网信部门组织的安全评估;对于其他需要跨境提供个人信息的,应由专业机构进行个人信息保护认证;个人信息处理者因业务需要向境外提供个人信息,须按照国家网信部门的标准合同与境外接收方订立合同;对我国缔结或者参加的国际条约、协定定向我国境外提供个人信息的条件等有规定的,可以按照其规定执行等。

第九,个人在个人信息处理活动中的七项权利。个人信息保护法全面构建了个人在个人信息处理活动中的权利,包括知情权、决定权(限制、拒绝和撤回)、查阅复制权、个人信息可移植权、更正补充权、删除权、规则解释权。这些权利的设定,表明了个人信息保护法对个人信息保护与利用的平衡,即在保护中利用,在利用中保护。

第十,重要互联网平台的“守门人”制度。鉴于重要互联网平台掌握海量用户数据,一旦发生信息泄露或滥用,可能引发严重后果,个人信息保护法专门要求其履行“守门人”角色,并承担更多责任。主要包括:应按照国家规定建立健全个人信息保护合规制度体系;成立主要由外部成员组成的独立机构进行监督;遵循公开、公平、公正的原则制定平台规则;对严重违法处理个人信息的平台内产品或服务提供者停止提供服务;定期发布个人信息保护社会责任报告并接受社会监督等。

此外,为了确保个人信息处理者遵守个人信息保护法之义务,严格规范个人信息处理者的处理活动,个人信息保护法设置了严格的行政和民事责任。例如,个人信息保护法第六十六条对违法处理个人信息,或者处理个人信息未履行法定个人信息保护义务设置了三项法律责任:一是由履行个人信息保护职责的部门责令改正,给予警告,没收违法所得;对违法处理个人信息的应用程序,责令暂停或者终止提供服务;二是拒不改正的,并处一百万元以下罚款;三是对直接负责的主管人员和其他直接责任人员处一万元以上十万元以下罚款。

如果有上述规定的违法行为,且情节严重的,设置了两项更严格的法律责任:一是由省级以上履行个人信息保护职责的部门责令改正,没收违法所得,并处五十万元以下或者上一年度营业额百分之五以下罚款,并可以责令暂停相关业务或者停业整顿,通报有关主管部门吊销相关业务许可或者吊销营业执照;二是对直接负责的主管人员和其他直接责任人员处十万元以上一百万元以下罚款,并可以决定禁止其在一定期限内担任相关企业的董事、监事、高级管理人员和个人信息保护负责人。(作者系浙江大学教授)

个人信息保护法实施应关注十大核心要点