

利益平衡视角下电商数据抓取行为的 法律规制研究

王馨茹

浙江理工大学启新学院, 浙江 杭州

收稿日期: 2025年7月14日; 录用日期: 2025年7月28日; 发布日期: 2025年8月27日

摘 要

随着电商平台数据价值的凸显, 国内数据保护与数据流通之间的冲突加剧, 并由此衍生出电商数据抓取不正当竞争行为。现行法律规制忽视数据的流通利用价值, 存在着法律规范体系不完备、认定标准和阻却判断标准模糊等问题, 亟需引入利益平衡理论重构规制路径。利益平衡理论指引下的电商数据抓取规制, 需在寻求开放数据使用、利用和分享和保护数据权益之间的平衡的基础上明晰电商数据抓取的本质特征, 在反不正当竞争法中增设电商数据抓取滥用条款, 并于司法实践中明晰阻却电商数据抓取违法性的判断标准, 构建统一的裁判标准。这样有助于为平台企业数据合规实践提供清晰指引, 也能更好地维护电商市场竞争秩序。

关键词

数据抓取, 利益衡量, 反不正当竞争法, 数据垄断

A Study on the Legal Regulation of E-Commerce Data Scraping Behavior from the Perspective of Interest Balancing

Xinru Wang

Qixin Honor School, Zhejiang Sci-Tech University, Hangzhou Zhejiang

Received: Jul. 14th, 2025; accepted: Jul. 28th, 2025; published: Aug. 27th, 2025

Abstract

As the value of data on e-commerce platforms becomes increasingly evident, conflicts between data

protection and data circulation have intensified in China, giving rise to unfair competition behaviors involving the scraping of e-commerce data. Current legal regulations overlook the value of data circulation and utilization, and suffer from issues such as an incomplete legal framework, unclear standards for determination, and ambiguous criteria for exemption judgments. It is urgently necessary to introduce the theory of interest balancing to reconstruct the regulatory framework. Under the guidance of the theory of interest balancing, the regulation of e-commerce data scraping should clarify the essential characteristics of such activities while seeking a balance between the open use, utilization, and sharing of data and the protection of data rights. This should be achieved by introducing provisions addressing the abuse of e-commerce data scraping into the Anti-Unfair Competition Law and establishing clear criteria for determining the legality of such activities in judicial practice, thereby establishing a unified standard for adjudication. This would provide clear guidance for platform companies' data compliance practices and better maintain the competitive order of the e-commerce market.

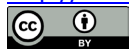
Keywords

Data Scraping, Interest Balancing, Anti-Unfair Competition Law, Data Monopoly

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

技术创新正在从根本上改变在线零售市场竞争格局，推动了以商品数据、价格数据、用户行为数据为核心的电商竞争模式的形成。头部电商平台依靠先发优势垄断了大量用户与经营数据。数字抓取技术应运而生，使得比价网站、数据分析公司、竞争对手等第三方主体得以低成本获取平台的数据资源，有效打破了数据垄断，但也引发了严峻的不正当竞争问题[1]。现有法律框架下电商数据抓取不正当竞争行为缺乏明确的规制，导致实践中出现认定标准混乱、边界模糊等问题。因电商领域数据抓取引发的不正当竞争纠纷不断增加，规范电商数字抓取乱象已是当务之急，相关部门需明晰电商数字抓取行为的本质属性，以合理兼顾电商竞争秩序、数据共享、消费者利益等多元价值为基础，制订出统一裁判规则。

2. 电商数据抓取乱象的规制困境

电商数据抓取是指平台经营者未经许可利用爬虫等自动化数字技术低成本、无对价获取其他电商平台用户个人信息、商品信息、价格数据、用户评价、销量统计等各类数据的行为，这是利用技术手段规避授权义务来获取他人投入大量资源形成的数据资源，直接攫取他人的衍生数据价值。它会使原始数据控制者的创新意愿受到打击，并破坏数据要素市场中价值决定的定价机制和正常有序的交易秩序。此类案件中涉及数据权益多样和各法益的利益平衡等因素，存在裁判标准不统一的问题。

2.1. 法律规范体系的不完备性

司法机关在实践中主要借助一般条款和兜底条款进行规制，这在一定程度上影响了裁判的统一性和可预见性，但归根结底是由于立法上的缺陷所造成的[2]。《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》)第十二条虽为互联网领域的专门条款，但是仍然没有规定清楚电商数据抓取的行为[3]。在最高人民法院关于反不正当竞争法司法解释的征求意见稿中，第26条有规制数据抓取行为的内容，但该条款最终未能保留在正式文本中。现《反不正当竞争法(修订草案)》第13条第4款涉及数据抓取的规定，

但该条款将获取并使用共同作为不正当竞争的成立条件,这显然门槛过高。这种立法态度的摇摆反映了规则制定的复杂性和困难性,最终给实践中的司法裁判带来困难。

2.2. 认定标准的模糊性

在审理涉及互联网领域新型不正当竞争纠纷时,我国法院通常将数据视作一种财产权益,以损害经营者利益、违反商业道德为由认定电商数据抓取的不正当竞争行为[4]。商业竞争本质上必然涉及对竞争对手既有利益的冲击,普通的商业竞争并不会破坏竞争秩序,单纯以损害竞争对手利益为由认定不正当竞争过于宽泛。最高院司法解释将“商业道德”界定为“特定商业领域普遍遵循和认可的行为规范”,并要求法院综合考虑行业规则、经营者主观状态、市场竞争秩序等因素进行判断。这种规定虽然为司法实践提供了一定指引,但在数据抓取这一新兴领域,尚未完全成型的行业准则和商业惯例使得普遍认可的标准难以把握。这种起源于侵权法的权利侵害认定范式以保护权益为核心,侧重限制第三方的数据抓取行为。这种方式曾是各国早期反不正当竞争法的主要实践路径[5]。电商数据抓取涉及平台经营者、第三方服务提供商、消费者、竞争对手等多元主体的利益博弈。单纯的权利保护模式容易形成对平台方的过度倾斜保护,忽视了消费者和社会公众的利益。新技术、新业态不断涌现,司法机关应引入利益衡量方法平衡数据保护和数据流通之间的冲突,把维护竞争秩序的思想贯彻到对电商数据抓取行为的判断之中。

2.3. 阻却事由的判断标准模糊

判断用户授权效力是对电商数据抓取行为是否违法造成阻却意义的关键。但除少量确定的用户数据外,多数电商平台数据的权属状态处于不明确的状态[6]。在司法实务中对用户授权能够产生对数据抓取行为违法阻却效力的观点也不相同。有的持否定说,认为用户授权同意并不具有阻却效力。在腾讯公司等与厦门联络易公司案中,法院认为被告收集的信息已获得用户的许可,但旗下软件具备的收集、存储、监控微信数据的功能导致这部分数据存在较高与其他数据的风险,该风险已经威胁到微信平台的安全运行,构成不正当竞争行为。¹有的持消费者利益说,以保护用户知情权和选择权为依据肯定用户授权的阻却效力。在 Facebook 诉 Power Ventures 案中,法院认为当被告的访问是得到用户授权且未规避技术封锁时,其数据爬取行为不违法。²还有的从数据利益分配权的角度出发,承认用户授权数据许可的合理性[7]。由于上述认识的差异,一方面造成了个案裁判分歧,另一方面也限制了数据要素的安全有效流动。为了维护数据要素的有效安全流动并发挥其效益作用,就需要建立起统一且有效的违法性阻却标准。

3. 优化规制电商数据抓取行为的立法路径

电商数据抓取的不正当竞争行为,系数字经济快速发展的重要产物。但是现有的法律规定并不充分,且缺乏对电商数据抓取行为的行为规范。要想从根本上治理该行为,完善的立法是根本性的解决方案。

电商数据抓取行为长期以来在立法上无具体涵义,不利于对执法、司法人员形成同一的认识,也不利于统一、稳定地预见可能的处罚后果。因此,最优先要做的就是确立科学精准的概念。《反不正当竞争法(修订草案)》第13条第4款规定:“以欺诈、胁迫、电子侵入等不正当方式,获取并使用其他经营者合法持有的数据”。虽然有对该数据抓取行为的规制,但仅是泛泛而谈,且对于取得使用数据的条件过高。电商网络爬虫是一种自动化发送数据请求,获取数据的高效手段,在高频次大流量的抓取电商数据的过程中就会导致对电商平台服务器资源的过度占用,破坏电商平台的正常运营[8]。例如 eBay 诉 Bidder's Edge 案中,后者每天发送 10 万次查询请求导致对方服务器资源过度占用。³因此,在《反不正

¹微梦公司诉字节跳动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 24530 号民事判决书。

²Facebook, Inc. v. Power Ventures, Inc., 844 F.3d 1058(9th Cir. 2016).

³eBay, Inc v. Bidder's Edge, Inc, 100 F. Supp. 2d 1058 (N.D. Cal. 2000).

当竞争法》应设置专门条款，把电商数据抓取定义为通过自动化技术手段[9]，对他人网站、应用程序或在线平台上的电商数据进行批量采集、复制、存储、使用或出售的行为。这一定义既包含了技术特征，又不过度依赖于特定技术实现方式，兼顾技术中立性与适度前瞻性。根据北京知识产权法院在审理奋韩网诉 58 同城不正当获取分类信息案时的观点，应避免技术固化关注行为本质与竞争效果⁴，印证了该立法思路的合理性。

从这一角度看，构建电商数据抓取不正当竞争行为法律规制体系要明确电商数据抓取行为的内涵，正视其本质以及产生的竞争效果，同时把握好数据流动与市场主体创新热情之间的均衡点。对上述规则加以规定之后，既可以为电商企业提供可供参考的操作性规范，也为执法与司法提供统一标准，以此来推进电商经济健康有序发展。

4. 优化规制电商数据抓取行为的司法裁判路径

司法作为法律规制的另一重要维度，在应对电商领域数据抓取不正当竞争行为具有不可替代的价值。面对司法实践中存在的认定标准不统一、阻却事由模糊等问题，司法机关可从以下两个角度规范电商数据抓取行为。

4.1. 利益衡量考察

利益衡量方法的引入，使得法院得以跳脱出单一主体利益保护的窠臼，转而关注数据要素市场的整体利益格局。电商数据抓取不仅关于经营者与其他经营者之间的利益，还关于消费者网上使用数据便捷和安全的利益以及公众利益[10]。通过衡量多种利益，可避免单一保护原数据持有者的利益，有助于协调不同群体之间的利益冲突，更有助于在个案正义与法律理性之间寻求平衡。法院在认定电商数据抓取行为的合法性方面应多方考量电商平台经营者、互联网消费者、社会公众利益。

第一，以实质性替代为判断标准保护原数据经营者的利益。随着数字经济的发展，电商平台间的竞争早已不是简单的同类竞争，已突破传统的直接竞争范畴延伸至多个相关市场。新浪微博诉脉脉案中法院提出了“三重授权原则”，第三方应用通过平台间接获取用户数据时需要获得用户、平台方和用户的三重授权，即用户对平台的授权，平台对第三方应用的授权，用户对第三方应用的授权[11]。该原则运用到电商数据抓取的争议解决中将人为抬高企业数据获取的正当性标准，降低反不正当竞争法的适用门槛容易导致反不正当竞争条款的泛化适用[12]。这种做法非但无助于问题解决反而可能扼杀数据流通的活力，与构建开放、协调的数字经济生态理念背道而驰。比起三重授权原则的复杂授权机制，实质性替代标准在间接竞争领域是判断电商数据抓取行为的合法性的不二之选。后者允许平台采集、利用他人电商网站上的非商业数据，但是采集、利用的行为不能对他人的电商网站造成实质性的市场替代[13]。典型如大众点评诉百度案，其实百度早期部分引用大众点评网信息的行为是在法律允许范围内的，因为查看完整的评价仍需跳转到大众点评网，此时还未对大众点评网构成实质替代。⁵当数据抓取行为对电商平台经营者只造成轻微损害时，竞争法应当保持克制，只有达到实质性替代标准时才有司法机关干预的必要[14]。这一标准既保护了原数据经营者的核心竞争利益，又为合理范围内的电商数据利用预留了空间。

第二，应当在兼顾信息获取与安全保障的前提下保护消费者利益。信息知情权和个人信息保护权是互联网消费者的主要信息权益[15]。一方面，电商数据爬取行为可以消除信息不对称现象。一部分数据爬取者获得数据后对数据进行再加工、编排、整理，丰富消费者获取信息的来源，提高消费者获取信息的

⁴青岛韩华快讯网络传媒有限公司诉北京五八信息技术有限公司不正当竞争纠纷案，北京知识产权法院(2017)京 73 民终 2102 号民事判决书。

⁵上海汉涛信息咨询有限公司诉北京百度网讯科技有限公司、上海杰图软件技术有限公司不正当竞争案，上海知识产权法院(2016)沪 73 民终 242 号民事判决书。

质量,促进数据流通,打破信息孤岛。比价平台的出现就为消费者提供了便利,有助于其获取多平台的商品优惠信息。另一方面,个人信息容易在不当的电商数据抓取过程中泄露[16]。部分公司或组织关于数据合规和数据信息安全的保障制度并不完善。在获取数据时容易受到黑客攻击以及不当的操作行为可能直接引发数据泄露风险;在获取数据后,其制度和技术的完备性也很难保证用户的数据安全。据 Threat Hunter 报告,电商行业是 2024 年数据泄露事件数量排名第二的行业[17]。因此,充分保障消费者的知情权以及信息安全应成为法院审理电商数据抓取案件的重要考量因素之一。

第三,社会公共利益的评估应聚焦数据要素市场的有效竞争和数据流通的有效价值。在电商竞争市场中,数据保护备受关注,但数据垄断的现象也愈发突出。大型互联网电商平台汇聚了大量的需要下游企业与之合作才能获得的资源性数据。但出于自身利益考虑,即使这些数据与下游企业间并不构成直接竞争,这些平台也都对向下游企业开放这些数据有所保留,甚至根本不提供数据开放服务[18]。又如天猫公司对聪明狗公司研发的比价插件拒绝授权许可并恶意屏蔽该插件,致使聪明狗公司难以开展正常业务。⁶差异化利用未构成实质性替代的电商数据抓取行为不应受到法律禁止。当抓取者利用数据的方式与原数据控制者的使用方式存在显著差异,且未对后者造成市场替代效应时,此类行为理应得到电商经营者的容忍[19]。因此,司法上应对电商数据抓取对于打破数据垄断和促进市场竞争的方面的作用给以必要的考虑,严格规制大型电商公司的不正当数据垄断优势地位的滥用。不过也不能单纯采取一刀切的方式去对待,还需结合电商行业特征、数据依赖程度以及有无其他获取同类数据途径等因素进行综合判断。

电商数据抓取行为涉及的多重利益往往存在冲突,法律难以通过预设价值位阶直接做出裁决,需要借助比例原则予以平衡[20]。作为利益衡量的核心工具,比例原则要求行为手段与目的达成合理对应关系。在淘宝诉美景数据纠纷案中,法院认为尽管被告抓取并转售电商平台交易数据有助于部分用户更高效地取得信息收益,但其侵犯了原告商业生态的做法却远远高于技术中立的所带来的少量增益,故判定构成不正当竞争。司法机关应当通过衡量电商数据抓取的行为手段之必要性、损害、行为方式和利益得失的大小等方式对各方主体的权利义务进行衡量,在个案中达到数据专有保护和数据共享利用的动态平衡,实现数字经济时代下多主体之间的互惠共赢、多元发展的状态。

4.2. 阻却违法性的认定考察

4.2.1. 用户同意抗辩

数据授权是判断电商数据抓取行为合法性的关键,电商平台数据库中的数据可依权属类型划分为:公共数据、用户专有数据、平台专有数据以及用户与平台共有数据。针对用户专有数据,基于私法自治原则,用户对其享有完整的支配权,其授权同意当然具有违法性阻却效力[21]。但用户对于其他数据的同意并不必然阻却数据抓取行为的违法性。以微梦诉字节跳动案为例,虽然被告主张已获得用户授权同意书,但法院认定该授权范围不应涵盖原告平台的全部数据的使用权。⁷这一裁判思路凸显了在处理用户与平台共有数据时的复杂性。

对于判断用户与平台共有数据的抓取合法性,应当着重考量数据形成过程中用户和平台的贡献度比例。若该电商数据主要源自用户提供或用户贡献占据主导地位,对这部分用户授权的数据,应扩大合理使用数据的范围。在平台未明确采取登录限制、使用 Robots 协议等技术保护措施且用户明确授权同意的情况下,对电商数据抓取行为应作出合法性评价。同理,若平台仅对用户数据进行简单复制整理,缺乏

⁶浙江天猫网络有限公司诉南京聪明狗网络技术有限公司不正当竞争纠纷案,北京知识产权法院(2019)京 73 民终 1490 号民事判决书。

⁷淘宝(中国)软件有限公司诉安徽美景信息科技有限公司不正当竞争纠纷案,杭州互联网法院(2017)浙 8601 民初 4034 号民事判决书。

创新性投入,则用户的授权同意足以阻却违法性。反之,若平台运用聚类分析算法、随机森林算法等分析手段对数据进行深度加工和价值重构,将原本无价值或低价值的数据改造为具有独特性的数据资产,则数据抓取者在取得用户同意的同时需取得平台数据许可[22]。

4.2.2. 技术措施的非合理使用范围抗辩

电商数据抓取纠纷主要集中在非公开数据抓取的纠纷上。数据资源以是否受到技术措施保护作为区分公开数据和非公开数据的判断依据[23]。在反数据抓取领域中,技术措施是保护数据资源的有效手段。绕过技术措施获取数据的行为在司法实践中基本上被认定为不正当竞争的观点为技术措施赋予了显著的合法性和排他性[24]。但在数据抓取与反抓取的对抗愈演愈烈的今天,技术措施滥用现象屡见不鲜,将企业收集的所有数据包括用户数据、平台数据、用户与平台共有的数据,甚至涵盖未匿名化的个人数据都纳入到企业的数据权益范围中。数据作为新时代的“石油”是可无损耗重复利用的社会资源,其价值并非静态存在,而是在流通利用中才能充分释放[25]。正如经济学家维克托·迈尔-舍恩伯格在《大数据时代》中指出的那样,数据的价值在于使用,而非占有[26]。第三方机构对数据的利用可能触及平台的核心商业利益,却也可能为平台构建更为完善的商业生态圈。谷歌搜索引擎通过抓取全网数据在为用户提供检索服务的同时也为网站带来了巨大的流量价值。平台应对第三方的数据抓取行为保留一定的容忍度,其技术措施防护强度应与数据的敏感程度和商业价值相匹配。技术措施所保护的数据范围是否合理应当是判断电商数据抓取行为阻却违法性的考量因素之一[27]。只有当第三方数据抓取者绕过技术措施爬取的电商数据超出合理数据保护范围并对平台和消费者造成实际损害时,反不正当竞争法才应当介入。

5. 结论

数据资源已成为电商平台经济发展中的核心竞争资源。大型电商平台企业凭借前期积累的海量用户数据建立起市场壁垒,而数据抓取技术的出现为打破数据垄断、促进数据要素流动提供了新的可能。为了维护正常的竞争秩序,立法机关有必要明确数据抓取行为的合法性边界。从判断框架的统一上讲,应该以利益平衡为基础,综合权衡利益,理顺数据开放共享和权益保护的关系,既要保障电商平台企业基于资源投入而形成的对自身数据的权益,又要保护消费者利益和社会公共利益,避免数据垄断带来的不良竞争态势对市场的正常竞争秩序造成损害。为提供明晰的裁判指引,司法机关应结合技术措施合理运用情况、用户授权范围以及数据公开程度等因素建立统一的评价体系。数据抓取行为的规制不应止步于个案正义,而应当立足长远的数字经济大局,在肯定技术创新促进数据流动的价值的同时,正确引导企业合理使用数据抓取的技术手段,建设开放、共享、有序的数据要素市场。

参考文献

- [1] 丁国峰,王静.大数据时代数据抓取行为的竞争法规制与反思[J].贵阳学院学报(社会科学版),2024,19(5):68-73.
- [2] 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):80-90.
- [3] 蒋舸.《反不正当竞争法》网络条款的反思与解释以类型化原理为中心[J].中外法学,2019,31(1):180-202.
- [4] 刘沛昀,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):44-58.
- [5] 弗诺克·亨宁·博德维希.全球反不正当竞争法指引[M].黄武双,刘维,陈雅秋,译.北京:法律出版社,2015:3-6.
- [6] 卢代富,张煜琦.从权益保护到利益衡量:数据抓取行为正当性认定的路径优化[J].河南财经政法大学学报,2022,37(6):59-70.
- [7] 高建成.竞争法视阈下数据抓取行为的合法性——从用户同意到可携带权的思路演进[J].江西财经大学学报,2024(1):112-123.
- [8] 洪延青.AI时代的数据爬取治理:法律冲突与利益平衡之道[J].政法论丛,2025(3):105-122.

-
- [9] 饶传平. 论数据抓取法律风险的流程化管理[J]. 东方法学, 2023(6): 28-42.
- [10] 孙晋, 马姗姗. 反垄断视野下数据开放与隐私保护的冲突与协调[J]. 武汉大学学报(哲学社会科学版), 2024, 77(6): 154-166.
- [11] 黄明健, 孙陈浩杰. 数据抓取行为正当性边界的法律分析——以中国、美国、欧盟模式为例[J]. 金融与经济, 2024(3): 75-84.
- [12] 何金海. 企业数据流通的实践困境与破解路径[J]. 西南金融, 2022(11): 43-56.
- [13] 李勇. 数据不正当竞争行为的实质性替代标准[J]. 中国流通经济, 2023, 37(6): 115-127.
- [14] 孔祥俊. 论反不正当竞争法的竞争法取向[J]. 法学评论, 2017, 35(5): 18-31.
- [15] 张守文. 消费者信息权的法律拓展与综合保护[J]. 法学, 2021(12): 149-161.
- [16] 詹红星, 王李. 网络爬虫行为的入罪路径及其限度[J]. 北京邮电大学学报(社会科学版), 2023, 25(1): 94-103.
- [17] 威胁猎人. [黑产大数据] 2024 年互联网黑灰产趋势年度总结[EB/OL]. 2025-01-16.
<https://www.threathunter.cn/blog/2024-e9ea9ca6-6081-438b-9417-9a4fdae8156d>, 2025-05-06.
- [18] 陈来瑶, 马其家. 平台企业数据共享的反垄断法规制[J]. 情报杂志, 2022, 41(6): 99-107.
- [19] 黄武双, 邱思宇. 行为区分视角下公开数据抓取的反不正当竞争法规制[J]. 科技与法律(中英文), 2025(3): 1-13, 25.
- [20] 李园园. 数据抓取行为不正当竞争司法认定规则探究[J]. 黑龙江社会科学, 2024(3): 59-67.
- [21] 韩世鹏. “三权分置”视角下的数据不正当竞争规制路径[J]. 南方金融, 2023(10): 88-99.
- [22] 包晓丽. 二阶序列式数据确权规则[J]. 清华法学, 2022, 16(3): 60-75.
- [23] 邱福恩. 商业数据的反不正当竞争保护规则构建[J]. 知识产权, 2023(3): 77-100.
- [24] 袁帅. 企业数据权益私人保护中技术措施滥用的“技术-法律”协同规制[J]. 暨南学报(哲学社会科学版), 2025, 47(2): 107-121.
- [25] 高富平. 数据持有者的权利配置——数据产权结构性分置的法律实现[J]. 比较法研究, 2023(3): 26-40.
- [26] 维克托·迈尔·舍恩伯格. 大数据时代: 生活、工作与思维的大变革[M]. 周涛, 译. 杭州: 浙江人民出版社, 2012: 78.
- [27] 李剑. 数据抓取行为的正当性判断——反思实质性替代规则[J]. 政法论丛, 2024(2): 63-74.