

算法应用风险与治理研究

李婧文¹, 李雅文²

1. 北京交通大学电子信息工程学院, 北京 100044;
2. 中国信息通信研究院政策与经济研究所, 北京 100191

摘要

随着信息技术的迅猛发展和深入应用, 算法被广泛地运用到交通、医疗、电子商务、信息服务和公共管理等领域。但其自身呈现的数据依赖性、逻辑黑箱性和通用目的性等特征, 也引发了算法训练中的数据滥用、公共管理中的算法歧视、信息服务中的算法操纵、市场竞争中的算法合谋等问题。为了应对算法应用带来的问题, 各国积极采取措施, 加强对算法应用的治理。推进算法的合规利用, 有必要进一步强化顶层设计、完善治理格局、丰富治理手段、加强国际协作, 不断健全算法治理体系。

关键词

算法; 人工智能; 风险; 治理

中图分类号: TP391

文献标志码: A

doi: 10.11959/j.issn.2096-0271.2023028

Research on algorithm and application risk and its governance

LI Jingwen¹, LI Yawen²

1. School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China
2. Institute of Policy and Economics, China Academy of Information and Communications Technology, Beijing 100191, China

Abstract

With the rapid development and in-depth application of information technology, algorithms are widely used in transportation, medical care, e-commerce, information services, public management and other fields, but their own data dependence, logical black box and general purpose characteristics also bring about data abuse in algorithm applications, algorithm discrimination in public management, algorithm manipulation in information services, algorithm collusion in market competition. In order to deal with the problems caused by the application of algorithms, countries have actively taken measures to strengthen the governance of algorithm applications. To promote the effective use of algorithms and strengthen algorithm governance, we need to further strengthen top-level design, improve governance pattern, enrich governance means, strengthen international coordination, and constantly improve the algorithm governance system.

Key words

algorithm, artificial intelligence, risk, governance

0 引言

算法的概念在中西方均由来已久，主要是指某种运算规则的集合^[1]。1970年前后，随着计算机的发展与普及，计算机算法得到广泛应用，先后经历技术化、产品化、工具化的演进过程。20世纪80年代后，生物学的神经网络和深度学习技术被创新性地应用到人工智能领域，由于算法能够不断从数据和经验中重构已有知识结构，大幅提升解决问题的能力，因而被更加广泛地应用于网络检索、信息推荐、对象识别、内容生成等自动化的决策和执行等方面。近10年，随着大数据、云计算、物联网、人工智能等信息技术的快速发展，算法技术作为智能时代的“控制中枢”得到深入应用。

从实践来看，当前计算机算法的大量应用是在推荐技术方面，具体可以分为生成合成类、个性化推送类、排序精选类、检索过滤类、调度决策类等。综合算法的应用可以发现，算法具有数据依赖性、逻辑黑箱性、通用目的性3个显著特征。一是数据依赖性。算法需要监测收集和分析处理海量的用户信息、产品信息和动态数据，数据是算法运行的基石。算法应用者运用海量数据对机器学习算法进行训练，让机器学习算法进行自我调整，以获得更好的性能，最终获取更优解决方案^[2]。二是逻辑黑箱性。复杂的算法决策模型涉及数百万个数据点和数千行代码，用户被置于算法决策的“黑箱”之内，既可能使人无法完全了解决策过程，也无法预测输出结果^[3]。在一些情形下，算法的运行者甚至开发者都很难解释算法的决策过程，更不用说监管者或其他人，算法也因此成为难以监察的“黑

箱”。三是通用目的性。通过引入数据训练模型，机器学习算法能够不断增强问题解决能力、拓宽适用范围，逐渐覆盖日常生产生活、经济社会运行的各方面，如个性化内容和产品推荐、自动驾驶、数字医疗应用等。算法逐步成为跨行业、跨领域的底层应用，为各类通用目的的实现提供了解决方案。

1 算法应用中的风险问题

算法应用与算法涉及的原始数据、技术特征、逻辑架构、平台市场等关键要素密切相关。在算法架构中，数据是训练算法的基础，自动化运行是算法应用的具体方式，目标价值是算法运行的核心，网络平台是提供计算、存储、传输功能的载体。算法正确、合理地运行能够大幅度提高决策的精准性和高效性，而其运用不当则会带来一系列的问题，可能会侵害使用对象的合法权利，甚至会给整个社会造成严重的影响。美国学者巴尔金在对算法应用进行全面梳理和研究后指出，不当的算法应用将会产生数据滥用、算法歧视、算法黑箱、算法操纵等严重问题^[4]。

1.1 算法训练中的数据滥用

算法高度依赖数据，海量数据是训练算法的重要基础，也是算法权力形成的根源。算法既成为收集处理数据和挖掘数据价值的工具，又演变成为影响甚至支配数据流动的权力^[5]。随着新一代信息技术的高速迭代，数据的要素资源禀赋进一步凸显，汇聚、挖掘和变现愈加频繁，同时也会带来隐私保护、数据集中、数据挖掘等一系列风险。其中，大型互联网企业基于其服务模式、业务范畴和数据体量，相较于用户和中

小型企业,掌握更多的数据权力。

一是数据的不当收集。在算法向生产、生活全方位渗透的过程中,个人数据甚至隐私数据的收集显著增多。出于获取更多数据以创造更大收益的目的,一些平台从“告知收集”到“强行采集”再到“隐秘采集”,肆意收集用户数据^[6]。2021年3月,央视3·15晚会曝光科勒卫浴、宝马、Max Mara等商家在消费者不知情的情况下,通过人脸识别摄像机秘密收集用户人脸信息,这些收集活动以自动化的程序进行,经算法处理后应用于商家的营销活动。二是数据的违规共享。平台内部、平台之间违规共享数据的现象普遍存在,如工业和信息化部先后多次通报和批评一些互联网应用程序通过第三方软件开发工具包,违规共享用户个人信息和行为数据。三是数据的非法使用。数据已成为平台开展业务的重要资源。一些平台为了拓展业务,获取流量,大量收集用户数据,在网络用户和数据主体不知情、未许可甚至已明确表示拒绝的情况下擅自和非法使用这些数据,损害了其他平台和用户的合法权益,破坏了公平公正的市场秩序,甚至对国家安全和社会稳定造成危害。

1.2 公共管理中的算法歧视

算法自身的复杂隐性结构、逻辑黑箱性导致其演算和结果无法预测和控制,形成“黑箱”,进而在决策和执行中出现不公平倾向。随着算法决策逐渐渗透至公共管理和社会经济运行领域,其“随机”分类和执行导致的歧视性风险递增。

一是事前难以保障各方公平参与、知情。算法的数据前置性与技术依赖性,使得部分决策的权力中心变为数据和技术的掌握者,并导致算法应用目的和标准不透明、决策结果缺乏救济途径等问题。

如2022年9月,Facebook母公司Meta宣布解雇60名劳务派遣人员,而决策结果完全由算法决定,被裁员工无法公平参与,也无法获知涉及其权益的决策过程。二是歧视性决策结果损害公平正义。部分不当的程序设计和数据集标注,导致错误的价值判断结果,尤其在司法、医疗、就业、劳动力等部分公共管理领域,极大损害被决策者的基本权益,破坏社会公共秩序。如2018年亚马逊的招聘软件被发现读取到带有女性的相关词汇时,算法会自动降低这些简历的权重,歧视女性就业者^[7]。三是难以分配和有效追究不当决策的责任。在司法、医疗、自动驾驶等领域,算法的介入改变了传统的公共管理和服务关系,导致了交通事故、医疗事故中的责任分配难题。2000-2013年,美国共发生10 624起机器人医疗事故,其中出现8 061次机器人失灵的情况,虽然可以通过签署事前协议的方式解决一部分此类事故的责任问题,但仍无法从根本上解决涉及实质性的算法责任归属问题。

1.3 信息服务中的算法操纵

近年来,算法逐步成为互联网信息分发的重要工具,一套全新的内容生成和传播规则随之被构建。算法的广泛性使用使算法成为信息系统运行的“神经中枢”,利用算法实施网络监控,形成信息茧房甚至思想操纵的风险大幅提升。

一是大量传播违法和不良信息。在网络时代,部分互联网信息服务提供者将获取流量作为核心目标,缺乏正确的价值引领。通过算法优先推送博人眼球的虚假信息、极端内容、低俗泛娱乐等信息,甚至生成侵害他人肖像权的虚假合成信息,成为此类互联网信息服务提供者拓展新用户、增加用户黏性的重要途径,极大损

害了网络空间秩序的有序构建。二是导致内容单一,形成信息茧房。算法的自我学习使其对目标群体关注内容的标签化、过滤化进一步加剧,信息茧房将受众屏蔽于不同意见之外,不断固化现有的自我认知,形成具有共同认知的不同群体,甚至导致群体极化、对立化,带来社会焦虑情绪加剧、社会矛盾激化、社会稳定受影响等风险。三是存在舆论和思想操纵等风险。通过数据分析和信息推送,算法能够对特定个人或群体实施精准思想操纵。如在“剑桥分析”事件中,算法通过识别摇摆群体,使用定向信息推送诱导网民支持特定候选人,实现了政治营销目的。

1.4 市场竞争中的算法合谋

在市场竞争领域,算法能够有效提升服务个性化和精准度,通过预测用户行为和 market 发展趋势,辅助降低决策和运营成本。与此同时,平台经济领域自我优待、算法共谋、限制交易等问题对市场竞争秩序带来严重负面影响。

一是大数据杀熟。平台经营者通过收集用户画像、支付能力、支付意愿等“标签”信息,利用算法对消费者进行歧视性定价,对高频用户采取高于低频用户、新用户的定价,导致同一商家、同一产品,不同的消费者可能面对不同的价格,形成“一人一价”“千人千价”的不正常现象,侵犯消费者知情权和公平交易权。近年来,“大数据杀熟”成为社会热点问题,不当算法应用在其中起到了推波助澜的作用。二是自我优待。平台经营者利用其掌握的信息筛选、展示的权力,通过搜索算法、排序算法等,对自营业务予以“优待”,以获得竞争优势,平台同时扮演了“参赛选手”和“裁判员”的双重角色,一定程度上破坏了公平竞争的环境。如2017年欧

盟委员会指出,谷歌滥用其在欧洲经济区(European Economic Area, EEA)的搜索引擎市场支配地位,在搜索结果中优先排序和显著展示与本企业的比较购物服务,降级其他同业者同类服务的排序与展示,损害其他同业者合法权益,并对谷歌处以24.2亿欧元的罚款处罚^[8]。谷歌的行为无疑滥用了市场支配地位,扼杀了市场竞争,剥夺了消费者的选择权和获益于创新的权利。三是算法共谋。共谋是竞争对手之间实施的损害消费者利益的共同利润最大化策略行为,算法的应用扩大了共谋的应用场景和社会危害。平台经营者和其他利益相关者达成联合价格共谋协议,利用算法定价软件收集竞争者定价信息,以提高、维持商品服务价格,实施横向固定价格共谋和纵向价格共谋。相比传统的共谋行为,算法共谋甚至无须明确沟通或互动,这加大了监管部门发现线索、收集证据和确认违法行为的难度。借助算法精度,算法共谋可以比传统的市场操纵方案更加有效。正是基于这种精确性,再加上算法交易的自动化和日益增强的人工智能性质,使算法操纵成为一个特别有害的问题^[9]。

2 各国的算法治理实践

2.1 国外的算法治理实践

算法治理已经成为世界各国共同关注的重要问题。当前全球已有超过50个国家、政府间国际组织、非营利性组织、行业协会发布了与人工智能和算法治理相关的原则和公约。从治理实践看,各国普遍遵循“软法”和“硬法”相结合的治理思路,在强化伦理道德约束的基础上不断完善法律法规和制度规范,推进依法治理。

(1) 美国：注重平衡产业创新发展与安全

美国针对人工智能等高速发展的具体应用场景，通过制定相关立法及解释、发布总统行政命令，推动司法判决，对算法进行创新治理。一是审慎的人工智能应用监管。通过科技政策办公室强化人工智能开发和治理，同时鼓励联邦行政机关建立健全人工智能治理机制。2020年，白宫科技政策办公室发布《人工智能应用监管指南》，明确提出建立和维护公众对人工智能信任的目标，确立审慎监管原则^[10]。二是创新的自动化决策监管。自动化决策系统正越来越多地对公众的健康、财务、住房、教育机会等做出关键决策，带有缺陷或偏见的算法风险可能进一步被放大。2022年，美国参议院引入《2022年算法问责法案》（以下简称《法案》），对算法在社会生活中的深度应用和广泛控制进行再评估，对适用主体做出更加精细化的规定，并设立新的监管机制。针对算法带来的数据滥用、算法歧视、操纵和合谋等问题，该法案对监管机构进行明确授权，规定监管机构可以干预州检察长根据本法案提起的任何民事诉讼；在介入后就民事诉讼中出现的所有事项参加听证，以及就民事判决提出上诉；同时，探索构建自动化决策系统影响评估机制，对自动化决策系统及其开发过程进行评估，以衡量其在设计和数据训练中，对决策结果准确性、公平性、偏见、歧视、隐私及安全的影响。此外，《法案》还提出双线监管问责机制，建立民法、行政法联合的问责规则。

(2) 欧盟：构建良好生态以保障基本权利

欧盟对算法的治理注重公平的技术主权体系构建，关注算法给人类自主性、基本人权和尊严带来的威胁，以及日益突出的极端言论、政治生态等问题^[11]。一是可信、

风险可控的人工智能生态。2020年2月，欧盟委员会发布的《人工智能白皮书》尝试建立一个统一性的欧洲监管框架，旨在提高公民技术信心的“信任生态系统”，并将“高风险”作为监管范围限定标准。二是透明的排序算法治理。欧盟将算法黑箱作为治理重点，力图通过提高算法的透明度治理人工智能应用中的算法歧视、算法操纵和算法合谋等问题。2020年12月，欧盟发布《排序算法透明度指南》（以下简称《指南》），要求算法服务提供者向用户描述排序机制的原理、特征、参数等信息，从而提高用户对排序算法的合理预期。《指南》规定，每个算法服务提供者要提供一个清晰的描述，使用户充分了解排序机制，其中应当重点说明算法服务提供者提供的商品或服务的特征；这些特征与使用该服务的用户的相关性。《指南》对算法服务提供者的算法解释义务进行了详细的规定，包括考虑特定服务的用户类型、技术能力和需求、用通俗易懂的语言起草、对起到决定性作用的参数进行解释，遵循相对突出性和最重要相关性的原则等。欧盟的《数字服务法》也明确规定，在线平台尤其是超大平台应当提高算法透明度，对所使用的算法进行审核并接受主管机关的监管。欧盟的上述措施，有效规范了算法应用，提升了算法应用的透明性和合规性。

(3) 日韩：强调在线交易公平与消费者保护

2021年2月，日本《改善特定数字平台上的交易透明度和公平性法》实施，明确数据平台供应商的信息披露义务，要求平台公平使用排序和分类方法，避免出现自我优惠或歧视性待遇的情况，如平台不得使用优先自身的排序系统等。2020年，韩国颁布《在线平台公平交易法》（以下简称《交易法》），旨在提高平台与在线商店之间交易的透明度和公平性，要求通过透明

公开的合同条款和条件来防止纠纷。《交易法》要求,平台运营商需草拟书面合同并提交给在线商店,合同中应当明确相关算法标准,包括用于确定诸如商品等信息的展示方式和在线平台上顺序的标准、费用对信息呈现方法和顺序产生的影响的标准等。

(4) 国际组织:推动形成统一的可信伦理标准

针对人工智能和算法治理,相关国际组织相继发布原则、指南,推动实现算法透明可释、公平公正和安全可控的目标。2019年5月,经济合作与发展组织(Organization for Economic Co-operation and Development, OECD)发布首个政府间人工智能政策指南《人工智能原则》(以下简称《原则》),《原则》发布之后即有42个国家签署,随后被二十国集团(Group of 20, G20)采纳。OECD的人工智能原则包括5项基本原则,以及5项公共政策和国际合作建议,涉及决策的透明性、合理化、公平性等涉及算法技术应用的问题。电气与电子工程师协会(Institute of Electrical and Electronics Engineers, IEEE)发布《人工智能设计的伦理准则》并组建IEEE P7000工作组,迄今为止已发布多项旨在解决系统设计过程伦理问题的标准,如IEEE P7001-自主系统透明性标准、IEEE P7003-算法偏见的处理标准等^[12]。

总体来看,随着算法逐步成为经济社会的底层架构,算法治理已经成为各国政府共同关注的话题。近年来,各国关于算法的政策指引和立法、执法数量激增,针对广告营销、搜索引擎、在线社交等应用发布一系列立法及行业操作指南,指导行业积极应对算法给隐私、版权保护带来的威胁和不当的算法操纵、交易等问题^[13]。各国政府对算法的重视和规制,有效提高了算法应用的透明度和规范化,防范了算法应用带来的风险^[14]。

2.2 我国的算法治理实践

近年来,我国高度重视算法治理工作,在完善算法治理规则、加强算法应用管理、部署算法违法行为整治等方面推进算法治理和监督,形成综合治理格局,特别是从规则完善和平台监督两方面深入推进算法治理。

(1) 出台算法相关伦理规范和专门规定

2021年9月,国家新一代人工智能治理专业委员会发布《新一代人工智能伦理规范》,对人工智能开发、应用中“避免偏见歧视”进行具体规定。该规范提出在数据采集和算法开发中,应当避免可能存在的数据与算法偏见,努力实现人工智能系统的普惠性、公平性和非歧视性。同月,国家互联网信息办公室等九部委联合印发《关于加强互联网信息服务算法综合治理的指导意见》,提出健全治理机制、构建监管体系、促进算法生态规范发展的算法安全综合治理格局,加强算法综合治理。2021年12月,国家互联网信息办公室、工业和信息化部、公安部和国家市场监督管理总局联合印发《互联网信息服务算法推荐管理规定》,这是我国首部专门的算法治理法规,也是全球第一个综合性、全面性、系统性规制算法的法律规定,标志我国的算法治理迈入法律治理时代。2022年6月,修订后的《中华人民共和国反垄断法》明确规定,经营者不得利用算法实施垄断行为,进一步强化对算法应用的法律规制。

(2) 丰富数据与算法司法治理实践

近年来,我国加快推进数据与算法的司法治理,审理解决了一批关于数据权益、数据使用行为的纠纷,以厘清不同主体的数据权益边界、防止数据垄断和不正当竞争。同时,规范数据抓取行为,通过司法手段推动建立公平、诚信的数据要素

市场秩序,促进数据开放、共享和有效利用,推进数字经济健康发展。面对互联网平台“大数据杀熟”的问题,司法机关通过具体案件审理和典型案件示范,进一步厘清电商平台经营者在检索服务领域的责任边界,平衡平台经营行为与消费者保护之间的关系。2022年8月,杭州互联网法院公布的数据和算法典型案例中,就有多起关于平台算法自动化决策、电子商务平台算法歧视、游戏算法生成权利纠纷的涉算法典型案例。

(3) 开展算法治理专项行动

为了应对算法带来的问题,维护广大网民的合法权益,国家相关部门加快推进算法专项治理。2022年4月,中央网信办下发《关于开展“清朗·2022年算法综合治理”专项行动的通知》,部署开展算法综合治理专项行动。专项行动明确具体、针对性强,推动了互联网企业梳理算法应用情况,开展算法安全能力评估,整改存在的问题,消除安全隐患。

从治理成效看,目前我国的算法治理体系加速构建,覆盖算法分级分类管理、算法备案、风险防范、违法行为惩治等环节的综合治理格局初步形成。现有体系将规范互联网信息服务和加强用户权益保护相结合,将常态化监管与专项行动相结合,切实压实平台主体责任。司法治理以具体案件为切入,对自动化决策、算法歧视等问题进行司法审查,实现对算法应用的有效监督。与此同时,算法治理工作在提升算法透明度和可解释性、结合算法多元应用场景进行针对性治理、强调社会组织等第三方共同参与等方面稍显薄弱,仍需持续发力。

3 加快推进算法治理的思考

随着算法在经济社会和生产生活中

日益广泛的应用,为应对算法应用带来的风险问题,各国需要进一步加强对算法应用的治理,以实现新技术新应用的趋利避害、为我所用。加强算法治理,推进算法的健康发展和有效利用,需要不断强化风险防范意识,将算法治理纳入国家战略,按照系统治理的思路^[15],探索开展算法规制立法、监管体系构建、算法产业发展引导等工作,多措并举实现对算法的有效治理。

3.1 强化顶层设计,完善政策法规体系

作为大数据时代的底层技术,算法对推动数字经济发展和国家治理能力现代化具有重大意义。因此,加强算法治理,要树立发展和安全并重的原则。一方面,要深入分析算法风险的根源和特征,防范算法应用带来的个体性和社会性问题;另一方面,要给予算法技术足够的创新空间,针对性地总结算法治理经验,推进算法在经济社会和生产生活中的有效利用。在发展环境方面,着力开展针对人工智能、算法等新技术、新应用的政策法规建设,强化顶层设计,为算法技术的创新发展和深入应用提供政策指引和法治保障。及时出台支持算法技术发展和引导算法技术在不同行业、不同领域、不同场景应用的指导文件,鼓励算法技术发展应用。同时,制定规范算法技术具体应用关键环节的专门规定,特别是针对数据安全和个人权利保护、市场竞争秩序等的痛点难点问题,加快推进相关法律法规修订解释和出台,实现法律与政策、立法与伦理、国内立法与国际规则相衔接的规则体系。

3.2 完善治理格局,推进多元共治

算法治理是一个系统的工程,既应当加

强外部监管,又应当完善算法应用自身^[16]。算法治理需要从互联网服务的特点出发,构建主管部门、行业协会、网络平台、网络用户协同互动、多元共治的治理格局。按照“谁主管谁负责、谁服务谁负责、谁运营谁负责”的原则,进一步明确和落实主管部门的管理责任,压实互联网企业的主体责任。借鉴欧盟经验,提高算法的透明性,推进算法公开,促进应用公平。建立平台算法责任体系,完善算法标准,推进技术管网^[17],不断建立和完善技术治理、动态监测等治理手段,基于算法涉及的场景丰富、主体复杂等特征,探索构建多手段、多维度的综合算法治理框架。同时,推进行业自律,充分发挥行业协会在算法治理中的作用,建立算法治理行业自律机制,制定行业规范和自律公约,通过建立伦理道德委员会、调节委员会等工作机制,推动行业规范、自律公约、伦理指南等规范规则落地。引入第三方监督主体,充分发挥具备专业技能的第三方机构的外部监督作用,对数据访问、算法模型等进行监督审查,有效弥合监管部门和公众的专业能力鸿沟。培育广大网民特别是算法用户的“算法素养”,强化公众在算法应用中所需具备的思维培养和风险教育,帮助公众认识算法可能带来的风险,建立算法治理的公民参与机制,有效平衡平台与公众间的信息不对称、权利义务不对等问题,推进算法技术和应用的社会共治。

3.3 立足算法特性,丰富治理手段

从算法的数据依赖性、逻辑黑箱性和通用目的性特征出发,借鉴国外实践经验,探索创新治理手段。构建基于算法分类分级的备案体系,研究算法分类分级标准,探索构建针对涉及国家主权、国家

安全等方面的重点算法备案审批体系,实现对算法风险预警和风险回溯。借鉴美国《2022年算法问责法案》相关立法经验,探索建立算法安全、影响评估机制,评估企业在算法应用方面的监测预警、应急处置等安全技术能力,主动介入和防范算法在信息传播、舆论引导方面存在的安全风险隐患。加强数据安全治理,强化对个人信息和重要数据的收集、处理、共享等关键环节的合法保护,完善数据分类管理、分级管理、授权管理,防止因数据泄露、滥用而危害个人、国家安全。开展算法监管技术手段建设,研究部署国家级算法技术监测平台,构建监测发现、风险管控、追踪溯源相结合的协同监管能力,有效防范算法滥用带来的风险。

3.4 强化国际协调,实现共同治理

信息技术的共同性决定了治理规则的共同性。在国际层面,公开透明、公平公正已成为各国算法治理的共同目标。为实现这一目标,政府间国际组织可以通过开展国际对话、协调与合作的方式,加快推进算法研发和应用。在尊重各国的数据、算法、人工智能治理原则的前提下,加强各国在算法透明度、可解释、可问责等方面的交流合作,推动在全球层面围绕算法形成治理共识,促进人工智能行业在全球范围的健康有序发展。同时,积极倡导世界各国各方摒弃对算法技术、应用、产品的“歧视性做法”,反对和破除算法技术贸易壁垒,推进算法技术及应用共享共用,推动建立多边、平等的沟通渠道,交流治理经验。按照国内规则和国际规则一体推进的原则,在推动各国形成有效治理规则的同时,形成广泛适用的算法技术和应用的国际治理规则,最终实现对算法的国际性治理。

参考文献:

- [1] 布拉萨德, 布拉特里. 算法基础[M]. 邱仲潘, 柯渝, 徐锋, 译. 北京: 清华大学出版社, 2005: 1-4.
GILLES B, PAUL B. Fundamentals of algorithms[M]. Translated by QIU Z P, KE Y, XU F. Beijing: Tsinghua University Press, 2005: 1-4.
- [2] BURRELL J. How the machine “thinks”: understanding opacity in machine learning algorithms[J]. *Big Data & Society*, 2016, 3(1).
- [3] BATHAEE Y. The artificial intelligence black box and the failure of intent and causation[J]. *Harvard Journal of Law and Technology*, 2018, 31: 889.
- [4] BALKIN J. 2016 Sidley Austin distinguished lecture on big data law and policy: the three laws of robotics in the age of big data[J]. *Ohio State Law Journal*, 2017, 1217(78): 1237-1239.
- [5] 张凌寒. 权力之治: 人工智能时代的算法规制[M]. 上海: 上海人民出版社, 2021.
ZHANG L H. Governance of power: algorithm regulation in the age of artificial intelligence[M]. Shanghai: Shanghai People's Publishing House, 2021.
- [6] 王磊. 个人数据商业化利用法律问题研究[M]. 北京: 中国社会科学出版社, 2020.
WANG L. Research on legal issues concerning the commercialization and utilization of personal data[M]. Beijing: China Social Sciences Press, 2020.
- [7] DASTIN J. Amazon scraps secret AI recruiting tool that showed bias against women, REUTERS[Z]. 2022.
- [8] JAMES V. Google fined a record €2.4 billion by the EU for manipulating search results, the verge[Z]. 2022.
- [9] YADAV Y. The failure of liability in modern markets[Z]. 2015.
- [10] 朱巧玲, 杨剑刚. 算法陷阱与规制跨越[J]. *经济学家*, 2022(1): 104-114.
ZHU Q L, YANG J G. Algorithm trap and regulatory breakthrough[J]. *Economist*, 2022(1): 104-114.
- [11] 胡坚波. 多措并举推进我国算法治理[J]. *人民论坛·学术前沿*, 2022(10): 18-28.
HU J B. Take multiple measures to promote algorithmic governance in China[J]. *Frontiers*, 2022(10): 18-28.
- [12] 徐凯, 蒋洁. 人工智能应用的社会性风险与监管对策[J]. *信息通信技术与政策*, 2021, 47(5): 15-21.
XU K, JIANG J. Research on social risks and supervising countermeasures of AI applications[J]. *Telecommunications Network Technology*, 2021, 47(5): 15-21.
- [13] SAURWEIN F, JUST N, LATZER M. Governance of algorithms: options and Limitation[M]. [S.l.:s.n.], 2015.
- [14] DOUGLAS H. Deepfakes: false pornography is here and the law cannot protect you[J]. *Duck Law and Technology Review*, 2019, 17(1): 99-127.
- [15] BEVIR M. A theory of governance[M]. Berkeley: Global, Area, and International Archive, 2013.
- [16] 丁晓东. 基于信任的自动化决策: 算法解释权的原理反思与制度重构[J]. *中国法学*, 2022(1): 99-118.
DING X D. Automatic decision-making based on trust: principle reflection and system reconstruction of the right to algorithm interpretation[J]. *China Legal Science*, 2022(1): 99-118.
- [17] 衣俊霖. 数字孪生时代的法律与问责: 通过技术标准透视算法黑箱[J]. *东方法学*, 2021(4): 77-92.
YI J L. Law and accountability in the digital twin era—perspective of algorithm black box through technical standards[J]. *Oriental Law*, 2021(4): 77-92.

作者简介



李婧文(2002-),女,北京交通大学电子信息工程学院本科生,主要研究方向为计算机技术应用、电信技术与应用。



李雅文(1990-),女,中国信息通信研究院政策与经济研究所高级工程师,主要研究方向为数字治理、网络法治、数字经贸规则。

收稿日期: 2022-11-28

通信作者: 李婧文, bjtuljw@sina.com