

电商平台数据抓取行为的法律认定与困境

黄 典

扬州大学法学院, 江苏 扬州

收稿日期: 2025年5月27日; 录用日期: 2025年6月10日; 发布日期: 2025年7月7日

摘 要

数据在电商平台间的竞争价值日益凸显, 部分经营者利用数据抓取技术获取别的电商平台上的大量数据。单纯从《个人信息保护法》视角进行法律规制难以保护电商平台经营者对于数据处理后的增值利益, 同时对于初始数据过分保护也会导致数据作为生产要素在市场内难以流通, 不利于相关产业的发展。此前的法院在处理数据抓取类案件中常用《反不正当竞争法》中的一般条款和互联网专条, 然而条文内容的宽泛性使得在数据抓取行为的认定和裁判领域内, 法官自由裁量权的范围不断外延。本文将在明确数据权益的法律属性和归属基础上, 进而判断数据抓取行为正当与否, 是否存在法律规制必要与可能, 最终在裁判领域提出相关纠纷的裁判规则思考。

关键词

数字经济, 《反不正当竞争法》, 个人信息保护, 数据抓取, 电子商务

Legal Determination and Dilemma of Data Scraping Behavior on E-Commerce Platforms

Dian Huang

Law School of Yangzhou University, Yangzhou Jiangsu

Received: May 27th, 2025; accepted: Jun. 10th, 2025; published: Jul. 7th, 2025

Abstract

The competitive value of data among e-commerce platforms is increasingly prominent. Some operators use data scraping technology to obtain a large amount of data from other e-commerce platforms. Regulating solely from the perspective of the *Personal Information Protection Law* is difficult to protect the value-added interests of e-commerce platform operators after data processing. At the

same time, excessive protection of initial data will also make it difficult for data as a production factor to circulate in the market, which is not conducive to the development of related industries. Previously, courts often used the general provisions and internet-specific provisions of the *Anti-Unfair Competition Law* in handling data scraping cases. However, the broadness of the provisions has led to the continuous expansion of the judge's discretionary power in the field of determining and adjudicating data scraping behaviors. This article will first clarify the legal attributes and ownership of data rights, and then determine whether data scraping behaviors are legitimate and whether there is a need and possibility for legal regulation. Finally, it will propose thoughts on the adjudication rules for related disputes in the field of adjudication.

Keywords

Digital Economy, "Anti-Unfair Competition Law", Personal Information Protection, Data Scraping, E-Commerce

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. 问题缘起

不正当竞争现象在当前的电子商务平台间已屡见不鲜，不仅危害到其他经营者的权益，还破坏了经济秩序，而在数字经济时代大背景下，数据已逐渐成为关键的生产要素，其在各行各业中的核心价值正日益受到企业经营者的深度挖掘与重视。尽管数据的价值实现和增值在流通之中，且数据的产生本身源于公众并具有一定的公开性，这就意味着数据的原始性质应当是一种个人信息，故而对于原始数据的收集采用仅需得到用户的授权或认可。但电子商务平台经营者若对庞杂繁琐的数据信息进行一定的加工、融合等措施后，数据信息的商业价值将会大幅提升，若此时还仅从《个人信息保护法》视角进行法律规制会导致超过必要限度的数据抓取行为造成平台经营者对应利益的受损，破坏市场竞争的秩序，助长“搭便车”“不劳而获”的不良之风在电商平台竞争内的盛行。然而当前对于数据信息的保护在现有的法律体系中还不甚完备，针对数据抓取行为的规制，主流观点倾向于通过《反不正当竞争法》来进行约束。在司法实践中，数据抓取行为主要依据《反不正当竞争法》中的通用规定以及专为互联网环境设计的“互联网专条”来进行规范，但是该规定过于笼统且具有模糊性，这就放任了法官的自由裁量权，难以统一裁判标准。

2. 数据权益的法律属性

在数字经济时代，电商平台间竞争的焦点通常集中在数据资源和用户的争夺上。公开的个人信息处理活动必须是合法的并且是在合理的范围内进行的，对合理与否的判断应当依据必要原则与目的限制原则在权衡不同利益的基础上进行，尤其是要考虑个人信息的公开目的^[1]，经营者通常采取数据抓取技术来获取大量数据，然而数据本身的公共属性以及开放性使得数据资源在一般情况下更呈现出一种公共资源的外貌，因此，在探讨数据抓取行为是否正当之前，首要任务是对数据权益的法律性质进行明确界定。这决定了当事人是否享有数据权益及内容以及是否存在诉权，这将直接影响到案件的审理结果。

2.1. 财产权

数据本身具有的半透明性以及其使用遍布现今互联网产业的各个环节决定了无法存在单一主体对其

具有绝对排他性权利。参照知识产权的权利效力,其仅具备相对的支配性和排他性,因此,传统的物权保护手段难以直接套用于知识产权的保障上。鉴于智力成果的无形特质,权利人对其的支配并非通过物理上的实际控制来实现,而是体现在对其的认识与利用上[2],在排他性层面,权利人对于智力成果的利用,并不排除其他非权利人对其进行合法利用或作出改进的可能性。故而可以得出,绝对支配性或者绝对排他性并非是界定财产权的唯一标准[3]。在数字经济时代,数据作为可供人们收集与利用并创造巨大价值的新型财富形式,确定其财产权属性得到了学界普遍支持。物权以权利人排他性使用为原则,他人的合理使用为例外,与此相对,数据权利则是以他人的合理使用为原则,权利人排他性使用为例外。因此,在构建数据财产权时,虽然可以借鉴物权保护的一些核心理念,但绝不能直接套用其保护模式[4]。根据数据的属性,电商平台对不同内容和产生方式的数据的生产投入有着巨大差别,因而不能“一刀切式”认定平台对数据的财产权。

2.2. 特许经营权

有学者倡导的特许经营权理论认为,数据权益的获得是基于法律明确授权,据此,相关从业者能够拥有对数据的经营权以及数据资产的所有权[5]。但是这一模式意味着将初始的数据权利赋予国家,不符合数据产生和利用的过程以及规律,故而在本文中不予讨论。但在我国,特许经营权还有另外一种表现形式,即一家企业有限期地或永久地授权另一家企业使用其商标、商号、专利权、专有技术,按照合同规定,在特许者统一的业务模式下从事活动,并向特许人支付相应费用。此种表现形式或许可以同数据抓取类商业行为适配,某一电商平台可以将自身深加工的数据资源授予第三方特许经营并获取相应收益。但鉴于数据资源的原始产生来源于公众,故而对于特许经营合同的内容尤其是费用的规定应该受到国家的监管,避免平台为追求利益过高设定所谓的加盟金或权益金,导致第三方的成本激增竞争力及创新力减弱。同时,一旦特许经营合同的内容不受监管,第三方成本难以负担时,或许选择正常特许经营模式的第三方或减少,反而增加了不正当数据抓取行为的增加。

综上,在不同的视角下分析了数据权益的法律属性,虽然都具有一定的合理之处,但数据权益的产生和来源十分复杂,我们不能将所有数据统一纳入人格权或者财产权的范畴,而应当以数据财产权为基础,根据数据内容的不同以及在流通阶段由不同主体对其进行的加工和投入确定其法律属性。如用户在使用网络服务时所产生的涉及个人隐私相关数据,应当赋予该数据人格权的属性,而在经营者在生产经营过程中,对收集到的数据加以过滤、调整、加密、筛选等加工产生的与企业经营业务或能创造重大利益有关的数据,此时该数据的主要商业价值明显并非诞生即存在,而是经过了经营者的商业投入创造的,则应当赋予其财产权的属性,肯定其财产性权益。在认定财产权的基础上,引入特许经营合同模式来合理分配平台收益以及第三方可使用数据资源的内容与权限。但对于合同内容的规定不可完全由合同双方设定,避免平台强势地位导致第三方成本增加使得最终难以合理使用数据抓取行为。

3. 数据抓取行为合法性判定的困境

数字经济的开放特性导致了对数据抓取行为正当性的判定变得复杂而难以明确。笔者将从对数据抓取行为正当性认定较为集中的竞争关系、违法性判定和竞争损害三个方面展开分析。

3.1. 竞争关系的存在并非判断行为正当性的绝对先决条件

传统的竞争关系通常被阐释为,那些提供可替代商品或服务的经营者之间,为争夺交易机会和市场份额而展开的相互竞争关系。由此可见商品或服务的可替代性是决定传统竞争关系是否存在的关键性要素。但数字时代互联网产业内各领域之间的界限并非完全清晰或具备很高门槛使得其他企业难以进入,以抖音这一网络平台为例,我们很难界定其到底属于电商平台或是社媒平台。正是因为经营者经营领域

可能随着市场风向随时变化,因此在司法实务中逐渐突破对传统的竞争关系的认定,不再局限于对商品或服务的可替代性的关注,而在对竞争关系进行更宽泛的理解时,例如在新浪公司起诉脉脉的案件中,法院依据双方用户和经营范围存在高度重合的实际情况,判定两者之间存在竞争关系[6]。因此不能以某一时间点下,平台与第三方经营者业务是否存在竞争关系来认定第三方经营者的爬取行为是否正当。然而,竞争关系的存在同样是评估数据抓取行为是否正当的关键因素。故而对于竞争关系存在与否产生的认定作用应当根据案件具体应用场景的不同灵活适用。对于数据抓取行为的违法性评估应当以行为特点为中心结合数据的流通与非竞争性价值综合分析。

3.2. 把握数据抓取行为的违法性边界在司法实践中尤为关键

由于数据抓取行为的技术色彩高度浓重,无法确定第三方的数据抓取行为系违法行为的情况下只能通过技术手段规范而非法律手段,这也是学界存在部分学者认为数据不可法律赋权的原因[7]。司法实践中主要通过技术中立原则,实质性替代原则,公认的商业道德等对数据抓取行为的违法性边界进行考量。

3.2.1. 技术中立原则

数据抓取技术实质是科技发展带来的科技创新。对于技术本身而言,立法与司法应当坚持中立立场,并在实际操作中持谨慎态度,以避免因过度干预而阻碍技术的创新与发展。本文中所谈的技术中立原则源自1984年美国的“索尼案”,该案是美国版权法中最早引入技术中立概念的案例。“索尼案”确立了“实质性非侵权用途”的标准,该标准指出,如果一项技术具有重大的、商业上的非侵权用途,那么即使该技术被用于侵权行为,也不应被视为间接侵权。国内部分学者认为,技术中立原则的理解应该基于“双重用途技术”的概念,即技术同时具备侵权使用和非侵权使用这两种用途。技术是否构成侵权,关键在于使用者的行为和价值取向,而非技术本身。在我国司法实践中,技术中立原则的适用界限已相对明确。通常,仅当技术提供者保持中立时,才可能基于该原则免除其法律责任。然而,对于技术的实际使用者,则需根据其具体行为是否合法来进行判断[8]。因此,当数据抓取技术的实施者存在恶意竞争、损害社会公共利益、违反商业道德或破坏公平竞争秩序等行为时,他们便偏离了技术提供者的中立立场,违背了技术中立原则的初衷。在这种情况下,技术中立原则不能作为免责的依据。所以,在处理具体的数据爬取纠纷时,必须综合考虑技术使用的具体方式、实施目的,以及该技术对市场竞争产生的实际影响。

3.2.2. 实质替代性原则

实质性替代指的是,当事人所提供的信息或数据相关产品、服务之间存在实质性的可替代关系。以“汉涛诉爱帮案”为例,爱帮网使用了来自大众点评网的内容,其中商户简介与大众点评网完全一致,用户点评也无实质性差异。通过爱帮网,用户能够直接获取到商户的完整简介及大部分用户点评,从而基本满足了获取信息的需求,导致网络用户通常无需再点击大众点评的链接。因此,爱帮网上的商户简介和用户点评实际上构成了对大众点评网相应内容的实质性替代,严重损害了汉涛公司的合法权益。这一案例典型地展示了反不正当竞争法视角下电商平台数据抓取纠纷的模式。在此种情景下实质替代性原则的适用对于判定行为的违法性和不正当性十分有效。然而互联网产业的一大特性在于不同营业领域之间的门槛较低,经营者可以快速识别风向并进入一个新营业领域。故而行为人的营业范围或用户群体等在实施抓取技术时可能并不与平台具有替代性,但其实施抓取行为的目的是获取较为成熟的数据信息资源以开展该领域内的营业活动,此时适用实质替代性原则不能有效判定该行为的不正当性。

3.2.3. 商业道德

由于当前法律体系中缺少直接针对数据抓取行为不正当竞争的具体规定,法院在实践中往往会引用

《反不正当竞争法》的第二条，该条规定：“经营者在生产经营活动中，应当遵循自愿、平等、公平、诚信的原则，遵守法律和商业道德。”在这里，诚信原则往往体现在商业道德之中。值得注意的是，电子商务行业的商业道德内涵与传统行业存在差异。考虑到产业互联共通的特征，这里谈到的商业道德最鲜明的特色要考虑开放包容性。在数据抓取领域最直接的表现形式即 Robots 协议，也称为爬虫协议，机器人协议，因 Robots 协议在互联网行业发展得较为成熟，因而实践中常把它作为衡量数据抓取行为违法与否的标准之一。

3.3. 判断行为正当性的考量因素：竞争损害效果

市场竞争的激烈本质决定了参与者在竞争过程中利益的增减是不可避免的，这是市场竞争的固有属性。纵使在有些情况下技术的革新可以带来全行业新商业模式和新的利益增长点，实现“双赢”乃至“共赢”。合理的市场竞争不可避免会产生一方利益受损，这不意味着竞争行为不正当。由损害结果推导数据抓取行为是否正当应该综合考量损害结果的各方面。一方面，第三方的数据抓取行为若造成对电商平台原有数据或技术的破坏、流失、错误等损失，显然未纳入合理利用的范畴，或行为人的抓取行为是服务于其营业内容，该行为导致行为人获得了电商行业内具有重要价值的数据后竞争力增强，与原电商平台争夺交易机会或是提供相似内容或服务导致原平台用户流失或黏度降低，减少了平台通过自身合法的生产经营活动能够取得的预期收益，可视为评估其数据抓取行为是否正当的关键基准之一。

4. 合理裁判规则的构建

在明确数据权益的法律性质后，从反不正当竞争法的角度审视，若第三方数据抓取行为逾越了合理边界，为保护原电商平台的相关权益，在裁判程序中，应当寻找合理路径处理该纷争。

4.1. 数据产品形态时的裁判规则

所谓数据产品形态是一种将数据直接作为商品投入到市场竞争中的模式，例如，消费娱乐信息、商品简介乃至用户评价等均是以数据产品形式呈现的产品。在此类模式下，数据竞争主要体现在经营业务相似或相同的竞争者之间，他们通过运用数据抓取技术来获取相关数据资源，从而展开激烈的横向竞争。最突出的特征就在于行为人与原平台提供的产品或服务之间具有高度的可替代性。故而此种形态下采取实质替代性原则作为裁判规则最为恰当。第三方出于“不劳而获”的恶意竞争目的采取数据抓取行为直接将原平台上的数据产品批量搬运至自己的平台展示。但使用这一规则时也应当注意一定界限，避免将一切主营业务或提供产品和服务相近的经营者实施的数据抓取行为识别为不正当竞争行为。若第三方的数据抓取行为控制在合理范围内，则不应过度苛责。2016 年大众点评诉百度一案，对“合理范围”的标准进行了重要补充和完善。在判断相关行为性质时，应全面考虑以下几个关键要素：首先，需评估利用者的行为是否产生了积极影响，比如是否推动了商业模式创新、提升了用户体验或为消费者提供了更多选择；其次，要判断使用涉案信息是否超出了必要限度，在理想情况下，利用他人资源应遵循“最少、必要”的原则，避免过度采集；最后，还需分析该行为对市场秩序可能产生的影响，特别是是否会导致其他市场主体因担忧成本而减少在信息收集方面的投入[9]。

4.2. 数据要素状态时的裁判规则

数字技术的快速发展以及人工智能的创新使得对于数据信息的利用方式也发生了巨大变化。利用算法技术可以通过用户授权或使用产生的原始基础数据演算出服务于经营业务需求的产品。也就是说，数据的运用方式已从直接的原始利用转变为更具创造性的衍生性使用。这意味着，即便不是处于同一竞争层面或上下游的经营者之间争夺数据这一生产要素，数据抓取者的目的也可能并非简单地将原平台的

数据产品原封不动地复制到自己的网页上。相反，他们更可能是利用原平台的数据，开发出全新的数据产品或服务，以提供给用户，从而创造出衍生产品，并充分挖掘数据的再生价值。这一过程中或许双方的经营业务或是利用抓取技术创造的产品与原产品之间都不存在替代关系，因而无法适用实质替代性规则。针对日益复杂的衍生性数据使用行为，裁决应基于具体情境进行，这其中包括了对“破坏性利用”行为的考量以及“三重授权”规则的适用。

4.2.1. “破坏性利用”规则

根据《反不正当竞争法》第12条，经营者不得利用技术手段妨碍、破坏其他经营者合法提供的网络产品或者服务的正常运行。因此，如果第三方通过数据抓取手段获取数据，并利用这些数据开发的衍生产品干扰或破坏了原平台所提供产品或服务的正常运行，该行为将被视为不正当竞争。司法实践中，对于“破坏”的边界把握应当牢记网络经济的特殊属性，数据资源本身具有开放性与共享性，如果第三方经营者进行数据抓取，并利用原平台的数据资源进行经营，但在这一过程中，并未对原平台的数据资源造成破坏性利用，也未违反相关法律规定，同时能够为用户提供全新的体验，那么这种行为通常不应被认定为不正当竞争。¹在实践中对于破坏行为的认定目前主要根据行为人的数据抓取行为是否突破了平台设置的限制措施，例如，绕过平台设定的登录障碍、IP访问频率限制、验证码验证等技术手段，或违反Robots协议进行数据抓取的行为。²探究其破坏性背后的原理在于，正是因为行为人的超限数据抓取行为可能会导致平台正常的访问运行机制被破坏，而平台对此只能投入更多的成本与精力与其对抗这无疑就增加了平台的运行成本；或者超限度的数据抓取行为对于数据的浏览阅读强度之大远超正常用户访问的负担会造成平台服务器工作量激增，有可能引起原平台服务器瘫痪等工作故障等，数据抓取行为也可能造成大量抓取数据后原平台所提供的产品或服务面貌全非，此时也应当被认定为具有破坏性。

4.2.2. “三重授权”规则

在淘友天下与微梦案中，法院明确了涉及个人信息数据获取的“三重授权”原则。该案判决指出，当第三方通过开放数据接口(OpenAPI)获取用户信息时，必须遵循“用户授权”+“平台授权”+“再次用户授权”的流程。具体而言，数据提供方需首先获得用户的明确同意，才能向第三方开放数据。同时，第三方平台在使用这些用户信息时，必须向用户清晰说明使用的目的、方式和范围，并再次获得用户的明确同意。³“三重授权”原则最初是应用于OpenAPI模式的，在此模式下，第三方需与平台签订协议，然后通过平台所提供的开放接口来获取数据。然而，“三重授权”规则目前主要局限于用户身份信息的应用，对于用户行为信息的适用性较为有限。故而笔者认为，只有在涉案数据的内容更倾向用户的个人信息，有可能造成个人信息安全遭到威胁的情景下才能适用这一高标准裁判规制。

5. 结语

随着数据资源对于数字经济的重要性日益凸显，电商行业内掀起的对数据的狂热追逐在正向上刺激了包括数据抓取技术在内的多种新兴技术的出现。然而为了在竞争博弈中获取优势地位，经营者一旦恶意使用数据抓取技术势必会导致正常有序的市场秩序被破坏。本文首先厘定了数据资源的法律属性，分析数据抓取行为的正当性界限，避免中立技术或是对数据资源合理利用的行为进入诉讼程序，减少无谓的纠纷，平衡个人隐私保护与刺激市场竞争活力，最后在裁判环节提供了处理相关纠纷的规则思考。通过完善补充现有法律框架下对于此类纠纷处理难题的不足，框定使用数据抓取技术的合理范围，不仅能够保护数据权益者的正当利益，也能够实现数据在更广泛的范围内被合理收集、开发、利用。

¹ 杭州铁路运输法院(2019)浙8601民初1987号民事判决书。

² 北京知识产权法院(2019)京73民终3789号民事判决书。

³ 北京知识产权法院(2016)京73民终588号民事判决书。

参考文献

- [1] 程啸. 论公开的个人信息处理的法律规制[J]. 中国法学, 2022(3): 82-101.
- [2] 王利明. 论数据权益: 以“权利束”为视角[J]. 政治与法律, 2022(7): 99-113.
- [3] 吴汉东. 无形财产权基本问题研究[M]. 北京: 中国人民大学出版社, 2013: 535.
- [4] 张新宝. 论作为新型财产权的数据财产权[J]. 中国社会科学, 2023(4): 144-163+207.
- [5] 钱子瑜. 论数据财产权的构建[J]. 法学家, 2021(6): 75-91+193.
- [6] 龙卫球. 数据新型财产权构建及其体系研究[J]. 政法论坛, 2017, 35(4): 63-77.
- [7] 北京知识产权法院(2016)京 73 民终 588 号民事判决书[Z].
- [8] 梅夏英. 数据的法律属性及其民法定位[J]. 中国社会科学, 2016(9): 164-183+209.
- [9] 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.