



算法透明机制的困境与克服：聚焦公众信任的机制重构

热点聚焦

□ 戴维 王锡梓

数字时代，算法已深度介入社会生活的方方面面，在给人们生活带来便利的同时，其背后的“算法黑箱”问题也日益凸显。“算法黑箱”源于算法技术的复杂性以及算法知识的高壁垒，使其运行逻辑和决策理由难以被公众所观察和理解。鉴于此，学界发展出以技术披露和决策解释为核心的算法透明模式，以打开“黑箱”。但透明模式未能充分考虑算法可解释性障碍和公众认知局限，在实践中面临来自技术和制度的双重挑战，可能导致透明机制的失效和透明度期待落空。

算法透明机制的理论逻辑与立法演进

透明机制一向被视为解决“决策黑箱”的有效手段。传统透明理论认为，通过向公众披露决策程序，论证决策合理性，可以保障公民知情权的实质性权利，并促进决策机关的可问责性。算法加剧了决策权力的封闭性和秘密性，以及决策权力和公民权利关系的失衡。学者们尝试以包含事前技术披露和事后说明理由的双线透明模式，以防范算法权力的滥用。

在规范层面，早期算法透明机制主要借助个人信息保护和信息公开立法来间接建构，但效果不佳。例如，我国不少学者尝试从个人信息保护法中推导出算法告知义务和算法解释权，但这需要多个条款“嫁接合成”，规范基础复杂且要求模糊，不便适用。有观点认为算法本质仍为行政规则，应当根据政府信息公开条例予以公开。但算法代码能否直接等同于立法文件而适用信息公开规则存疑。

近年来，随着人工智能立法兴起，算法特别是机器学习类算法，有了更具针对性的规范基础。欧盟《人工智能法》根据人工智能风险程度设置四类透明义务，包括：第一，针对高风险人工智能系统的技术信息披露义务，服务于人工智能系统设计者、提供者以及监管者；第二，针对特定人工智能系统的标识和告知义务，服务于公众；第三，针对高风险人工智能系统的登记义务，服务于公众；第四，针对高风险人工智能系统决策建议稿都将透明原则作为一项基本原则，要求提供者向公众告知算法基本信息、监管信息、用户权利及救济途径等；并规定当人工智能的产品或服务对个人权益有重大影响时，个人有权要求提供者解释说明。

算法透明机制的规范表达及其现实困境

尽管算法透明机制在理论和立法上有所发展，但仍面临诸多问题。

首先，立法设置的透明标准低于理论预期。在技术披露方面，欧盟《人工智能法》不要求算法提供者向公众提供技术材料，而是要求提供关于算法的说明、描述或摘要。信息的可靠性、准确性和真实性存疑，可能被有意误导公众；在算法解释权方面，立法未能将解释权行使提供充分的制度保障。

其次，算法可解释性技术瓶颈制约算法透明机制。可解释人工智能(XAI)技术虽有进展，但普遍性和准确性不足。模型越大，层级越复杂，解释效果越差。应用不同的XAI技术于同一模型，可能凸显不同的影响因素或权重。可解释人工智能确实是充满希望的发展方向，但目前它尚不足以彻底打开“黑箱”。此外，算法可解释并不等同于公众可理解，公众需要通俗易懂的语言理解算法模型，进而引发了算法披露在“信息的真实性”与“信息的可理解性”之间的矛盾，使算法透明陷入两难境地。

最后，算法透明当下难以以风险规制为核心的算法治理体系间存在张力。规范意义上的风险划分，并不一定符合公众的风险感知。当公众对某个算法的潜在风险产生疑虑，并不会因为它属于规范意义上的低风险算法，就放弃获得说明解释的诉求。因此，以风险等级区别设置透明义务的方式，可能会在整体上对算法信任产生负面影响。

以公众信任为导向的算法透明机制重构

基于双线透明模式可行性与有效性困境，需要我们打破过往透明机制的思维惯性，正视可解释人工智能技术

术发展现状，明确以公众为受众，重塑算法透明机制。

算法透明机制应以提升公众信任为目标。当下，“可信人工智能”已成为全球人工智能发展的核心要求与原则，算法信任旨在为公众营造安全可信的制度环境，让公众能以更轻松的心态享受先进技术带来的便利。这种信任至关重要，它为算法技术的开发与应用创造了空间。算法信任的提出，还标志着算法治理理念从以技术为中心向以人为本的转变。毕竟，公众才是算法风险的最终承担者。

算法透明能够为公众建立算法信任搭建桥梁。算法信任是公众对算法可信度评价后形成的主观态度。适度的信任对算法生态的健康发展尤为关键；盲目信任会让个体毫无防备地暴露在算法风险之下；过度警惕则会催生不必要的恐慌和抵制情绪。算法透明机制可以成为展示算法可信度的重要窗口，帮助公众建立适当程度的算法信任。

具体而言，直接评估算法的性能、运行状态及目标达成情况，需要具备深厚专业能力，投入大量时间和精力，这一方式更适合专家审议场景。因此，公众算法信任的建立主要依赖对算法可信度的间接评估，即借助知识信任与制度信任，通过“中介背书”，帮助公众形成对算法的信任。

知识信任，亦可称为专家信任，其构建要求对算法研发、部署和评估等重要环节责任主体的身份信息披露以及对专家意见质量的严格把控。算法责任主体的身份信息、团队专业资质、过往研发履历等背景信息，有助于公众形成自己的信任判断。此类披露还能将以往隐于幕后的责任主体推向台前，接受公众的审视与监督。特别是在政企合作的场景中，对于公众而言，保障政企合作的透明度，远比向他们披露算法资料更有实际意义。

完善的算法治理体系，对于增强公众对算法的整体信任起着举足轻重的作用，这便是制度信任的价值。在算法应用场景中，通过制度约束，让公众知晓无论技术提供者的主观意愿如何，其在研发和部署算法时，都必须遵循既定的规范、标准和流程，可以有效安抚公众对算法的不安情绪。当前，算法治理正在向着行政监管、企业自我规制以及社会监督三方协同的综合治理体系迈进。算法透明机制需要与这一治理体系实现妥善衔接。首先，应当加强行政监管中的信息披露。一方面，披露内容需与前置的许可管理、备案审查、行政处罚等算法监管手段紧密衔接；另一方面，要整合分散的监管信息沟通渠道，搭建统一的披露平台。积极运用先进的数据处理技术和创新的可视化手段，以“用户友好”的方式进行信息呈现，为公众提供一个全面“考察”算法产品及其研发部署主体的系统性视角。其次，算法透明机制呼吁企业加强自我披露。科技企业作为前沿技术的掌握者，对技术问题和漏洞具有敏锐的洞察力，在风险控制上具有天然优势。最后，算法信任的建立离不开第三方机构的参与。第三方机构参与算法评估和审计工作，既能弥补行政监管在专业知识和资源上的不足，又能规避企业自我约束过程中可能出现的客观公正性问题，有助于增强公众对算法的信任。

公民知情权



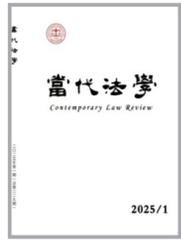
合治理体系迈进。算法透明机制需要与这一治理体系实现妥善衔接。首先，应当加强行政监管中的信息披露。一方面，披露内容需与前置的许可管理、备案审查、行政处罚等算法监管手段紧密衔接；另一方面，要整合分散的监管信息沟通渠道，搭建统一的披露平台。积极运用先进的数据处理技术和创新的可视化手段，以“用户友好”的方式进行信息呈现，为公众提供一个全面“考察”算法产品及其研发部署主体的系统性视角。其次，算法透明机制呼吁企业加强自我披露。科技企业作为前沿技术的掌握者，对技术问题和漏洞具有敏锐的洞察力，在风险控制上具有天然优势。最后，算法信任的建立离不开第三方机构的参与。第三方机构参与算法评估和审计工作，既能弥补行政监管在专业知识和资源上的不足，又能规避企业自我约束过程中可能出现的客观公正性问题，有助于增强公众对算法的信任。

以上重构路径，有望打破算法透明机制的现有僵局，避免“透明但无用”的制度窘境；有助于提升公众在算法治理中的主体性，使其摆脱因知识壁垒而被边缘化的无力感，为公众参与算法监督创造条件；最终形成“公众信任—算法透明”互相促进的正向循环。

(原文刊载于《华东政法大学学报》2025年第1期)

观点新解

冯晓青谈知识产权制度的有效运用——为未来产业创新生态培育提供法律保障



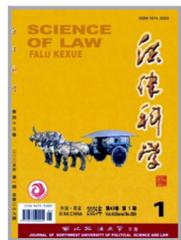
中国政法大学冯晓青在《当代法学》2025年第1期上发表题为《未来产业创新生态培育的知识产权制度因应》的文章中指出：

随着经济社会和科学技术的发展，传统产业被不断改造和升级，新兴产业不断涌现。新兴产业的形成和发展以及传统产业的现代化改造和升级，始终是一个国家或地区经济增长和社会发展的引擎与动力。现代化产业的形成和打造是以面向未来的具有前瞻性、先进性的重大技术变革和创新作为前提和基础的。未来产业是我国发展新质生产力，实现中国式现代化不可或缺的重要新兴产业，其形成和发展离不开一系列重大科技创新，特别是涉及前沿性、先导性技术甚至颠覆性技术的关键核心技术突破。基于此，我国未来产业的发展，需要构建良好的创新生态系统，该系统应具有开放协同性、动态平衡性以及竞争与合作的双重性。

在未来产业创新生态系统构建和创新生态培育中，知识产权制度的有效运用能够发挥其独特的功能和作用，尤其是其内涵的激励创新机制、保护机制、市场机制、促进公平竞争机制和平衡机制，能够有效地促进未来产业领域科技创新和创新成果的转化利用。知识产权制度的有效运用，有利于维护未来产业安全，为未来产业创新生态培育提供坚实的法律保障。

针对现行知识产权制度在鼓励和保护创新、促进创新成果转化利用以及市场机制等方面存在的问题，需要进行制度创新：首先，建立和完善未来产业底层技术知识产权保护规则。由于数据和人工智能技术是未来产业底层技术的根本性支撑，为此，需要建立数据与人工智能知识产权保护制度和规则。其次，建立和完善未来产业新业态、新模式知识产权保护规则。这些新业态和新商业模式是前所未有的，亟待通过构建相关的知识产权保护规则，为未来产业不同领域提供知识产权制度的有力保障。最后，建立和完善未来产业重点领域知识产权保护规则。通过上述规则的制定完善，以更好地优化创新资源配置，为我国建设创新型国家提供坚实的制度保障。

张璋谈对电子商务平台服务协议的规定问题——应发挥服务协议程序性条款的实际效力



上海交通大学凯原法学院张璋在《法律科学》2025年第1期上发表题为《电子商务平台服务协议的行政法功能》的文章中指出：

互联网平台履行行政职责，实现公共目标，已成为数字经济协同治理的重要方式之一。以电子商务平台为例，针对平台内经营者的售假行为，电子商务平台实施的下架商品、限制提现以及扣除保证金等措施在制度效果上与行政处罚相似。然而，电子商务平台针对违规行为的处理措施不同于行政处罚，是以平台内经营者在入驻平台时与平台签订的电子商务平台服务协议为依据，往往缺乏法律基础。

电子商务平台服务协议兼具交易属性与管理属性，交易性规则与管理性规则发挥着不同的功能作用。具有管理属性的“管理性规则”塑造平台的内部秩序。管理性规则使得电子商务平台承担了监管和规范交易行为的职责，从而在电子商务平台内部形成了自上而下的管理与被管理的关系，其发挥着建立规则体系、制裁违规行为以及塑造市场秩序等功能。虽然这些功能与行政法的功能具有一定的一致性，但也存在区别，因此有必要通过平台主体参与行政监管时形成的“国家—市场—社会”三元框架，对电子商务平台履行行政职能的主体地位加以解释。服务协议中的管理性规则可以细化行政法律制度，承接行政执法体系，衍生行政管理权力，促使电子商务平台服务协议从“形式上的协议”转变为“实质上的法”，此亦为电子商务平台服务协议行政法功能的成因。

基于上述成因，针对电子商务平台服务协议的规定问题，应从以下方面应对：其一，拓展服务协议的实质审查路径。通过行政监管制度建立服务协议的实质审查机制，通过电子商务平台内部监督机制的建立确保服务协议的合法性和透明度，通过社会第三方监管为服务协议的实质审查提供条件。其二，发挥服务协议程序性条款的实际效力。程序性条款的制度应当具有灵活性和适应性，能够随着技术发展和创新成果变化而适时调整。其三，构建服务协议纠纷化解方式。司法实践中电子商务平台与平台内经营者之间的纠纷通常被纳入民事诉讼的范畴。其四，形成服务协议的权利控制机制，应明确服务协议规则制定权的权限范围，优化服务协议变更与修改权的监督控制机制，对服务协议执行过程中出现的问题建立问责机制并确保服务协议的内容符合监管要求。

(赵珊珊 整理)

前沿观点

□ 李致君

随着互联网技术的快速发展，基于境外架设的网站和平台进行的跨境网络犯罪呈现高发态势，对国民的人身和财产安全造成了巨大威胁。网络犯罪的特殊性在于，其发生场所从现实空间转向了网络空间，电子证据成为网络犯罪中的重要物证。然而，境外实体掌握的电子证据具有所有权模糊、提取手段复杂、时效性强等特点，传统的跨境电子取证模式因行政成本高、耗时长等缺陷，难以对网络犯罪进行及时有效的打击。因此，我国跨境网络犯罪治理体系亟须因势而变，以更高效地应对跨境网络犯罪的挑战。

传统跨境网络犯罪治理体系的困境

在司法实务中，我国现阶段的跨境网络犯罪治理体系主要面临以下问题：

(一)传统司法协助体系的低效率

基于国际条约的传统司法协助体系虽然为打击传统跨境犯罪提供了基础，但其程序冗长、流程缓慢，难以适应网络犯罪的高效性和跨国性。特别是在Web3.0时代，跨境网络犯罪往往涉及多国管辖和海量电子证据的移交，传统司法协助体系的效率缺陷尤为突出，严重影响了我国打击跨境网络犯罪的力度。

跨境网络犯罪治理的困境与规范化路径

(一)数据控制者的消极配合
各国间罪刑标准的不同，导致同一行为是否构成犯罪存在冲突。例如，开设赌场在我国构成犯罪，但在某些周边国家则不被视为犯罪。这使得我国对境外网络赌博网站的查处难以获得境外主管机关的司法协助。此外，跨境网络犯罪的影响地通常与服务器架设地不同，当地主管机关缺乏打击此类犯罪的内生动力。

(二)公私协作模式的程序合法性问题
网络犯罪的数字化和全球化特征使各国认识到传统国际刑事司法协助制度的局限性，因此积极探索跨境数据取证的公私合作模式。然而，这种模式在程序构建和理论依据上尚不完善，证据收集合法性随着理论界观念的变化和实务需求不断调整，程序合法性备受争议。

跨境网络犯罪治理的规范化路径

为针对性解决跨境网络犯罪的治理难题，实现对跨境网络犯罪的高效打击，应从创新司法协助途径的角度出发，结合现有国际司法协助体系，构建公私协作、多方参与的综合治理体系。

(一)规范化司法协助与电子证据固定流程
规范化跨境网络犯罪治理中的司法协助流程与所需材料，并制定互相认可的电子证据固定和移交程序，是提高司法协助效率的重要举措。具体包括：建立适应涉外网络犯罪的程序合法性问题，简化司法协助的行政成本，明确标准和术语以降低错误率，明确审查要求，最终保障国际间网络犯罪主管机关程序的稳定衔接

和效率提升。

(二)加强公私协作体系中配债机构的监督和培训
对于网络服务提供商等配债机构，应在立法层面积极推进相关资质申请体系与培训体系建设，有针对性地预防和规制侦查权力的不当使用，不断完善监督体系，促进传统主管单位与配债机构之间的协调配合。在司法实践层面，应总结公私协作中的典型特征和突出问题，明确所提取证据的审查要求，促成公私程序之间的无缝衔接。

(三)规制单向跨境电子取证制度
单向跨境电子取证制度是我国现阶段重要的境外电子取证途径，具有高效、便捷的特征，但也存在不尊重他国数据主权的隐忧，标准引发争议。主管机关在进行此类证据提取活动时，应严格遵照《公安机关办理刑事案件电子数据取证规则》，只能提取公开发布的电子数据，不得擅自获取受保护的数据。

(四)建立、加入打击网络犯罪的国际组织，提升国家间互信

通过“一带一路”倡议、国际刑警组织等区域性合作平台，加强国家间在打击网络犯罪领域的合作，简化审查程序，提升各国主管机关的积极性。同时，积极参与《联合国打击网络犯罪公约》等国际条约的制定与实施，凝聚全球共识，推动更深层次的司法合作。

结语

解决传统司法协助体系的效率缺陷，既不能违背尊



重数据主权的原则，也不能以降低电子证据的可靠性为代价。随着国际交流的不断加深，多维度、多主体的跨境网络犯罪打击体系是符合时代特征的治理路径。探索具有中国特色的网络犯罪治理道路，需规范并改进传统的司法协助流程，基于国际条约和国际组织对分布在多国的跨境网络犯罪采取多方共管、多维度治理的方式，从而进一步普适国际化时代下的网络犯罪治理模式，以保护网络社会安全与数据主权，方能满足人民群众的网络安全需求，保障人民财产不受侵害。

行政争议调解制度创新路径

前沿关注

□ 马文哲

行政争议调解作为化解行政纠纷的重要方式，其制度创新直接关系到社会治理效能的提升和法治政府建设的推进。当前，随着经济社会的快速发展，行政争议日趋复杂多元，这对调解制度的科学性、规范性和效能性提出了更高要求。

完善制度设计：提升调解工作的科学性和规范性

完善制度设计的核心目标是建立科学规范的调解工作体系，实现调解程序的标准化和调解方式的多元化。

建立科学高效的分类调解机制是首要任务。要以提高调解效率为导向，构建科学的行政争议分类标准体系。根据争议主体、争议对象、争议性质等要素，建立行政争议分类目录，实现争议类型的精准识别和科学分类。针对不同类型的行政争议，设计差异化的调解规程和工作流程。

完善调解前置程序是重要环节。要以促进行政争议高效解决为目标，科学界定调解前置的适用范围。建立调解前置与行政复议、行政诉讼的有效衔接机制，确保当事人救济权利得到充分保障。对于涉及国家利益和社会公共利益的重大行政争议，建立调解前置豁免制度，保持程序设置的灵活性和科学性。

创新调解方式方法是关键突破。要以提升调解效能为目标，积极推进调解方式的创新发展。充分运用现代信息技术，建设智能化调解平台，实现调解工作的在线化、数字化、智能化。开发电子档案管理系统，

推进调解案件信息的标准化管理和高效利用。建立健全调解案例库，加强典型案件的总结提炼和经验推广，为类似争议的调解提供参考借鉴。同时，要注重调解技巧的创新和调解人才的培养，不断提升调解工作的专业化水平。

健全运行机制：提高调解工作的整体效能

健全运行机制的核心目标是构建高效协同的调解工作体系，实现调解资源的整合共享和调解效能的整体提升。

构建协同高效的组织协调机制是基础保障。要以提升工作效率为导向，建立健全跨部门协作机制。设立调解工作联席会议制度，定期研究解决调解工作中的重点难点问题。成立专门的协调机构，统筹协调各方资源，推动调解工作有序开展。建设信息共享平台，实现调解信息的互联互通和资源共享。

优化规范的调解程序机制是重要支撑。细化调解工作的操作规范，实现调解工作的标准化。建立健全调解档案管理制度，做好调解文书的制作、归档和管理等工作。设置科学合理的调解期限，原则上不超过30日，特殊情况可适当延长。建立调解终结制度，明确终结情形和程序要求，确保调解工作依法有序开展。

创新多元化的工作运行机制是发展方向。整合各类调解资源，构建多层次、多渠道的调解网络。建立专业调解人才库，吸纳法律专家、行业专家参与调解工作。推行“调解+法律援助”模式，为经济困难的当事人提供法律援助。建立调解与信访工作衔接机制，实现矛盾纠纷多渠道化解。充分发挥各类调解组织的作用，形成调解资源整合共享、优势互补的工作格局。

强化保障体系：提供有力的制度支撑

强化保障体系的核心目标是构建完善的制度支

撑体系，为调解工作提供坚实的法律保障、监督保障和资源保障。

健全法律制度保障体系是根本要求。加快制定专门的法规或规章，为调解工作提供明确的法律依据。明确规定调解协议的法律效力，赋予其与行政协议同等的法律地位。建立调解协议司法确认制度，增强调解结果的强制执行力。完善调解协议履行监督机制，确保调解协议得到有效执行，切实维护当事人的合法权益。

完善监督管理机制是重要保证。要以提升监督实效为导向，构建全方位的监督体系。建立调解工作考核评价机制，将调解工作纳入行政机关绩效考核范围。建立调解质量第三方评估机制，定期开展调解工作效果评估。实行调解工作信息公开制度，主动接受社会监督，提高调解工作的透明度和公信力。

加强资源要素保障是基础支撑。设立专项经费，为调解工作提供充足的经费保障。建立专业调解人员培养机制，加强业务培训和交流研讨，提升调解队伍的专业化水平。建立调解工作激励机制，调动调解人员的工作积极性。

优化服务机制：提升调解工作的服务质量

优化服务机制的核心目标是构建优质高效的服务体系，提升调解工作的服务水平和群众满意度。

创新服务理念和方法是思想引领。要以提升服务质量为导向，将群众满意作为检验调解工作的根本标准，不断强化服务意识，提升服务能力。建立便民利民的工作机制，最大限度方便当事人参与调解。创新服务方式，运用现代技术手段，提供多样化、个性化的调解服务。

完善服务体系是重要内容。要以优化服务保



给为导向，构建全方位的服务体系。建立“一站式”调解服务中心，提供咨询、受理、调解等全流程服务。设立调解工作便民窗口，开展预约调解、上门调解等个性化服务。建立特殊群体服务保障制度，确保服务覆盖全面、便捷高效。

建立服务评价机制是有效抓手。要以提高服务效能为导向，完善服务评价体系。制定调解服务规范，明确服务标准和要求。建立服务质量评价机制，定期开展满意度调查。完善服务反馈机制，及时发现和解决服务中存在的问题。建立服务改进机制，持续提升服务水平，不断增强群众对调解工作的信任度和满意度。

行政争议调解制度创新是一项系统工程，需要从多个维度协同推进。要坚持目标导向，着眼于提升调解工作的科学性、规范性和效能性，通过制度创新推动行政争议调解工作质量和效率的整体提升。