

引文格式:周佳颖,吴道霞.竞争法视角下企业数据爬取行为合规研究[J].山西高等学校社会科学学报,2024,36(3):38-46.

竞争法视角下企业数据爬取行为合规研究

周佳颖,吴道霞

(中国人民公安大学法学院,北京 100032)



摘要:企业在市场竞争中扮演着数据的控制者和爬取者的双重角色。然而,从竞争法的视角来看,企业不当利用数据爬取技术可能引发不正当竞争和数据垄断问题。为了实现企业的可持续发展,必须重新定义大数据环境下企业的实质竞争关系;健全企业数据爬取的正当标准,即设立“善意爬虫”的三重标准,以确保数据爬取范围合法合规;完善企业数据开放义务,即结合企业经营类型,明确数据授权的必要性,改进反爬措施,实现数据的开放共享。只有正确应用数据爬取技术,将其作为推动企业发展的工具,企业才能获得长足的发展。

关键词:数据爬取;合规要求;竞争关系;数据开放;信息保护

DOI: 10.16396/j.cnki.sxgskxb.2024.03.006

中图分类号: D922.294 **文献标识码:** A **文章编号:** 1008-6285(2024)03-0038-09

一、问题的提出

随着数据时代的发展,人们越来越认识到“数据即资产”的重要性。海量数据的收集成为企业进入市场、获取竞争优势的关键。相比传统的调查问卷等方式,数据爬取技术因其具备快速整理和收集海量数据的能力而备受欢迎,已成为主流的数据收集方式。然而,这种中立的技术也容易被错误地使用,使其脱离中立性成为违法犯罪的工具。

从内容来看,企业数据既包括企业自身发展所涉及的企业发展前景和行业预测等与企业相关的数据,也包括为发展自身业务而广泛收集的公民个人信息和个人数据。在市场竞争法的视角下,一家企业可能同时扮演数据的控制者和数据的爬取者的角色。从数据控制者的角度来看,企业出于对自身数据的保护需求,以及对在数据收集过程中投入的人力、物力和财力资源的珍视,往往希望将数据锁进保险箱,不被任何人所探知,以确保在市场竞争中保持优势地位。尽管某些核心数据确实需要保护,但在实践中,这种行为已被垄断市场中的企业演变为数据垄断的手段。而对于数据爬取者而言,部分爬取行为是为了突破垄断企业的该种数据垄断,从而获

取进入市场从事正当竞争的“门票”。但也不乏部分企业意图使用数据爬取技术“搭便车”式获取商业信息数据,进行不正当竞争,扰乱市场秩序。数据控制者与数据爬取者之间进行的技术博弈既涉及企业间的正当竞争秩序判断,也包括对个人信息的收集以及合理的利用。企业既希望能够通过技术快速获取有效数据进入市场,又不愿自身数据被同样的竞争者攫取。在数据博弈之中,数据爬取技术被异化,成为违法的工具。在竞争法视角下,我们应当结合当前实际对大数据场景下的企业实质竞争关系重新定义,设立企业数据爬取行为的正当边界,完善企业的数据开放义务,以平衡数据控制者与数据爬取者之间的利益冲突,使得技术真正成为推动发展的工具。

二、数据爬取技术下企业间数据保护同数据共享的博弈

随着大数据时代来临,我国数字经济快速发展,党和国家大力推动数字产业化、产业数字化,数字经济也增强了我国的产业创新力与竞争力。2020年7月,国家发展改革委等十三部门联合发布的《关于支持新业态新模式健康发展 激活消费市场带动扩大就业的意见》中指出,要加快传统企业数字化

收稿日期: 2023-06-24

作者简介:周佳颖(1999—),女,山西晋中人,中国人民公安大学法学院硕士生,从事民商法学研究。

吴道霞(1967—),女,吉林安图人,中国人民公安大学法学院教授,硕士生导师,从事民商法学、数据法学研究。

转型步伐,培育一批数字化服务企业和创新应用企业,发挥引领带动作用;鼓励发展开源社区,支持开放软件源代码、硬件设计和应用服务^[1]。而2022年12月发布的《中共中央 国务院关于构建数据基础制度更好发挥数据要素作用的意见》中则进一步指出,要发挥我国海量数据规模和丰富应用场景优势,在维护国家数据安全、保护个人信息和商业秘密的前提下,推动数据合规高效的流通使用;坚持共享共用的原则,合理降低市场主体获取数据的门槛,强化反垄断与反不正当竞争^[2]。

这种政策大趋势对企业数据保护与共享提出了更高的要求。企业在市场竞争过程中扮演着双重角色。企业既是数据控制者,也可能是数据处理者、数据爬取者。在当今市场上,企业收集数据后再进行加工将会把一些单独无经济价值的数据整合成为决定企业占据市场竞争力的核心要素。当企业扮演着数据控制者的角色时,可能会采取设置“robots”协议、树立安全壁垒等反爬措施保护企业数据,防止竞争对手采取爬取手段大量收集自身数据,从而影响企业的市场竞争优势。但也不乏大型企业将反爬措施作为进行数据与市场垄断的手段,影响市场的正常竞争秩序。在企业扮演数据处理者时,如何合法合规地获取数据进入市场成为难题。企业希望同类竞争者能够适当履行数据开放、数据共享义务,但在经济效益至上的市场竞争中这一希望往往落空。由于企业间保护数据与共享数据发生争执,在其上升为不可调节的矛盾时,数据爬取技术便被异化为不择手段的工具。为此,我们首先应理解数字经济下企业间竞争模式的改变以及企业由于异化数据保护原则和数据爬取技术给竞争市场所带来的不利影响,以寻找数据保护与数据开放之间合理共存的支点。

(一) 大数据下企业实质竞争关系的扩张

在探讨如何最大化地发挥数据爬取这一中立技术的优势时,我们也必须认识到在互联网平台这一领域,企业间竞争模式较我们以往的认知有所改变。在传统的竞争模式下,被认定为竞争关系的企业大多体量相当、经营范围相近,且多数情况下以直接竞争者的关系出现^[3]。故而《反不正当竞争法》明确列明的不正当竞争行为多数是对产品的混淆、虚假误解宣传、损害商誉等一系列与企业所生产产品直

接相关的不正当竞争行为。但是在互联网时代,企业间的竞争模式已发生了改变。在数字平台中,能形成竞争关系的企业不仅公司规模相距甚远,而且提供的产品与服务之间可替代性不强,服务的顾客群体也大不相同^[4],甚至连主营业务的范围都可能天差地别,数据企业之间的竞争更多的是一种跨服务领域的竞争。可以说数据爬取技术参与下,企业竞争法意义上的竞争关系的定义标准较之以往应当有所改变。

首先,数据爬取技术下的企业竞争关系不局限于对服务和产品的直接竞争,无法单纯以产品、服务可替代性作为竞争关系或不正当竞争行为的判断标准。在数据市场早期,数据的利用方式主要是原生使用,数据被作为一种产品直接提供给大众,从而能够满足人们的各种需求^[5]。在这个竞争阶段,竞争企业最常用的手段是爬取其他企业数据,从而产出具有同类替代功能的产品或服务,挤兑原企业市场份额,攫取竞争优势。例如,在“深圳谷米科技有限公司诉武汉元光科技有限公司”^①一案中,被告元光科技有限公司旗下“车来了”APP为了提高下载与使用量,通过数据爬取技术,突破了谷米科技有限公司设置的技术壁垒,广泛抓取了“酷米客”背后的海量公交车行驶信息、到站时间等实时公交数据为己所用,削弱了“酷米客”APP的用户黏性,破坏了其本身具有的市场优势。二者形成了一种实质竞争关系,故而法院认定,即使被告辩称其所提供的服务并不具有营利性,两企业间仍存在竞争关系。而元光科技有限公司的行为扰乱了市场竞争秩序,未遵守商业道德,构成了不正当竞争。

但是随着数据的发展,数据更多地作为一种加工要素被使用而非对数据的直接使用,市场上的数据竞争由同类产品、服务的横向竞争转变成了产品、服务生产过程的纵向竞争。在数据初始收集阶段,在未经过加工的情况下,数据本身的单纯聚合可能无法产生很大的经济效益,企业数据爬取行为能产生交易价值的渠道在于数据的后续运用^[3]。小型企业通过对大型企业的数据库爬取而获取数据,但是在后续加工过程中,小型企业利用相同数据加工生产出了与大型企业原先提供的产品完全不同类型不同领域的产品,此时二者在产品和服务提供方面完全处于不同领域,小型企业产品与服务并未对原产

^① 案号:(2017)粤03民初822号——深圳市谷米科技有限公司(以下简称谷米公司)与被告武汉元光科技有限公司(以下简称元光公司)、邵凌霄、陈昂、刘江红、刘坤朋、张翔不正当竞争纠纷民事判决书。

品进行同类替代。例如,在2021年的“新浪微博诉蚁坊”^①一案中,被告通过大量收集微博内容数据,开发了舆情监测系统软件,这时两公司对数据的使用并不属于同业竞争。而案件后续认定“蚁坊”构成不正当竞争的缘由在于“蚁坊”获取微博数据的手段是突破了微博公司自己设置的技术壁垒,从而被认定为是对微博公司数据的不正当获取。而在2022年“蚁坊”公司则以微博公司限制访问用于分析舆论的数据为由,以反垄断提起了诉讼。可以看出,在类似案件中数据控制者与数据爬取者之间的竞争关系能否成立需要结合更多的条件去判断,企业实行不正当竞争行为的判断要件并不仅限于所爬数据产生的产品与服务是否构成对原产品的同类替代,还有对数据再生价值的判断^[5],此外法院会更侧重于考察获取和利用数据行为的正当性。

《反不正当竞争法》第十二条规定了互联网不正当竞争行为,但仔细阅读会发现该规定仍旧将目光更多地聚焦于网络产品或服务本身,带入数据爬取的背景下,该条的适用具有非常大的局限性。竞争企业所爬取的企业数据可能包含企业所收集的客户数据,包含可直接识别特定人的个人信息以及仅能泛泛体现个人喜好习惯的个人数据,例如运动习惯、用餐爱好等。企业数据也可能包含被界定为商业秘密的企业发展规划、市场前景预测等。可以看出,运用数据爬取技术所获得的数据似乎与企业所能提供的产品和服务并无直接线性关联,而数据爬取者也并非利用诱导、欺骗、恶意不兼容等手段影响被爬取数据企业的产品与服务,而是通过对数据的后续加工、利用开发新的产品和插件影响原企业的用户黏性,分流用户群体。在这种情况下,一旦原企业起诉数据爬取者不正当竞争,法院将只能适用《反不正当竞争法》第二条中“扰乱市场竞争秩序”这一原则性规定进行审理。

然而对于原则性规定的不明晰也造成了对该条适用的一些问题。竞争法的本质是一种对行为的规制,无论是《反不正当竞争法》还是《反垄断法》,可以从法条以及立法目的中清楚地看出通过对不当行为的规制达到维护市场正常竞争秩序、保护诚信经营者的合法利益进而维护消费者权益的目的。所以对相关行为的判断也应当是从行为出发,判断行为是否具有合法性,进而判断该行为是否因造成损

害而具有可罚性。当前立法对于数据领域的竞争损害判断的规定较为模糊,司法实践中似乎出现了一种“以损害存在”倒推“竞争行为具有不正当性”的现象^[3]。例如,在“大众点评诉百度”一案^②中,法院审判要点中点明,百度公司未经许可在“百度地图”和“百度知道”中大量使用了来自大众点评网的信息,实质替代大众点评网向用户提供信息,对原告公司造成损害,具有不正当性,构成不正当竞争。这里说理的逻辑在于被告的行为对原告产生了损害,所以被告的行为具有不正当性,构成不正当竞争,恰好是对竞争法从“行为到法益”判断的颠倒。这也体现出我们需要对数据领域的竞争损害判断进行进一步明确的需求。

其次,企业由于数据被爬取而遭受的损失难以直接衡量。在不正当竞争案件中,由于数据的独特性,使得数据爬取所造成的损失难以直接衡量。数据由于其特殊性,较之以往的生产产品与服务而言,难以单纯直接对数据本身进行经济化,数据的经济化往往与它的流动以及后续被加工所形成的产品、服务紧密相关。而正是这种特性使得被爬取数据的企业难以直接明晰自身所损失的究竟是何种经济利益。一是其直接的经济损失似乎没有办法衡量,因为数据本身的价值确定带着一种因人而异的特性。同样一份数据,对于在该行业打拼、了解其价值的企业来讲或许价值连城,而对于不同行业、不了解该数据如何运用的企业来讲或许一文不值。二是如果以可期待利益来定义数据的价值似乎也会遭受重重阻碍。因不正当竞争或者垄断行为而造成的经济损失本身就是难以估计的。因为这种价值只是企业估算的一种如果能够正确合理的运用该数据可能给本企业带来的利益数,这种数目并没有实际产品经过市场流转而获得真正的价值检验,而动态的市场变化则使得这种估算存在太大的出入。可期待利益本身是一件虚无缥缈、无所定论的东西,如果我们认可通过可期待利益来定义数据企业的损失,则可能导致数据企业漫天要价,肆意抬高可期待利益,从而更有可能助长不正之风。

在“深圳谷米科技有限公司诉武汉元光科技有限公司”一案中,由于谷米公司没能提供确切的证据证明其确有的自身损失的具体数额,也无法提供元光公司因窃取数据而获得的侵权利益是多少,故

① 案号:北京知识产权法院(2019)京73民终3789号民事判决书。

② 案号:(2016)沪73民终242号——北京百度网讯科技有限公司与上海汉涛信息咨询有限公司其他不正当竞争纠纷二审民事判决书。

而对谷米公司所主张的“数据损失评估”“数据技术维护费”等不予支持,酌情确定由被告赔偿原告一定的费用作为经济损失与合理维权费用。

无独有偶,在“大众点评诉百度”一案中,原告大众点评网所属汉涛公司所举证的审计报告中虽然载明了汉涛公司的营业成本,但这并不能成为其损失的直接经济依据。除此之外,汉涛公司无法提供百度公司通过爬取其数据进行不正当竞争所造成的经济损失相关证据,故而法院酌情确定赔偿及其合理维权费用。

数据时代下,企业间的竞争模式已较以往发生了转变,纠纷过程中各个方面都体现了数据的独有特性,所以我们在研究企业数据爬取合规行为时也应当注意到这种特性,应从新的角度去看待企业间的数据竞争。

(二)企业数据爬取技术的异化

当企业扮演数据爬取者的角色时,我们必须正视目前实践中确实有不少企业滥用数据爬取技术,海量爬取竞争对手的数据,从而不费吹灰之力优化自身产品与服务,形成利用他人数据优化自身与他人形成同类竞争的局面,通过“搭便车”“蹭流量”抢占市场份额,攫取他人已经深度加工过的成果并进一步夺取市场优势。例如,在2021年12月杭州市余杭区人民法院判决的“抖音诉小葫芦直播数据抓取”一案^①中,抖音公司发现“小葫芦”的开发者六界公司未经其许可,长期采取不正当技术手段,非法收集抖音直播平台用户打赏记录、主播收益情况等数据,损害了抖音的数据安全,同时也造成了用户对于“抖音”软件的不信任,造成“抖音”的流量流失。法院审判认为,六界公司通过技术手段,非法获得了本不被公开的数据,对“抖音”公司的经济和流量产生了巨大影响,已经构成了不正当竞争。被扭曲使用的数据爬取技术除对公开数据直接进行爬取外,还视“robots”协议与反爬措施于无物,违背企业对数据的开放授权,突破企业设置的防御措施爬取企业不愿意公开的数据,或者是通过解密等技术手段,绕开防御措施进行数据的爬取^[6]。目前,由于我们并没有对“robots”协议与反爬措施予以明确的法律效力认可,“robots”协议仅为互联网行业的一种倡导性规定,而反爬措施更只是企业自己设置的技术壁垒。如果可以以一个网站或平台单方面的意思来作为爬取行为的合法与否的判断标准,则将容易忽视网

站或平台是否有正当理由采用技术手段拒绝他人的访问,以及访问者是否有正当理由无视被拒绝而继续访问^[7]。突破或绕过这些反爬措施,爬取数据的行为应该如何被评价成了重点,而这也正是我们是否可以用不正当竞争行为来定性数据爬取行为的关键。如果数据爬取者是为了进入市场,或者数据爬取者与数据所有者之间无法构成实质竞争关系,或所爬取的数据没有体现数据所有者的创新性与深度加工,这时能否以不正当竞争来评价数据爬取行为呢?

(三)企业数据保护原则的异化

对于数据所有者而言,利用数据盈利不仅仅是依靠自己先前的收集行为,其实能够给自身带来利益的是收集数据之后进行的加工整合。当这些加工整合后的数据形成并且成为市场上不可替代的独一份时,这些数据便会拥有极高的商业价值,会给企业带来无可替代的市场竞争优势。市场机制也会以最快速度对市场上的变化做出反应,当这些数据变成一种生产要素,数据流动能够产生巨大的经济利益时,企业便会迅速认识到这些数据的重要性,从而将自身数据全部纳入保护范围。这种行为的起因在于企业意图保护自己的投资,这种投资既包括对数据收集以及加工所注入的经济成本,还包括在此期间的的时间成本等,同时企业所保护的数据中也可能包含企业自身的发展规划、上市意图等重要的商业秘密。企业对于商业秘密的保护无可厚非,但是不乏一些大型互联网企业将这种保护异化为数据垄断的手段,他们对数据的过度保护已经使他们成为市场创新的阻碍者^[3]。对于那些经过企业深加工的数据,企业遵循着数据保护原则,会通过设置各种保护壁垒的手段来防止其他数据采集者收集自己加工的数据,比如设置“robots.txt”或者防火墙式的反爬措施。但是拥有着海量数据的大型企业,却将所有的信息、数据一股脑地全部加密,间接导致后来者难以进入市场,或者必须花费许多时间重新收集用户信息。我们当然要保护企业在加工数据时的投入,但是在大数据时代,花费大量时间重新收集已经公开的一些数据必然会导致企业进入市场滞后,使得进入市场的门槛被大大提高,可能影响正常的市场竞争。而拥有海量数据的大型企业,则可能会滥用市场支配地位,通过采取技术手段拒绝其他企业数据爬取的方式,使得其他企业难以获得数据,从而巩固其在市场上所占据的优势地位。

^① 案号:浙江省杭州市余杭区人民法院民事判决书(2021)浙0110民初2914号。

且因为市场竞争者的减少,使得大型企业在服务与产品的提供上拥有着绝对优势,这种从上游数据垄断性保护的手段及至下游产品、服务上的垄断,使得消费者在相关产品服务上没有其他选择,而只能选择大型企业所开发的相关产品服务。这不仅伤害了消费者的自主选择性,企业也会利用这种独特优势将其演变成更大的经济优势,进行价格飙升或者“大数据杀熟”等操作,最后深受其害的则可能是位于产业链最终端使用产品与服务的消费者。“数据控制者对其持有的数据享有的应当是一种‘相对的权利’,行使权利时仍应注意保护其他利益主体的权益,避免造成数据垄断。”^[8]《反垄断法》强调的并不是对单个企业的利益保护,更多的是对市场正当竞争秩序的保护。所以当企业对数据的保护原则过度利用而进行数据垄断时,则应当落入《反垄断法》的规制领域^[9]。

三、企业数据爬取行为合规建议

企业在市场竞争时同时扮演着两种角色,但是无论是哪种角色,都应当明白“月圆则亏,水满则溢”的道理,企业既要正确地将数据真正运用得当,将之化为商场上使自己的轮船高歌远航的利器,又不能过度使用数据使自己身陷囹圄,必须要做到把握好技术使用的界限,以免过犹不及。企业需要正确运用数据爬取技术,作为数据爬取者时,想要快速进入市场的心态可以理解,但是要从爬取行为的正当性角度给自己设下限制,不可滥用技术,使得原本中立的技术变成淬毒之剑,危害其他公司的正当权益。而企业作为数据所有者时,也要注意数据保护的限度,遵循政策导向,在保护自身合理权益时,不过分地扩张数据权益边界,积极履行数据开放义务,促进市场正当竞争秩序,加快创新步伐。我们应当首先明确大数据下市场竞争关系应该如何判断,从而明了何为不正当竞争行为、何为垄断行为,继而明确爬取行为的正当性与企业应当履行的数据开放义务,对企业扮演的两种角色都提出规制意见,才能尽量做到中立导向,维护市场正常竞争秩序。

(一)重新定义大数据场景下企业的实质竞争关系

从竞争法的角度而言,规制数据竞争的核心目标是维护数据要素市场的竞争秩序和竞争机制,这既需要兼顾数据控制方与数据使用方的利益诉求,也需要均衡数据资源的产出激励效率和配置使用效率^[6]。在大数据场景下,企业间的竞争不再如原先一

般表现为横向的产品服务竞争,而是更多地以要素竞争、跨领域竞争为特征,所以法院在审理相关的数据竞争案件时,判断双方企业是否构成竞争关系是审理不正当竞争案件或者垄断案件的必要前提。如果双方企业被判断为不具有竞争关系,那么进行不正当竞争便无从谈起。而《反垄断法》中对于相关市场的定义也同样局限于在一定时期内就特定商品或服务进行竞争的商品范围和地域范围,也就是说至少是同业内才有可能形成滥用市场支配地位的垄断行为。如果企业间分属于不同的市场行业,提供不同的服务,双方在同业内不构成竞争关系,那么也无法以《反垄断法》进行规制。因此,在竞争法视角下,要想判断数据爬取技术是否会对双方企业造成影响、企业合规数据爬取的前提为何,就应该先明确大数据场景下企业间实质竞争关系应该如何定义,只谈论爬取行为是否正当而不顾及企业爬取的具体场景以及竞争性质,这种判断只会是空中楼阁、镜花水月。

首先,对企业竞争关系的判断要结合对产品可替代性的考量。当前企业间的竞争所表现出的是对数据的后续加工利用,所以企业间的竞争关系也由此被宽泛定义。从数据用户的视角来观察,用户对于数据产品和服务的使用是多平台多维度的,不同的产品与服务提供给用户的使用体验与创新体验并不同。所以在考量数据竞争关系时,应当将产品可替代性考量在内,如果企业间所开发的产品具有可替代性,提供的服务会使得用户放弃原先所使用的某一企业产品而转投另一方,则可以认为两企业间存在竞争关系。例如,在“深圳谷米科技有限公司诉武汉元光科技有限公司”一案中,元光公司与谷米公司都在做公交实时信息服务,开发的应用所提供的內容基本为同质的,对于用户而言,同功能的产品只需选择一个即可,这就是说两家的产品存在可替代性,可以用一个产品替代另一个而不会影响用户的使用体验,所以可以认定二者之间存在竞争关系。但如果双方公司所提供的服务并不存在可替代性,用户放弃某一产品将会影响其用户体验,用户要“兼收”的时候,二者间的竞争关系就需要依靠下一个特征来进一步判断。

其次,需判断双方企业所提供产品的经营模式是否相同。数据爬取者通过利用竞争对手服务器中的用户数据或其他已加工的实时数据,建立类似的商业模式或者完全进行与数据所有企业相同的商业模式的经营,该数据成为经营者开展经营活动的重要的投入要素,此时,在数据领域这两方构成竞争关

系^[10]。从判断企业提供的产品的可替代性扩大到判断企业的经营模式,是基于大数据时代互联网企业独特的特性。互联网企业所开发的产品多样,涉及的领域广泛,呈现为一种“去中心化”的上下游协同作业的趋势,因而一个大型的互联网企业可能将其经营模式覆盖整个产业链全环节,所以判断双方企业竞争关系时除了考量具体的产品替代性之外,也应该将企业经营模式纳入竞争关系判断的考量范围。当双方企业在经营模式上重合度越高时,则越有可能在各个领域形成竞争与垄断关系。例如,搜索引擎与音乐软件针对不同的需求群体提供服务并不一致的服务,但是搜索引擎却可能通过在界面上不断提醒无法打开音乐软件或者无法读取音乐软件内的内容的方式,迫使用户卸载该音乐软件,这已经构成一种竞争。这种竞争关系,可能是企业为拓宽经营范围,为自己接下来的商业战略提前铺路清除竞争对手,也有可能是企业因为与另一企业达成垄断协议,以竞争手段排除其他企业,使得协议中的企业能够快速进入市场经营变现,从而形成由上游至下游的垄断,本质上都是在市场上以不正当手段竞争。

最后,要考虑双方产品服务之间是否存在影响用户黏性、妨碍原产品销售、原服务正常提供,意图使原产品流量或经济损失的行为。在这里需明确,致使原产品流量或经济损失应当是行为的目的,而非是采取行为后所产生的结果。否则将会陷入前面论述的以结果倒推法益侵害的不合理逻辑。这一判断标准主要是从用户群体的角度出发考虑。“多重归属感”是数据平台的服务常态,数据用户可在平台上自由支配其数据、使用多个不同的提供商来提供不同服务甚至同一服务^[3]。如果某一企业所提供的服务中明确限制另一企业服务的正常提供,逼使用户必须在产品中择一使用,即使此时两企业并不属于同一产品、服务提供领域,也应当认为企业间存在竞争关系。因为这种“二选一”的行为,本质上是对企业间经营利益的影响。不同企业本应以不同的产品服务不同的用户群体,却以不正当行为影响对方企业的用户黏性,就是一种市场份额的争夺,背后体现对市场经济利益的抢占与商业上的排斥和歧视,应当被认为双方存在竞争关系。虽然企业经营过程中难免出现经营利益的变动,但是企业自身的发展不应当以不正当手段侵犯他人正当利益为跳板。

(二)健全企业数据爬取的正当标准

数据爬取作为一种中立的技术^[11],本来并无善恶之分,但在实践中人们往往会根据技术使用人员

的运用行为将数据爬取分为善意爬取与恶意爬取。善意爬取是指数据爬取的行为人出于善意的目的进行数据收集,即正确地利用数据爬取的技术便利大量收集信息,从而达到自身学习或者其他善意目的。当前所见的各种搜索引擎即是利用数据爬取技术搭建起来的。利用数据爬取工具,快速精准的“扫描”全网出现的所有信息,并从中提取需要搜索的关键词,按照关联度整理排序,然后整理“打包”呈现在网页上。如引起讨论的舆论检测系统也有学说认为其属于善意爬取。日常生活中利用数据爬取技术在学术网站广泛的搜集已经公开的论文资料也成为当前学术界学习与研究必不可少的一种方式。恶意爬取则是指数据爬取人海量搜集数据,并非出于个人学习或者工作目的,而是希望靠海量数据营利或达成其他非法目的。恶意爬取行为人突破网站反爬虫限制,违背数据所有者的意愿,肆意搜集未公开的数据,同时还有可能对正常运行的网站造成破坏或者影响普通用户的正常访问。更有甚者,通过数据爬取技术爬取特定的受法律严格保护的数据,如国家安全信息等,以此达到不法目的。

对于数据爬取行为本身,不能评判其是否合规,而应在实践中从其使用方面进行评价。要想使数据爬取技术真正发挥其作用,在企业数据竞争方面就应当严格限制爬取行为的合规要求。因为只有这样,才能保证数据爬取技术不会直接异化为刑事犯罪的手段,而竞争法领域当前所存在的数据垄断问题,则可以通过赋予企业数据开放义务而达成。结合国内学者与美国学者提出的善意爬虫的三种标准^[12],我们认为企业数据爬取行为正当应当至少满足以下条件。

首先,企业竞争所需爬取的数据应当是合法公开的数据。数据公开意味着数据所有人主动分享数据信息,许可该数据被第三方获取,并由他自己承担数据公开所带来的收益与风险。此时通过数据爬取获得数据等同于获得了数据所有者给予的查阅、复制、收集该信息的授权,对于该数据的使用会再次产生的经济收益原数据所有者已经有了足够的心理预期,且此种数据信息一般经过数据所有者处理,或许已经模糊了关键特征,即使被爬取、再加工使用,也已经较难与原数据所有者直接形成竞争关系的对立。如果数据爬取的是非法公开或未公开的数据,往往会需要通过技术手段突破或绕过原数据所有者设置的数据保护壁垒,此时的数据爬取行为便很难被评价为善意爬取行为,需要通过后续使用来判断

是否应当从竞争法、民法甚至刑法角度来规制。

其次,数据爬取行为本身不应当具备侵入性与破坏性。数据爬取技术就像一把双刃剑,一旦使用者心思不正,就会使其变成损人不利己的利刃,伤人伤己。要想使数据爬取技术真正成为能够给人带来便利且有益处的技术,就应当从使用方法上对其进行规制。爬取行为本身是一种对计算机信息系统的访问^[7],正确利用技术促进发展的企业,不应对他人设置的技术壁垒进行强行突破。因为对于公开数据,数据持有者通常不会设置访问网站或者内容的技术壁垒,也很少会采取措施禁止或者限制相关开放数据的获取,所以此时采用数据爬取技术收集就可以直接获取到开放数据,其行为也自然就不具备侵入性与破坏性。而对于那些受保护不开放的数据,数据持有者定会设置许多保护手段来保护数据,因为这些数据可能是企业的高度商业秘密,此时采用数据爬取技术强行突破侵入的行为不仅会涉及窃取数据、盗取商业秘密,也可能会涉及破坏计算机系统从而落入刑法规制领域。而敢于冒法律风险从事这项工作的人多数不会正当利用爬取来的数据,此时这一部分别有用心的人当然不是出于正当使用目的进行数据爬取。所以,使用者一旦为获取数据不择手段,通过爬取技术侵入网站并破坏相应的保护措施,甚至有可能对计算机系统造成破坏的,其爬取行为不能被评价为善意。但随着技术的发展,许多技术人员为规避这种情况,选择绕过技术壁垒,此时的爬取行为可能仅涉及直接对数据的获取,而并未对计算机信息系统造成破坏,对这种爬取行为善恶的评价需要结合其行为表现出的其他特征进行综合考量。

最后,数据爬取行为应当出于正规合法使用目的。在数据爬取行为满足以上两点客观要件后,还应当以是否具备正规合法的使用目的这一主观因素进行判断。如果所爬数据既属于开放数据,又在爬取过程中没有侵入性与破坏性,但是数据爬取者在整理所爬取的数据后,将其用来生产具有高度市场替代性的产品进行不正当竞争,甚至是售卖信息非法牟利,投入黑灰产业产生不当利益,从而侵犯原数据所有人的权利,更有甚者对社会公共利益产生了众多不利影响,那么这种爬取行为也是不合规甚至是违法的。

(三)完善企业数据开放义务

企业运用数据爬取技术严格按照上述的要求进行数据爬取,可以从技术层面避免企业将原本中立的技术变成进行不正当竞争的非法手段。但也有

可能造成企业以此为依据,将所有数据全部保护起来,进行数据垄断。所以也应当赋予企业一定的数据开放义务。但是这种开放义务应当严格界定开放数据的类型,这样才能在保护企业权利与经济效益的情况下促进市场良性竞争。

首先,数据开放的类型应当与企业经营类型相关联。现在的数据更多地呈现出复合性经济效益的特点。即单独的数据很难给企业带来有效的经济效益,数据必须进行复合整理与分析,形成企业所在领域的独特优势,才能给企业创造价值。所以我们认为要赋予企业数据开放义务必须限定数据开放类型与企业经营类型相关。对于一些未能体现企业加工整理价值的信息,且单凭这些数据无法影响企业经营的,应当让企业承担数据开放义务。因为这些数据是企业从用户处直接收集的,而企业在这个过程中仅起到将单一数据汇总的作用,这些数据与其经营范围无线性直接关联,也无法体现企业在相关数据中所注入的创新性。那么从竞争法角度而言,为了促进市场合理、良性竞争,应当要求企业对相关数据予以开放,使得更多善意竞争者进入市场。在良性竞争中,企业也会为了提升自身竞争力而积极谋求创新,这样有益于市场整体发展与市场秩序良性构造。但是如果相关数据涉及企业立身之本,与企业在市场竞争中的竞争力息息相关,则相关数据不应当要求企业必须开放共享。例如,对于投行而言,客户名单是他们竞争的关键要素,所以即使仅是客户的名字与联系方式等在其他经营范围内看起来很普遍的内容,也不能要求投行将其开放。而对于知乎、微博等用户所编辑内容具有独特格式的应用而言,也不应当将其相关内容的开放视为必须,应当允许企业对其下属应用做一定的数据展示限制,防止其他不良企业直接爬取用户编辑内容而不展现原平台从而进行“搭便车”式不正当竞争。

其次,衡量数据是否授权。此处的数据来源主要针对个人信息相关数据,企业爬取涉及个人信息的数据时,应当仔细衡量数据是否授权。对于某些敏感个人信息以及隐私信息,用户只授权将其信息公开给数据所有者,则数据爬取者不应当爬取相关信息,这样才能最大程度保护个人信息不被不法利用,同时最大限度尊重用户的个人信息自决。但是对于一般个人信息,如用户对颜色的喜好,对于出行的偏好,如果没有用户明确的否认,可以认为用户的信息公开授权不只针对数据所有者,数据所有者也不应当以此为由将信息锁入“茧房”拒不履行数据

开放义务。在要求企业衡量所爬数据是否拥有授权时,也应当要求数据所有企业在收集用户数据时对用户进行数据用途阐明,保证用户对于自身数据授权拥有足够高的可知性。

再次,判断所公开数据内体现的企业创新性。如果数据经过企业的深度加工,凝聚了企业的创新结晶,则该数据势必与企业未来发展及未来规划、市场竞争息息相关,因为企业不会无故加工自己所用不到的数据,加工后的数据体现了企业创新性,则不能对相关数据要求企业进行数据开放。但这种创新性的判断也应当结合企业类型进行实际考量,如果相关数据与企业经营类型相关,能够使企业在竞争市场中具有独特的竞争优势,则应当认为相关数据高度体现了企业的创新性,不能要求企业对其进行数据开放。

最后,要在竞争法上完善“robots”协议与反爬措施的效力。目前,我国并没有对反爬措施以及“robots”协议的法律效力给出明确的规定。反爬措施是各数据所有者在其客户端或者网站上自己设定的限制访问的一种措施。其并非限制所有人访问,而是限制或禁止采用某种特定方式进行访问的行为^[7],而“robots”协议目前依旧是互联网行业的一种自律协定。我们必须承认二者在保护企业数据利益方面起到了很大的作用,但是正是因为反爬措施与“robots”协议的法律效力不明确,所以企业过度利用反爬措施和“robots”协议进行数据垄断的行为难以被规制。突破反爬措施或违反“robots”协议在何种程度上需要被法律规制,哪些数据被纳入上述措施不予以开放可能构成数据垄断,上述措施在竞争法领域、民法侵权领域与刑法犯罪领域应当如何衔接,这都是我们应当思考的问题。我们要求企业履行数据开放义务,应当在法律上完善反爬措施与“robots”协议的效力,使得企业明确哪些数据被保护不会进入反垄断法的规制领域,表明涉及哪些部分的数据应该被写入协议并具有强制性的保护力。当对不法行为有了明确规制后,合法开放的领域自然显现出来。

四、结语

在竞争法的视角下,对企业数据竞争的考察不

应局限于个别企业,而是要通过完善制度来促进整个市场的合规合理竞争和良性发展。企业在数据竞争中扮演着双重角色,我们需要确立规则以确保其合规开展竞争,同时避免过度束缚,使企业成为盲目追逐利益的无脑追随者。首先,重新定义大数据场景下的实质竞争关系。其次,设立数据的爬取标准,从数据爬取的角度限制企业的不当竞争行为,使得爬取合法、合规。最后,明确企业一定的数据开放义务,使其在合规框架内进行数据开放。数据保护和数据开放并非完全对立,过度偏向任一极端都会对企业和市场造成巨大打击。因此,我们应在两者之间寻找微妙的平衡,以促进企业发展、发挥创新动力,推动市场经济繁荣运行。这样才能确保企业在合规范围内发挥数据竞争的积极作用。

参考文献

- [1] 国家发展改革委,中央网信办,工业和信息化部,等.关于支持新业态新模式健康发展 激活消费市场带动扩大就业的意见:发改高技〔2020〕1157号[A/OL].(2020-07-15)[2023-05-16].https://www.gov.cn/zhengce/zhengceku/2020-07/15/content_5526964.htm.
- [2] 中共中央国务院关于构建数据基础制度 更好发挥数据要素作用的意见[N].人民日报,2022-12-20(1).
- [3] 蔡川子.数据抓取行为的竞争法规制[J].比较法研究,2021(4):174-186.
- [4] 王磊,贡绍海.数据竞争中的个人信息处理规则问题研究[J].法律适用,2021(12):37-46.
- [5] 周樾平.数据爬取的不正当竞争认定规则研究[J].南大法学,2023(2):87-102.
- [6] 杨志琼.数据时代网络爬虫的刑法规制[J].比较法研究,2020(4):185-200.
- [7] 孙杰.数据爬取的刑法规制[J].政法论丛,2021(3):115-125.
- [8] 解正山.个人信息保护法背景下的数据抓取侵权救济[J].政法论坛,2021,39(6):32.
- [9] 陈兵.保护与竞争:治理数据爬取行为的竞争法功能实现[J].政法论坛,2021,39(6):18-28.
- [10] 李慧敏,孙佳亮.论爬虫抓取数据行为的法律边界[J].电子知识产权,2018(12):58-67.
- [11] 王立梅.爬取他人作品数据传播牟利行为的评析[J].人民检察.2021(10):43-44.
- [12] 童云峰.大数据时代网络爬虫行为刑法规制限度研究[J].大连理工大学学报(社会科学版),2022,43(2):88-97.

Research on the Compliance of Enterprise Data Scraping Behavior from the Perspective of Competition Law

ZHOU Jiaying, WU Daoxia

(School of law, People's Public Security University of China, Beijing 100032, China)

Abstract: In the market competition, enterprises play the dual role of data controller and data scraper, but from the perspective of competition law, the improper use of data scraping technology may lead to unfair competition and data monopoly. In order to achieve sustainable development of enterprises, it is necessary to redefine the substantive competitive relationship of enterprises in the big data environment, establish and improve legitimate standards for enterprise data scraping, namely the triple standard of "well-meaning scarpers", to ensure that the scope of data scraping is legal and compliant, and improve the enterprises' obligation of opening data, that is, combined with the operation type of enterprises to clarify the necessity of data authorization, improve anti-scraping measures, and achieve opening and sharing of data. Only by correctly applying data scraping technology as a tool to drive enterprise development can enterprises achieve significant growth.

Key words: data scraping; compliance requirements; competitive relationship; opening data; information protection

[责任编辑 令狐国芳]



(上接第 23 页)

[16] 付文军.《资本论》的社会历史辩证法思想释义[J].东南学术, 2022(6): 54.

[17] 徐文越,高惠珠.论马克思劳动观三重意蕴之历史思维[J].上海师范大学学报(哲学社会科学版),2022,51(6): 61.

[18] 詹姆逊.重读《资本论》[M].胡志国,陈清贵,译.北京:中国人

民大学出版社,2013: 74.

[19] 何召鹏.唯物史观视角下马克思所有制理论的科学内涵及其当代价值研究:兼评“取消所有制分类”等错误观点[J].政治经济学评论,2019,10(4): 183.

[20] 马克思,恩格斯.马克思恩格斯全集:第 26 卷[M].北京:人民出版社,2014: 311.

Exploration of Marx's Thought on Labor Liberation of the Working Class

LI Shibo

(School of Marxism, Beijing Normal University, Beijing 100875, China)

Abstract: Marx believed that the principle that "labor creates wealth" advocated by national economists could not dispel the paradox that "workers' labor brings poverty". Through the critique of political economy, Marx corrected the abstract expression of national economists about labor, revealed the domination of capital over labor, explained the commodification process of labor and the inevitability of workers' labor being devalued, and pointed out the path to labor liberation for the working class. Marx pointed out that wage labor is a "low-level" and "temporary" form and it is destined to be replaced by higher forms of free association labor in society. The fundamental path for labor to overcome capital lies in that the working class must abolish the wage labor system through revolutionary political action. The ultimate goal of labor liberation is to establish a "united, social individual ownership", eliminate the class attribute of labor, and return labor to human autonomous activities.

Key words: Marx; the working class; labor liberation; labor; capital; wage labor

[责任编辑 张红霞]