

数据抓取不正当竞争纠纷的司法审查要素考察与反思

赵丹¹, 沈澄²

(1. 深圳市社会科学院, 广东 深圳 518000; 2. 上海市汇业律师事务所, 上海 200333)

摘要:《反不正当竞争法》的司法解释删除了原征求意见稿中关于数据抓取行为的规定。司法实践中关于数据抓取行为的正当性判断依然没有明确的具体规则。目前业已形成的审查模式已经将抓取数据的合理必要限度、抓取数据的类型属性、被爬取方所遭受的损害以及竞争秩序保护中多元价值的平衡作为主要考察要素。但各要素的审查仍有进一步完善的空间, 司法机关在确定合理必要抓取限度时, 可考虑引入合理替代与部分替代标准对实质性替代要素进行限制; 在确定内容存在交叉的数据抓取规则时, 可结合反爬方对其数据采取保护措施的力度及其有效程度对损害结果要素进行修正; 在平衡竞争秩序保护多元价值取向时, 可考虑从消费者权益出发。

关键词: 数据抓取; 不正当竞争; 抓取限度; 抓取对象; 价值取向

中图分类号: D 913

文献标志码: A

文章编号: 2096-9783(2023)02-0052-08

数据作为一种新型资源和生产要素, 与企业的竞争利益和竞争优势密切相关。而数据的可获得性是关键性问题, 是经营者进一步发挥数据效用, 获得竞争优势的前提和基础。因此, 经营者之间围绕数据获取产生了较为突出的利益冲突, 争议纠纷不断增多。其中, 数据权益的权属、权利边界以及数据抓取行为的正当性应如何判断, 已成为司法实践关注的焦点问题。现行法律并未对数据权益提供“权利法”保障的范式, 目前我国对数据保护主要有商业秘密、反不正当竞争和财产权化三种路径^[1]。而以《中华人民共和国反不正当竞争法》(以下简称《反不正当竞争法》)为依据, 对数据抓取^①和处理行为进行“行为法”层面的规制, 在理论与实务界达成初步共识^②。有学者专门统计过近

年来的我国数据权益保护案件, “基本上都纳入《反不正当竞争法》一般条款或者互联网专条进行调整”^[2], 而前者适用的情况比重更大^{[3][20]}, 呈现出向“一般条款”“逃逸”的倾向。然而, 该条款作为对不正当竞争行为的一般性、概括性规定, 使用了较多抽象性、模糊性的表达, 增添了司法适用的难度^[4]。2021年8月18日, 最高人民法院发布的《关于适用〈中华人民共和国反不正当竞争法〉若干问题的解释(征求意见稿)》(以下简称《司法解释(征求意见稿)》)提供了一种解决路径, 即要求商业道德、用户同意、合法来源和非实质性替代等要件同时满足。然而, 2022年3月20日, 正式实施的最高人民法院《关于适用〈中华人民共和国反不正当竞争法〉若干问题的解释》(以下简称《司法解释》)却删除了

基金项目: 2021年深圳市哲学社会科学规划课题“人工智能科技的治理困境与法伦理学应对研究”(SZ2021D022); 国家社会科学基金资助青年项目“智慧司法的伦理风险及其应对研究”(22CFX079)

作者简介: 赵丹(1988—), 女, 黑龙江鹤岗人, 助理研究员, 法学博士, 研究方向: 国际法, 竞争法;

沈澄(1992—), 男, 江西九江人, 法学硕士, 研究方向: 国际法, 竞争法。

① 在研究数据领域不正当竞争问题时, 有学者使用“抓取”, 参见李慧敏, 孙佳亮: 论爬虫爬取行为的法律边界, 《电子知识产权》2018年第2期; 也有学者使用“爬取”, 参见刘源: 数据爬取行为规控的法解释路径, 《法律方法》2021年第36卷第4期; 还有学者认为应当使用“获取”更为合适, 参见傅显扬: 《大数据领域不正当竞争行为的法律规制》, 武汉大学出版社, 2021年版。本文使用“抓取”一词。

② 仲春, 王政宇: 数据不正当竞争纠纷的司法实践与反思, 北京航空航天大学学报(社会科学版), 2022年第1期; 刘源: 数据爬取行为规控的法解释路径, 《法律方法》, 2021年第36卷第4期; 张建文: 网络大数据产品的法律本质及其法律保护——兼评美景公司与淘宝公司不正当竞争纠纷案, 苏州大学学报, 2020年第1期; 李扬: 日本保护数据的不正当竞争法模式及其检视, 《政法论丛》, 2021年第4期; 侯媛: 反不正当竞争法视野下用户数据获取行为解读, 《经济法学评论》, 2018年第1期; 翟巍, 刘一诺: 反不正当竞争法视角下公开数据爬取行为的合法性边界, 《价格理论与实践》, 2021年第8期。

《司法解释(征求意见稿)》第24—26条规定的“恶意不兼容”“数据抓取”和“互联网专条兜底条款”。因此,在司法实践中,数据抓取的正当性判断仍然缺乏明确具体的法律规定,而只能依赖于裁判者对“一般条款”的要义把握和价值平衡。

本文通过对数据抓取不正当竞争纠纷案件的梳理分析,试图总结出在“行为法”规制层面,数据抓取正当性判断模式中主要要素及其内容的审查思路和存在的问题,进而对如何适用反不正当竞争规则来解决经营者之间因数据获取和使用而产生的利益冲突提出建议。

一、数据抓取正当性审查模式中的主要要素考察

目前,数据抓取不正当竞争纠纷案件的司法裁判中已经形成基本的正当性审查模式,包括:双方是否具有竞争关系、是否享有竞争性权益、抓取行为的损害后果、抓取行为是否具有不正当性等审查内容。而在具体的判断步骤中又会涉及具体要素的审查,即数据抓取行为的合理限度、数据抓取对象的类别差异、数据抓取的损害结果和竞争秩序的多元价值取舍与平衡。

(一)数据抓取行为的合理限度——实质性替代考察

数据流动是数据实现价值的重要方式。因此,在识别数据抓取及其后续利用的正当性边界时,我们必须考虑数据抓取行为的合理限度与损害结果的适当容忍之间的平衡关系。当一个产品抓取另一个产品数据,并对其长期积累的劳动成果产生实质性替代时,则可能被法院判定为构成不正当竞争^[5]。

“大众点评诉爱帮网不正当竞争纠纷”案中,一审法院首次提出了关于“实质性替代”的考虑。一审法院认为,“使用垂直搜索技术的网站对于特定行业网站的信息的利用,应控制在合理的范围内,不得对该网站造成市场替代的后果。而爱帮网已对全部商户简介内容和绝大部分点评内容进行了充分展示,网络用户一般不会再选择点击大众点评链接标识。因此,爱帮版的商户简介和用户点评已经构成对大众点评网相应内容的实质性替代,必将不合理地损害汉涛公司的商业利益……爱帮科技公司使用大众点评网商户简介和用户点评,构成不正当竞争。”^③虽然在本案中,一审法院未能就实质性替代给出具体的量化标准,但其中关于“全部”“绝大部分”的措辞,逐渐成为之后类案的参照。如在“大众点评与百度上诉”案中,法院认为,“百度公司在使用来自大众点评网的评论信息时,理想状态下应当遵循“最少、必要”的原则,即采取对汉涛公司损害最小的措施……而百度公司在百度地图和百度知道产品中大量使用来自大众点评网用户的评论信息,已对大众点评网构成实质性替代,这种替代必然会使汉涛公司的利益受到损害。”^④在“饭友APP抓取微博数据”案中,法院认为,“基于正常的阅读习惯,明星粉丝用户在饭友APP中浏览完相关内容后再回到新浪微博查阅相关内容的概率很低……因此,就涉案92个明星微博而言,饭友APP已对新浪微博构成实质性替代,并结合其他论点综合分析得出结论认为复娱公司抓取新浪微博数据并在饭友APP中进行展示的行为妨碍、破坏了新浪微博的正常运行,进而构成互联网不正当竞争行为。”^⑤在“微博诉云智联抓取用户数据”案^⑥中,法院对这一问题也作出了相同的判断。又如在“阿里巴巴诉南京码注1688”案^⑦中,法院认为,“码注公司计算企业活性

③ 上海汉涛信息咨询有限公司与爱帮聚信(北京)信息技术有限公司不正当竞争纠纷上诉案,北京市第一中级人民法院(2011)一中民终字第7512号。

④ 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不正当竞争纠纷上诉案,上海知识产权法院市浦(2015)沪73民终242号。

⑤ 上海复娱文化传播股份有限公司与北京微梦创科网络技术有限公司不正当竞争纠纷上诉案,北京知识产权法院(2019)京73民终2799号。

⑥ 北京微梦创科网络技术有限公司诉云智联网络科技(北京)有限公司不正当竞争纠纷案,北京市海淀区人民法院(2017)京0108民初24512号。

⑦ 杭州阿里巴巴广告有限公司、阿里巴巴(中国)网络技术有限公司等与南京码注网络科技有限公司等不正当竞争纠纷案,浙江省杭州市滨江区人民法院(2019)浙0108民初5049号。

值可以参考1688平台公开的数据,但其将1688平台公布的商家数据直接用于其网站,甚至可以直接替代1688平台的部分功能,显然超过合理限度。”

即使经营者使用正当手段进行数据抓取,也仍然需要考虑数据抓取给被抓取方所带来的损害问题。“如果他人抓取网络平台中的公开数据之行为手段系正当,则需要结合涉案数据数量是否足够多、规模是否足够大进而具有数据价值,以及被控侵权人后续使用行为是否造成对被抓取数据的平台的实质性替代等其他因素,对抓取公开数据的行为正当性作进一步判断。”^⑧

实质性替代除了作为损害结果的考察要素之外,有时也会贯穿不正当竞争审查的全过程,即既是“竞争损失”存在的一个方面,又是“竞争关系”存在的一个前提。比如,在“腾讯诉今日看点数据爬取”案中,法院依次考虑了“原告是否存在值得保护的竞争利益、原被告之间有无竞争关系、被告的竞争行为是否使原告竞争利益受损”^⑨这三个争议焦点,并且在后两个争议焦点的审理过程中均考察了替代效果问题。在“优酷诉蓝蓝盗链”案中,法院同样在竞争关系的判断中以替代关系作为标准,指出“近似商品就是具有替代性的商品,即这些商品在功能或者用途上可以互相替代。”^⑩

如果经营者进行大量投入所产生的竞争性权益(数据利益)不能得到有效保护,那么必然减损其进入或发展特定领域的主观意愿。而消费者将来可享受相关经营成果的渠道和数量亦会因此而减少。正如“额头出汗”原则所描述的那样:法律所应保护的,正是经营者通过努力所形成的劳动成果。在数据抓取不正当竞争纠纷案件中,劳动成果恰恰体现为经营者合法收集、编辑、分类存储的数据,以及基于该等数据向消费者提供的产品或服务。若经营者直接以复制、

下载或类似手段获取了这些劳动成果并替代提供产品或服务,则很可能构成不正当竞争的行为。

值得思考的问题是,实质性替代的程度边界应当如何界定?如果数据抓取后的使用场景并不形成直接的替代或竞争时,那么实质性替代的判断就是需要进一步细化的问题。

(二)数据抓取对象的性质差异——数据法律属性考察

根据属性特征和产生构成方式,数据可以被分为个人信息、用户生成内容、平台自身数据和衍生数据四类。而这四类数据的生成方式不同,所耗费的“劳动”不同,因此,对抓取行为的正当性评价也存在差异。

第一,抓取个人信息行为的正当性。在“新浪微博诉脉脉”案^⑪中,二审法院确立了个人信息抓取行为正当化的“三重授权原则”。根据该原则,数据抓取方首先应当获得特定个体的授权,从而收集其个人信息;其次,数据抓取方向被抓取方取得同意获得该个人信息;最后,被抓取方亦必须获得该特定个体的授权,从而向抓取方共享个人信息。根据司法实践的认定,“在Open API开发合作模式中,第三方通过Open API获取个人账户信息时应坚持‘用户授权’+‘平台授权’+‘用户授权’的‘三重授权规则’。该原则已成为开放平台领域网络经营者应当遵守的商业道德。”^⑫

第二,抓取非个人信息的用户发布数据行为的正当性。非个人信息的用户发布的数据(UGC数据),不是个人信息,但又是基于用户行为所产生的进入公共领域的数据,比如在“大众点评”案中,百度公司抓取的用户在大众点评上的评价信息;“头条数据抓取”案中,字节公司抓取的用户在微博上的发布内容,因此,UGC数据抓取行为的正当性判断需要在对公共

⑧ 北京微梦创科网络技术有限公司诉云智联网络科技(北京)有限公司不正当竞争纠纷案,北京市海淀区人民法院(2017)京0108民初24512号。

⑨ 腾讯科技(深圳)有限公司、深圳市腾讯计算机系统有限公司等诉广州合聚教育咨询有限公司等商业贿赂不正当竞争纠纷案,广州市越秀区人民法院(2020)粤0104民初46873号。

⑩ 北京蓝蓝科技有限公司与优酷信息技术(北京)有限公司不正当纠纷上诉案,北京知识产权法院(2019)京73民终3570号。

⑪ 北京淘友天下技术有限公司、北京淘友天下科技发展有限公司诉北京微梦创科网络技术有限公司不正当竞争纠纷上诉案,北京知识产权法院(2016)京73民终588号。

⑫ 深圳市腾讯计算机系统有限公司诉腾讯科技(深圳)有限公司商业贿赂不正当竞争纠纷案,天津市滨海新区人民法院(2019)津0116民初2091号。

资源的合理使用和平台授权之间寻求平衡。互联网企业经过长期投入运营、整合公共数据所形成的数据,凝结其劳动成果并形成市场价值。如果数据抓取方“不劳而获”地抓取了这些数据并对被抓取方形成实质性替代,那么,其行为很有可能被认定为构成不正当竞争。

第三,抓取衍生数据行为的正当性。与个人信息数据抓取要求三重授权不同的是,对于衍生数据,数据抓取方则应当遵守且仅遵守数据处理者的意志。在“生意参谋”案中,法院明确指出,“淘宝公司的数据产品是在巨量原始网络数据基础上通过一定的算法,经过深度分析过滤、提炼整合以及匿名化脱敏处理后而形成的预测型、指数型、统计型的衍生数据……系淘宝公司的劳动成果,其所带来的权益,应当归淘宝公司所享有。”^⑬因此,互联网公司基于原始数据整合处理形成的衍生数据产品,享有单方的权益。

需要说明的是,关于数据财产权问题,理论界也曾经出现过不同的观点,包括数据控制者权、企业数据权、大数据有限排他权、企业数据知识产权等。但《中华人民共和国民法典》(以下简称《民法典》)等法律并不主张对数据赋权。即便是衍生数据,也不能像知识产权那样,具有对世权利,否则将违反“物权法定”原则。理论上,比较好的解释可以将其视作一种“数据上请求权”^⑭,企业得以据此主张对合同外第三人的侵扰阻却。衍生数据的这一“权益”归属性质,也是现在各大数据交易所能够进行“数据交易活动”的基本前提。此类交易所交易的数据产品所输出的通常为交易方通过数据收集和处理生成的衍生分析结果,并不包含个人信息、商业秘密、运营数据等原始数据。

企业基于自主商业运作和投入,对于自采的运营数据也享有相应的权益,包括竞争性权益,但是如果相关数据构成《中华人民共和国网络安全法》(以下简称《网络安全法》)、《中华人民共和国数据安全法》(以下简称《数据安全法》)等法律法规意义上的“重要数

据”,则其数据的对外提供(即经过同意的抓取行为)仍应受到公法上的限制。一些重要的物流数据、金融数据可能会对经济命脉和民生造成重大影响,那么就不能仅凭经营者的同意即可抓取。

如果抓取的是除UGC数据或“重要数据”外的其他一般公共数据,且此类数据并不归属于任何主体,那么一般不应受到过多限制,否则将影响信息流动和信息共享,导致信息孤岛的出现。经营者在后续使用数据时,只要履行合理的注意义务即可。比如在“蚂蚁微贷”案^⑮中,法院提出,对“公共数据”的使用应当真实、及时、准确,这与资本市场领域的信息披露规则要求基本类似。再如,在“鹰击系统”案^⑯中,海淀法院和北京知识产权法院对数据做出了公开数据与非公开数据的划分,并以此作为区别两类数据抓取行为正当性边界的根据。

但是,数据有时呈现出混合性特征,使得已有的分类分级方式无法解决其抓取的正当性边界问题。因此,当混合型数据被抓取时,正当性审查的规则适用问题就值得进一步思考。

(三)数据抓取行为的损害结果——反爬方受有妨害及其异化考察

不少互联网企业的《服务协议》或《用户许可协议》里都会有类似的条款,即要求“未经许可,禁止对信息内容进行非法采集”。相应地,企业通过设置一定的反爬措施来限制或禁止特定的数据抓取行为,比如,在网页代码中设置 robots 协议,限制特定爬虫机器人的抓取。而 robots 协议是一种实现反爬的相对初级的非技术措施,并不能强制阻拦爬虫的脚本运行。除此之外,能够在技术上实现反爬的措施还有信息校验、动态渲染、文本混淆、IP限制、蜘蛛陷阱、加密算法、字体文件映射等方式。但是爬取方却不断通过多个代理IP(云服务)、伪造身份验证信息、人工辅助等操作绕开或突破平台设置的种种访问限制。可见,数据抓取方和反爬方始终在进行“道高一尺,魔高一

^⑬ 杭州铁路运输法院(2017)浙8601民初4034号淘宝(中国)软件有限公司与安徽美景信息科技有限公司不正当竞争纠纷一案《民事判决书》。

^⑭ 浙江蚂蚁小微金融服务集团股份有限公司、重庆市蚂蚁小微小额贷款有限公司诉苏州朗动网络科技有限公司商业诋毁及不正当竞争纠纷案,杭州铁路运输法院(2019)浙8601民初1594号。

^⑮ 湖南蚁坊软件股份有限公司诉北京微梦创科网络技术有限公司不正当竞争纠纷上诉案,北京知识产权法院(2019)京73民终3789号。

丈”的较量,而双方在这场“军备竞赛”中投入的水涨船高最终可能导致服务器不堪重负。如在“抖音诉小葫芦”案和“抖音诉六界”案中,抓取方不断突破或绕过反爬方的防护措施,通过多代理IP高频率、大批量的爬取造成了数据方的服务器的额外负担,被认定为不正当竞争行为^⑩。

关于数据抓取行为不正当性的损害结果的判断,需要审查其是否对反爬方服务器造成过重的访问负担,妨碍、破坏了其合法提供的网络产品与服务的正常运行。而被爬取方有时为了避免数据被抓取而采取的如 robots 协议等反爬措施本身也可能引起法律上的否定性评价。如“360引擎数据抓取”案中,反爬方的 robots 协议设置就被认定为是一种不正当竞争行为。因此,仅从反爬方“受害”的角度来审查数据抓取行为的正当性可能会产生悖论,即无论反爬方是不是真正的“受害”,爬取方的角色被限制为承担最终的举证说理义务,对于其数据爬取的正当性只有辩解的义务而无抓取的权利。

(四)竞争秩序保护的价值取向——三元利益均衡及其内在矛盾考察

在法律尚未规定权利范式的情况下,数据抓取行为正当性的判断主要依赖于价值工具分析。而《反不正当竞争法》中的价值工具主要是指消费者权益、竞争秩序和商业道德。

在“360搜索引擎”案中,法院指出“不应将 robots 协议作为限制信息流通的工具”^⑪,原因就在于其违反了以《互联网搜索引擎服务自律公约》为代表的商业道德。然而,在“头条数据抓取”案中,同一法院对商业道德的看法却发生变化,因为相关自律公约在该案中适用的领域不同。可见,细分行业的领域差异在对商业道德的要求上也产生影响。尤其是2017年的《反不正当竞争法》修订之后,将原规定中的“公认的商业道德”改为“商业道德”,去掉了“公认的”限定语,它允许对于“新市场和新产业等缺乏公认商业道德的领域,法官根据法律精神、市场需求等确定可资遵循

的市场道德准则,再据此判断竞争行为的正当性……使其在字面上既可以包括既存商业道德的情形,又可以包括法官认定的商业道德的情形”^⑫。这一修改在加强了反不正当竞争法“塑造”功能的同时,也更加加剧了商业道德在概念认定上的复杂性。

经营者竞争优势与消费者权益之间常常发生矛盾冲突。司法实践中经常出现视频网站经营者通过加框技术将其他视频方的付费观看视频“抓取”后提供给消费者免费观看的案例。在此类案件中,个别消费者短期利益的实现在整体上会伤害经营者和创作者经营优质服务和产品的积极性,并最终导致“劣币驱逐良币”的市场退出。

二、数据抓取正当性审查要素考察的反思及其进路展开

数据抓取行为的正当性边界非常复杂。原因在于:究竟对数据爬取行为应采取严格禁止的态度,强调对被爬取方的利益保护——遵从经营者利益保护优先的侵权法逻辑,还是应对爬取方的行为持竞争损害中立的态度——转向市场竞争秩序优先的竞争法逻辑,抑或引入其他多元利益,譬如国家利益、公共利益、用户利益等来综合权衡爬取行为的正当性,在实践中尚未达成一致意见^⑬。这为司法裁判带来巨大挑战。而通过对相关审查要素进行分析或许可以提供一些思路和启发。

(一)引入合理替代与部分替代标准

在考虑实质性替代的限度问题时,首先可以找到消费者权益保护层面的“阿基琉斯之踵”。《中华人民共和国个人信息保护法》(以下简称《个人信息保护法》)第四十五条赋予了用户数据迁移权,即用户可以请求数据平台将其个人信息转移至指定的平台(包括经营者的竞争对手)。那么,如果经营者通过合理运用该规则引导消费者将数据迁移到其自身平台时,所发生的替代行为似乎也将被合理化。据此,可以形成

^⑩ 深圳市腾讯计算机系统有限公司、腾讯科技(深圳)有限公司诉斯氏(杭州)新媒体科技有限公司不正当竞争纠纷案,杭州互联网法院(2021)浙8601民初309号;北京微播视界科技有限公司诉上海六界信息技术有限公司、厦门市扒块腹肌网络科技有限公司、浙江淘宝网络有限公司不正当竞争纠纷案,杭州余杭区人民法院(2021)浙0110民初2914号。

^⑪ 百度在线网络技术(北京)有限公司等与北京奇虎科技有限公司不正当竞争纠纷,北京市高级人民法院(2017)京民终487号。

实质性替代判断对基于消费者请求而产生的“合理替代”的忍让。

其次,在审查实质性替代因素时还应考虑受替代程度和竞争环境。如果抓取的数据仅仅在很小的维度上发生“替代”,那么此时可能需要引入“部分替代”的概念。比如,A公司作为一家综合性的社交平台,通过调用接口方式抓取了电商平台A公司存储的与用户运动装备偏好度和消费能力有关的部分数据从而用于改善对用户的间接画像。在这种情况下,审查实质性替代因素时,就应当考虑“部分替代”。对实质性替代进行审查时,还应考虑竞争者之间所共处的业务领域、特定的业务场景和模式、替代效果。仅当业务领域、业务场景和模式与替代效果相近时,才有理由发生实质性替代的主观恶性。一个行为是否落入反不正当竞争法框架应考察其是否匹配《反不正当竞争法》最基本的功能,即维护竞争制度和公共利益,因为在市场中有竞争必有损害才是常态,任何一个行为不因它造成的损害而轻易、直接被认定为不正当竞争^[8]。竞争本身就隐含着“损害”的内涵,所以在对“实质性替代”的审查中应当保持谦抑,为技术创新留下空间,从而实现对简单机械“复制粘贴”“搬运”数据的不劳而获、搭便车行为的规制,鼓励和支持提高技术手段深入挖掘和利用数据价值的行为。

(二)明确数据分类分级与抓取正当性的递进关系

经营者所控制的数据依照不同的标准可以划分为不同类型,而不同标准下的数据之间会产生交叉。虽然目前的司法实践对某些特定类型数据的抓取行为作出回应,并形成不同的审查规则和标准。但由于数据分类会存在一定的交叉特性,因此,建立一套裁判评判体系必须先将数据种类的划分标准进行统一。这一标准应当符合并反映出互联网产业中有序竞争关系的规范要求。

从数据领域的法律发展来看,进行数据分类分级保护已有共识。比如《数据安全法》第二十一条规定“国家建立数据分类分级保护制度”。美国已有的司法实践也在探索数据的分类分级保护。“Facebook 诉 Power Ventures”案^[9]中,法院认为,仅获取用户许可

的 Power Ventures 公司不足以让其行为合法,他们应该在获得 Facebook 公司的授权许可后才能正规获得数据并合法使用,因此该经营者明显触犯了计算机欺诈和滥用法。“HiQ 诉 LinkedIn”案中,美国第九巡回法院对“HiQ 诉领英(LinkedIn)”案作出判决,认为抓取可公开访问的数据是合法的,LinkedIn 不得依据 CFAA 来禁止 HiQ 从 LinkedIn 的网站上抓取会员数据。但在终审判决中,法院作出有利于 LinkedIn 的判决,HiQ 的抓取行为被认定为违法,原因在于其违反 LinkedIn 的用户协议以及其创建虚假身份的行为。数据的类型直接决定了抓取的可允许性的差异,也即从数据流通的角度而言,那些公开数据的抓取应当被允许,但如果数据本身附着于商业努力和经济价值之上,则因为价值归属而应有别于公开数据的保护。比如,以表 1 为例,当抓取方抓取的是第 4 顺位的数据时,则只要取得被爬方的许可即可;当抓取方抓取的是第 3 顺位的数据时,则需要同时取得个人信息主体和被爬方的双重许可;当抓取方抓取的是第 1 和 2 顺位的数据时,其抓取行为很难通过获得任何主体的同意而变得正当化。

表 1 数据级别与顺位

顺位	级别	示例
1	影响国家安全的核心数据	国家秘密
2	影响公共利益的重要数据	经济运行类数据
3	影响消费者权益的用户数据	个人信息
4	影响经营者利益的一般数据	企业自行整理挖掘的衍生数据
5	其他数据	公开数据

当被抓取的数据中包含了不同顺位的数据时,则应当考虑分别适用审查标准,除非这种切割的做法不可行或者显著困难。而不宜直接以严标准吸收宽标准的形式适用统一的严格标准,否则可能与数据流动的价值取向相冲突。由于竞争的广泛存在,数据保护的技术措施和反爬手段是必要的。但不同层级的数据所对应的保护力度也应有所区别。因此,当数据本身的级别难以被清晰界定时,应当结合反爬方对其数据采取的保护力度及其有效程度,对审查结果进行修正。此外,数据使用场景的变化也有可能对数据级别

^[8] Facebook Inc. v. Power Ventures, Inc., 828 F.3d 1068(9th Cir.2016).

调整产生影响。例如:虽然用户对某一产品功能的点击频率、在某一页面停留的时间长短并不属于敏感的数据,但如果此类数据与用户账号信息进行匹配结合就可以据此分析用户偏好并用于进行个性化推荐,那么数据敏感度显著提高,数据级别也应上调。

(三)从消费者权益出发平衡数据抓取正当性判断的多元价值取向

根据芝加哥学派的观点,竞争秩序的本质要求是对消费者福祉的增加。根据一般条款的规定,市场竞争秩序、经营者合法权益和消费者合法权益是认定不正当竞争行为的基本要素^①。从宏观角度看,良好的竞争秩序和促进消费者整体的利益增加是一体的,所以在考虑经营者利益的同时,消费者权益是否受损也应当被重点审查。

目前,互联网公司之间的数据不正当竞争行为基本上是在争夺存量而不是增量用户市场。基于互联网领域独特的动态竞争格局,消费者对经营者的忠诚度较低,而通过抓取数据后“免费提供”的模式往往能够快速实现平台竞争效果。其中,用户自主权和选择权仍然是《反不正当竞争法》所强调的消费者权益。在“e成招聘”案中,虽然被告有调用原告平台上用户简历数据的操作,但该等调用多是基于用户的授权和意愿,是用户自主选择、自主操作的结果,据此法院认定被告的行为不构成不正当竞争^②。因此,在消费者权益、竞争秩序和商业道德三者之间发生冲突时,应当首先把握消费者权益本身与竞争效果分析的内在一致性。在具体评判时,剔除冗余的干扰因素,并结合互联网产品与服务领域的特征,强调对消费者知情、选择和隐私等长远利益的保障。

在关于数据抓取行为造成损害结果的评价方面,其实同样需要进行消费者权益受损的潜在评估。“损害结果”不仅仅是一种私法意义上的侵权损失,因为竞争双方彼此采取的行为都可能对对方造成伤害,这

也是数据抓取和数据反爬都有可能构成不正当竞争的固有问题。评估数据爬取损害时,应在此基础上进一步衡量是否会对其他使用反爬方网络产品或服务的消费者造成潜在损害的效果。

我们还可以通过其他的规则来理解这样的价值取向。比如在刑事司法实践中,通常数据抓取行为有可能触犯的罪名是破坏计算机信息系统罪。从该罪名的结果要件看,如果破坏计算机系统导致用户不能正常运行的数量达到一定量级,则可能构成“后果严重”。举重以明轻,在刑事司法中尚且存在以后果论罪刑的情况,作为刑法的前置法,同样也会以用户使用体验下降作为评判标准。这也是许多突破或者绕过反爬措施的数据爬取方被认定为实施了不正当竞争行为的底层逻辑,他们伤害了大量用户正常使用互联网产品或服务的权益。

三、结语

通过对当前司法实践的观察,本文总结和归纳了司法机关在数据抓取不正当竞争纠纷案件中对实质性替代的必要限度、数据类型、竞争损害以及竞争秩序保护的审查思路、做法及存在的不足,并建议引入合理替代与部分替代标准对实质性替代的限度进行阐明;将数据分类分级与正当性顺位进行匹配,同时适用宽严渐次的审查标准;以消费者权益为中心衡量竞争秩序保护的价值取向,从而在个案中既恪守反不正当竞争规则的谦抑性并抑制裁判的随意性,又在终极取向上回归对消费者的关照。2022年11月22日,国家市场监督管理总局发布《关于公开征求〈中华人民共和国反不正当竞争法(修订草案征求意见稿)〉意见的公告》,其中第十八条^③也规定了类似的审查规则和价值取向。

^① 上海逸橙信息科技有限公司与北京智联三珂人才服务有限公司、北京网聘咨询有限公司其他不正当竞争纠纷案,上海杨浦区人民法院(2019)沪0110民初16688号。

^② 第十八条:经营者不得实施下列行为,不正当获取或者使用其他经营者的商业数据,损害其他经营者和消费者的合法权益,扰乱市场公平竞争秩序:(一)以盗窃、胁迫、欺诈、电子侵入等方式,破坏技术管理措施,不正当获取其他经营者的商业数据,不合理地增加其他经营者的运营成本、影响其他经营者的正常经营;(二)违反约定或者合理、正当的数据抓取协议,获取和使用他人商业数据,并足以实质性替代其他经营者提供的相关产品或者服务;(三)披露、转让或者使用以不正当手段获取的其他经营者的商业数据,并足以实质性替代其他经营者提供的相关产品或者服务;(四)以违反诚实信用和商业道德的其他方式不正当获取和使用他人商业数据,严重损害其他经营者和消费者的合法权益,扰乱市场公平竞争秩序。

参考文献:

- [1] 许可. 数据保护的三重进路——评新浪微博诉脉脉不正当竞争案[J]. 上海大学学报(社会科学版), 2017(6): 15-27.
- [2] 孔祥俊. 商业数据权: 数字时代的新型工业产权——工业产权的归入与权属界定三原则[J]. 比较法研究, 2022(1): 87.
- [3] 陈兵. 保护与竞争: 治理数据爬取行为的竞争法功能实现[J]. 政法论丛, 2021(6): 19-20.
- [4] 张倩雯, 吴少华. 企业数据爬取的反不正当竞争法规制——基于中美案例的比较研究[J]. 政法论丛, 2022(1): 81.
- [5] 仲春, 王政宇. 数据不正当竞争纠纷的司法实践与反思[J]. 北京航空航天大学学报(社会科学版), 2022(1): 25.
- [6] 马宇飞. 企业数据权利与用户信息权利的冲突与协调——以数据安全保护为背景[J]. 法学杂志, 2021(7): 164.
- [7] 孔祥俊. 反不正当竞争法新原理总论[M]. 北京: 法律出版社, 2019: 263-265.
- [8] 刘沛响, 沈瞿和. 涉数据竞争行为正当性的判定标准分析[J]. 电子知识产权, 2022(1): 44.
- [9] 张倩雯, 吴少华. 企业数据爬取的反不正当竞争法规制——基于中美案例的比较研究[J]. 科技与法律(中英文), 2022(1): 82.

Judicial Practice and Reflection on the Review Elements of Unfair Competition Disputes over Data Scraping

Zhao Dan¹, Shen Cheng²

(1. Shenzhen Academy of Social Sciences, Shenzhen 518000, China;

2. Huiye Law Firm, Shanghai 200333, China)

Abstract: Interpretation by the Supreme People's Court of Several Issues Concerning the Application of the Anti-Unfair Competition Law of the People's Republic of China deletes the provisions on data scraping which was provided in the Draft for comments, which has made the rules of the judgment absent. The existing judicial practice and experience has formed a mode that has taken some main elements into account, such as reasonable and necessary limits, the type of the data, the damage suffered by the crawled party, and the balance of multiple values in the protection of the competitive order. However, when the Judges review these elements, they need to improve the path. When determining the reasonable and necessary limits, the court judicial may consider introducing reasonable substitution and partial substitution standards to limit substantive substitution; when determining that there are overlapping data capture rules, they may based on the strength and effectiveness of the data protection measures taken by the anti-climbing party to revised the review results; when balancing the competition order and protecting multiple values, consideration can be given to consumer rights.

Keywords: data scraping; unfair competition; limit of data scraping; the type of the data; balance of value