

算法“黑箱”的法律规制：电子商务中的消费者权益保护研究

何彬宇

浙江理工大学法学与人文学院，浙江 杭州

收稿日期：2025年12月22日；录用日期：2025年12月31日；发布日期：2026年1月23日

摘要

在电子商务领域，算法的深度应用在提升效率的同时，也因其决策过程不透明而形成了“算法黑箱”。这对消费者权益构成了系统性威胁，并催生了价格歧视、信息茧房及搜索操纵等新型风险。本文通过剖析算法“黑箱”的风险呈现，指出我国现行法律框架在应对时，面临着规制工具滞后、信息不对称下的举证困境等难题。为此，本文尝试构建一个以“过程透明”与“结果公平”为双重支柱的法律规制新路径。该路径主张，通过从事前备案评估、事中分层解释到事后独立审计的全链条机制强化算法透明度；同时，以明确反歧视原则、保障消费者人工干预权和探索举证责任转移等方式，确立算法的实质公平性。最终旨在为实现技术创新与消费者权益保护的动态平衡提供可行的法律对策。

关键词

算法黑箱，消费者权益，算法透明度，算法公平性，法律规制

Legal Regulation of Algorithm “Black Box”: A Study on the Protection of Consumers’ Rights and Interests in E-Commerce

Binyu He

School of Law and Humanities, Zhejiang Sci-Tech University, Hangzhou Zhejiang

Received: December 22, 2025; accepted: December 31, 2025; published: January 23, 2026

Abstract

In the realm of e-commerce, while the profound application of algorithms enhances efficiency, it has also created an “algorithmic black box” due to its opaque decision-making processes. This poses a

文章引用：何彬宇. 算法“黑箱”的法律规制：电子商务中的消费者权益保护研究[J]. 电子商务评论, 2026, 15(1): 112-119. DOI: [10.12677/ecl.2026.151015](https://doi.org/10.12677/ecl.2026.151015)

systemic threat to consumer rights and gives rise to new risks such as price discrimination, “information cocoons,” and search manipulation. By analyzing the risks presented by the “algorithmic black box,” this paper points out that China’s current legal framework faces challenges in response, including lagging regulatory tools and difficulties in evidence presentation under extreme information asymmetry. To address this, the paper proposes a new legal regulatory path built on the twin pillars of “procedural transparency” and “result fairness.” This path advocates for strengthening algorithmic transparency through a full-chain mechanism encompassing pre-assessment and filing, tiered in-process explanations, and post-hoc independent audits. Concurrently, it aims to establish substantive fairness by clarifying anti-discrimination principles, guaranteeing consumers’ right to human intervention, and exploring a shift in the burden of proof. The ultimate goal is to provide a viable legal strategy for achieving a dynamic balance between technological innovation and the protection of consumer rights.

Keywords

Algorithmic Black Box, Consumer Rights, Algorithmic Transparency, Algorithmic Fairness, Legal Regulation

Copyright © 2026 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

数字经济时代，算法已成为驱动电子商务运行的核心引擎。从个性化推荐到动态定价，从供应链优化到售后服务匹配，算法深度渗透于电商交易的全流程，根据《中国数字经济发展报告(2024)》，我国超85%的主流电商平台已实现核心业务算法化¹。然而，算法在高效赋能商业创新的同时，其内在的复杂性与不透明性催生了新型消费者权益风险，这些风险的根源普遍指向“算法黑箱”这一核心难题。

“算法黑箱”使得决策过程被技术壁垒与商业秘密保护所隐匿，消费者往往在不知情中接受算法的预设安排：它为“大数据杀熟”等价格歧视行为提供了技术温床，据中国消费者协会2023年投诉数据统计，算法相关投诉量同比增长41.3%，其中价格歧视类投诉占比达67.2%²；通过“信息茧房”效应无形中限制消费者的自主选择，导致消费者知情权与公平交易权受到前所未有的侵蚀。当权益受损时，消费者因无法穿透“黑箱”而陷入举证困境，传统法律救济途径步履维艰。

我国《中华人民共和国个人信息保护法》《中华人民共和国电子商务法》等虽已就自动化决策作出原则性规定，但在规制实践中仍面临标准模糊、解释权虚化等难题^[1]。鉴于此，本文系统剖析算法“黑箱”对消费者权益的风险呈现，审视现有法律规制的适用性与困境，构建以“过程透明”与“结果公平”为双重支柱的法律规制路径，为实现技术善治、平衡创新发展与权益保护提供可行的法律对策。

2. 算法“黑箱”对消费者权益的风险呈现

2.1. 算法“黑箱”的成因与特征

算法“黑箱”是对算法运作逻辑外部不可见性的客观描述，其形成与演化是技术特性、商业诉求与监管局限共同作用的结果。算法“黑箱”并非一个纯粹的技术贬义词，而是对其内在运作逻辑外部不可

¹参见国家数据局发布《数字中国发展报告(2024年)》。

²参见《2023年全国消协组织受理投诉情况分析》，中国消费者协会，2024年。

见性的一种客观描述。理解其成因与特征，是准确评估其法律风险的前提。

首先，技术内在复杂性构成“黑箱”的核心壁垒。现代电商平台广泛采用的深度学习模型，决策逻辑并非由人类通过明确规则代码预设，而是在海量数据训练中自主学习演化形成。一个复杂的电商推荐算法可能包含数亿个参数，参数间的相互作用机制即便算法设计者也难以完全解构，这种“知其然不知其所以然”的技术特性，形成了“黑箱”最坚硬的内核。

其次，商业秘密保护需求强化了“黑箱”的封闭性。算法模型及训练数据集是电商平台投入巨额成本研发的核心资产，据艾瑞咨询测算，头部电商平台年均算法研发投入占营收比重超 8%³。平台为维护市场竞争优势，必然以商业秘密为由拒绝披露算法细节，在技术壁垒之外加筑了法律与商业的双重屏障。

最后，数据驱动的动态演化特性加剧了“黑箱”的隐蔽性。算法会根据实时用户行为数据持续优化调整，某电商平台的定价算法日均迭代次数可达数十次，这种动态性使得静态审查难以捕捉其真实运行状态，进一步放大了“黑箱”的规制难度。

2.2. 对消费者公平交易权的危害

公平交易权是消费者权益的核心，而算法“黑箱”催生的“大数据杀熟”等价格歧视行为，对该权利构成严重侵蚀。传统合法的价格差异化策略(如学生优惠、老年折扣)具有公开性与合理性，而算法驱动的价格歧视则是基于个人画像的不透明非对称定价。

其实现机制表现为：平台通过算法持续追踪分析用户消费历史、浏览偏好、价格敏感度、设备型号等数据，构建精准的“支付意愿画像”^[2]。据某第三方检测机构 2023 年实验数据，同一商品在不同用户账号中的价差最高可达 35%，使用高端设备、高频消费的用户更易遭遇高价待遇⁴。当消费者交易时，算法瞬时调用画像，对价格不敏感用户报出高价，对新用户或敏感型用户提供优惠，整个过程后台自动完成，消费者无从知晓价格的“量身定制”属性。

这种行为严重违反《中华人民共和国消费者权益保护法》确立的公平诚信原则，将数据资源异化为牟取超额利润的工具^[3]。算法“黑箱”在此扮演关键角色：决策过程的不透明使平台得以隐蔽实施价格歧视，被质疑时又能以“市场波动”“系统差异”搪塞，导致消费者维权困难。

2.3. 对消费者知情权与自主选择权的危害

消费者的知情权与自主选择权是交易达成的先决条件，前者保障消费者获得全面、真实的信息，后者则确保其能够基于自由意志做出选择。算法“黑箱”通过个性化推荐和搜索排序两种主要方式，对这两项权利构成了双重架空。

2.3.1. “信息茧房”的构建

个性化推荐算法的初衷是通过预测用户偏好，提升信息匹配效率。然而，其“优化用户参与度”的单一目标导向，极易导致“信息茧房”的产生。算法通过不断分析用户的点击、浏览和购买行为，持续向其推送相似或相关的内容。久而久之，用户的信息流被算法精心过滤和塑造，视野被局限在一个由其过往兴趣所构筑的封闭空间内。

这对自主选择权的损害是深远的。真正的自主选择，前提是拥有足够多样化的选项。但在“信息茧房”中，消费者的选择权会受到实质性的侵害^[4]。因为算法并非呈现一个广阔的、多样化的市场供消费者探索，而是构建了一个看似丰富、实则同质化的“回音室”。消费者以为自己是在自由选择，实际上却是在算法预设的、狭窄的轨道上滑行。这种对选择范围的无形限制，比强制交易更为隐蔽，也更难被察

³参见《2023 年中国电商算法应用发展报告》，艾瑞咨询，2024 年。

⁴同脚注 3。

觉。同时，消费者的知情权也在此过程中被削弱，他们对于整个市场的全貌、对于不同品牌和产品的多样性认知，都在不知不觉得被算法所塑造和窄化。

2.3.2. 搜索排序与搜索结果的操纵

搜索排序算法的操纵是对消费者决策的“硬性”干预。据市场监管总局2023年执法数据，32%的电商不正当竞争案件涉及搜索排序操纵⁵。例如，在2021年阿里巴巴反垄断处罚案中⁶，调查显示平台利用算法规则控制商品流量分配，对未执行“二选一”的商家进行搜索降权。这种“自我优待”或“信誉压制”，将商业干预包装成客观中立的搜索结果，直接剥夺了消费者做出真实判断所需的信息基础。消费者无法区分自然排名与人工干预，其基于信赖做出的交易决策实际上是被诱导的结果。

由此可见，对搜索算法的操纵行为会加剧消费者与经营者之间的信息不对称，直接侵害了消费者的知情权。消费者有权知道搜索结果的排序逻辑，有权区分广告与自然搜索。当算法“黑箱”将商业干预包装成客观中立的结果时，消费者就被剥夺了做出真实、无偏见判断所需的信息基础，其自主选择也就无从谈起。

3. 现行法律规制框架的适用与困境

3.1. 《中华人民共和国消费者权益保护法》：原则性保障与规制乏力

作为消费者保护的基础性法律，《消费者权益保护法》(以下简称《消法》)提供了原则性的权利宣言，但在具体适用中却显得捉襟见肘。

首先，《消法》确立的知情权、自主选择权与公平交易权等核心原则，无疑为抵制算法滥用提供了根本的法理依据。理论上，算法价格歧视是对公平交易权的直接违背，信息茧房是对自主选择权的无形架空，而算法决策过程的不透明本身就是对消费者知情权的挑战。这部法律成功地勾勒出了消费者权利的理想边界。

然而，从理想边界到现实保护之间，横亘着一道由算法“黑箱”造成巨大鸿沟。《消法》的困境根植于其诞生年代的局限性，它所设想的交易场景是买卖双方信息相对透明、交易条件相对明确的“前算法时代”。因此，其权利条款的抽象性与算法行为的隐蔽性形成了尖锐的矛盾。平台总能以市场波动、成本差异等看似合理的商业逻辑进行辩护，而外部观察者因无法窥探“黑箱”内部而无从反驳。同样，法律保障了选择的自由，但如何界定个性化推荐何时从“贴心服务”异化为“非法限制”，仍成问题。

更为致命的是，《消法》所依赖的“谁主张，谁举证”的传统民事诉讼原则，在算法场景下几乎完全失效。消费者若要证明“大数据杀熟”，需获取平台后台数据、算法代码等核心信息，但这些均属平台严格保护的商业秘密^[5]。中国消费者协会2023年调研显示，算法相关投诉的胜诉率不足15%，其中举证困难是主要障碍，这直接导致消费者权利被技术性悬置⁷。

3.2. 《中华人民共和国电子商务法》与《中华人民共和国反不正当竞争法》： 视角偏移与间接保护

《中华人民共和国电子商务法》与《反不正当竞争法》虽在特定层面触及算法规制，但其立法视角决定了其对消费者个体权益的保护是间接且有限的。这两部法律的重心在于维护宏观的市场竞争秩序，防止平台利用其优势地位进行不正当竞争，如通过算法操纵搜索结果以实现“自我优待”，或利用技术手段排挤竞争对手。当算法行为同时损害了市场公平竞争和消费者利益时，这些法律便能发挥作用，监

⁵参见《2023年反垄断与反不正当竞争执法报告》，市场监管总局，2024年。

⁶参见国家市场监督管理总局：《行政处罚决定书》(国市监处〔2021〕28号)，2021年4月10日发布。

⁷同脚注2。

管机构可以据此对平台进行调查和处罚。

然而，这种“公共执法”模式的视角，决定了其保护的落脚点是市场整体，而非具体的消费者个体。其目标在于纠正市场失灵，而非弥补个体损失。一个因算法推荐而错失了更优选项，或因价格歧视多付了几十元钱的消费者，其个体损失虽然微小，却难以触发这些法律的救济机制。例如，《中华人民共和国电子商务法》第三十五条禁止平台不合理限制交易⁸，但未明确算法操纵是否属于“不合理限制”^[6]；《中华人民共和国反不正当竞争法》规制平台“自我优待”“数据掠夺”等行为，如2023年阿里巴巴因算法驱动的不正当竞争行为被处罚182.28亿元⁹。但这些法律的“公共执法”模式，落脚点是市场整体而非个体消费者，个体因算法受损的小额损失难以触发救济机制^[7]。因此，两部法律虽为市场划定了竞争“红线”，却未给消费者提供直接维权的“利剑”，难以回应日常交易中普遍存在的算法不公问题。

3.3. 《中华人民共和国个人信息保护法》：精准回应与解释困境

作为我国应对算法挑战的“锐利武器”，《个人信息保护法》(以下简称《个保法》)实现了从原则宣示到精准规制的重要跨越，但其核心条款在实践中仍面临深刻的解释与执行难题。

首先，《个保法》第二十四条关于“自动化决策”的规定¹⁰，是我国在算法规制立法上的里程碑。它直面“黑箱”问题，赋予了个人对自动化决策的知情权和拒绝权，并明确要求平台保证决策的“透明度”和“结果公平、公正”，更直接禁止了“不合理的差别待遇”。这无疑为解决“大数据杀熟”等顽疾提供了迄今为止最强有力的法律工具。但其核心条款在实践中面临落地困境。法律要求平台对自动化决策作出说明，但实践中平台多以“基于您的偏好推荐”等模糊表述回应，对消费者理解决策逻辑毫无帮助。若要求过度披露，又会触及商业秘密，二者的平衡缺乏明确指引，导致解释权虚化^[8]。

其次是“公平、公正”标准的虚置之困。法律要求结果的“公平、公正”，但“公平”本身就是一个复杂的、多维度的概念。而在算法领域，算法公平存在个体公平、群体公平等多种技术定义，且可能相互冲突。一个在统计学上对所有群体都“公平”的算法，仍可能对特定个体造成不公。在缺乏可操作的、具体的衡量标准和技术规范的情况下，法律文本中的“公平、公正”更像是伦理宣示，而非司法裁判的精确标尺。

再者是“不合理”差别待遇的界定之困。法律禁止的是“不合理”的差别待遇，默许了“合理”差异的存在，但未明确界定标准，平台常以“用户画像差异”“消费场景不同”为由抗辩，给监管与司法带来裁量难题。综上，现行法律框架已形成初步围堵之势，但《消法》的原则性乏力、竞争法的间接保护、《个保法》的解释困境，共同揭示了现有法律工具箱对算法这一新型规制对象的系统性不适应，亟需构建更具针对性的规制路径，以“预防”与“惩治”相结合的方式推动形成政府、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^[9]。

4. 算法“黑箱”法律规制路径的构建

4.1. 强化算法透明度义务

透明度是撬动算法“黑箱”的核心杠杆，法律应追求“有意义的透明”，建立分层级、多维度的透明

⁸《中华人民共和国电子商务法》第三十五条 电子商务平台经营者不得利用服务协议、交易规则以及技术等手段，对平台内经营者在平台内的交易、交易价格以及与其他经营者的交易等进行不合理限制或者附加不合理条件，或者向平台内经营者收取不合理费用。

⁹同脚注2。

¹⁰《中华人民共和国个人信息保护法》第二十四条 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，不得对个人在交易价格等交易条件上实行不合理的差别待遇。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。通过自动化决策方式作出对个人权益有重大影响的决定，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。

度义务体系，平衡知情权与商业秘密保护。因此，法律规制应追求一种“有意义的透明”，即根据不同阶段和不同对象的需要，建立一个分层级、多维度的透明度义务体系，实现从“不可见”到“可解释”的转变。

4.1.1. 事前规制

规制的起点应前移至算法上线应用之前，建立起一道“预防性”的防火墙。具体而言，可借鉴欧盟《数字服务法》和我国金融领域的监管经验，将算法规制由事后转向事前，建立算法备案与影响评估制度。例如，欧盟《数字服务法》明确要求高风险平台算法需履行备案与评估义务，这为我国提供了成熟的制度参考。

此外，建立高风险算法备案登记制度也迫在眉睫。这要求明确高风险算法范围，涉及个性化定价、搜索排序、敏感个人信息处理的算法。电商平台需在算法上线或重大更新前，向市场监督管理总局等主管部门提交备案，材料包括算法功能、应用场景、数据类型、决策逻辑简要说明及安全自评估报告，实现监管“家底清、风险明”。

同时还需要强制开展算法影响评估。高风险算法备案时，需同步提交详尽评估报告，内容可涵盖算法设计目标、数据集来源与偏见分析、潜在风险预判、减损措施、可解释性方案及救济渠道等多方面。通过该制度倒逼平台将合规与伦理考量内嵌于算法设计，实现“设计向善”，从源头上预防和减少潜在的权利侵害。

4.1.2. 事中控制

针对《中华人民共和国个人信息保护法》说明义务的虚化问题，可引入计算机科学中的“局部可解释性”概念作为法律标准。根据计算机科学研究，虽然深度学习模型的全局可解释性(即理解整个模型的所有逻辑)很难实现且涉及商业秘密，但局部可解释性(即解释特定输入如何导致特定输出)在技术上已成熟。基于此，法律若确立“反事实解释”的义务标准，则更有利于“解释权”透明度的落实。在此基础上，还应构建一个分层级的解释机制，以平衡消费者的知情权与平台的商业秘密保护。

第一层级是对所有消费者的普适性、简化解释。平台应以显著、简明易懂的方式，在服务界面中向用户主动公示其数据，包括对算法代码、数据、决策树等信息进行公开，且就算法决策定义、原理、过程等作出说明[10]，以让消费者对算法的存在和基本逻辑有直观认知。

第二层级是对个体消费者的、针对具体决策的“可追溯”解释。当自动化决策对消费者个人权益产生重大影响时，或在消费者提出明确要求时，平台必须提供更具针对性的解释。这种解释不要求披露源代码，但应能清晰地说明“导致该决策的主要影响因素及其相对权重”，以便为消费者理解自身处境、提出有效申诉和修正错误信息提供了可能。

4.1.3. 事后监督

仅依靠平台的自我声明和被动解释是远远不够的，必须引入外部的、独立的监督力量，对算法的实际运行情况进行定期“体检”，比如成立专业的第三方机构进行审查监督，因为专业性社会组织在破解算法不透明性层面往往有着独特优势[11]。法律应明确，运营高风险算法的大型电子商务平台，有义务定期(如每年或每两年)委托具有资质的、独立的第三方机构对其关键算法进行合规性与伦理风险审计。此外，审计机构应独立出具审计报告，其摘要应向社会公布。这一制度通过引入市场化的专业监督力量，既能弥补政府监管机构技术能力的不足，又能通过声誉压力倒逼平台持续进行算法优化和合规整改。

4.2. 确立算法公平性审查标准

透明度解决了“看得见”的问题，而公平性则要解决“对不对”的问题。算法公平性的审查，应坚持

权利本位的原则,不仅要关注统计学意义上的结果公平,更要确保每一个体消费者的基本权利不受侵犯。

法律应在《中华人民共和国个人信息保护法》“不合理差别待遇”的基础上,进一步细化和明确算法应用中禁止歧视的实质性原则。首先,应明确列举受保护的特征,除了传统的种族、性别等,还应纳入数字时代特有的敏感特征,如地理位置(地域歧视)、设备型号(终端歧视)、过往消费水平(客户忠诚度歧视)等。任何基于这些受保护特征,对处于同等交易条件下的消费者实施差别待遇的行为,除非有法律明确规定或具有绝对正当的理由,否则均应被推定为“不合理”。其次,应确立“基本生活服务普惠”原则,即对于网络约车、线上购药、基础食品外卖等关乎民生的基本消费场景,应严格限制平台利用算法进行差异化定价的能力,保障所有消费者都能以公平、非歧视性的价格获取基本服务。

4.3. 构建协同治理机制: 强化规制实施保障

上述制度路径的实现,离不开强大的监管能力和平台内生的治理动力。法律规制的设计,必须辅之以协同机制的建设。一方面,应大力提升监管机构的技术赋能水平,推广“监管科技”的应用。监管部门应组建包含数据科学家、算法工程师和法律专家的跨学科团队,开发自动化工具对平台的算法进行在线监测和“沙盒”测试,实现以技术对抗技术。另一方面,应进一步压实电子商务平台的主体责任,推动平台建立健全内部的算法治理体系,设立专门的算法伦理委员会,将算法的公平、透明、无歧视原则内化为企业文化和产品开发流程,从被动合规转向主动担责。

总之,通过构建“事前备案评估、事中分层解释、事后独立审计”的透明度义务链条,辅之以“明确反歧视原则、保障人工干预权、转移举证责任”的公平性审查体系,并在监管能力和平台责任上提供支撑,我们有望逐步破解算法“黑箱”的规制难题,最终在一个更安全、更公平、更值得信赖的数字消费环境中,实现技术创新与消费者权益保护的和谐共生。

5. 结论

算法“黑箱”作为电子商务时代的核心技术特征,其在提升效率的同时,也因内在的不透明性深刻动摇了传统的消费者权益保护根基[12]。本文通过审视发现,现行法律框架虽已作出回应,但在规制工具的精准性与有效性上仍显不足,难以穿透技术壁垒,解决信息不对称下的举证困境。

为此,构建一个以前瞻性法律规制路径为核心的治理体系势在必行。该路径主张,应以“过程透明”与“结果公平”为双重支柱:通过引入算法备案评估、分层解释、独立审计等机制强化透明度;并以明确反歧视原则、保障人工干预权、转移举证责任来确立公平性标准。最终,在一个法律、技术与伦理协同共治的框架下,方能实现技术创新与消费者权益保护的动态平衡,促进数字经济的健康与可持续发展。

参考文献

- [1] 强若晨. 消费者权益保护视野下的算法歧视行为规制研究[D]: [硕士学位论文]. 无锡: 江南大学, 2023.
- [2] 岳蓝馨. “算法黑箱”背景下平台个性化推送的法律规制[D]: [硕士学位论文]. 武汉: 中南财经政法大学, 2024.
- [3] 王利明. 算法时代消费者权益保护的挑战与应对[J]. 中国法学, 2023(2): 123-140.
- [4] 郭明瑞, 娄逸骅. 算法消费者信息茧房的治理困境与制度因应[J]. 深圳大学学报(人文社会科学版), 2025, 42(4): 86-95.
- [5] 李佳羽, 刘子聪. 大数据“杀熟”的规制困境与协同治理路径[J]. 海南开放大学学报, 2025, 26(4): 96-102.
- [6] 王鹏飞. 电商平台算法歧视法律救济的困境与出路[J]. 科学决策, 2023(5): 213-223.
- [7] 黄锫. 数字平台算法侵害的行政法律责任论纲[J]. 比较法研究, 2023(6): 184-198.
- [8] 汪庆华. 算法透明的多重维度和算法问责[J]. 比较法研究, 2020(6): 163-173.
- [9] 周汉华. 为个人信息保护搭建法治屏障[J]. 中国网信, 2023(11): 58-61.

-
- [10] 贾慧智. 算法治理重在“透明” [J]. 法人, 2022(4): 18-21.
 - [11] 张卫涛. 算法黑箱的成因、危机及治理路径选择[C]//中国智慧工程研究会. 2025 数字时代的社会结构变迁与治理创新学术交流会论文集(下). 北京: 中国公安大学出版社, 2025: 374-377.
 - [12] 刘建, 吴理财. 算法治理的黑箱及规制: 基于治理界面的视角[J]. 学习与实践, 2025(11): 34-45.