Open Journal of Legal Science 法学, 2023, 11(6), 5185-5191
Published Online November 2023 in Hans. https://www.hanspub.org/journal/ojls
https://doi.org/10.12677/ojls.2023.116741

Hans汉斯

人工智能算法侵害的立法规制机制研究

陈思羽

南京信息工程大学法政学院，江苏 南京

收稿日期：2023年8月30日；录用日期：2023年9月13日；发布日期：2023年11月10日

摘　要

人工智能算法是社会生产力发展和科技进步的产物，其广泛应用极大地促进了社会生产力水平的跃升与生产关系的变革，但也催生了一系列新型风险，包括个体层面的算法歧视和算法偏见，市场层面的算法共谋与算法垄断，国家层面的算法黑箱、算法霸权等，亟待采取立法规制。为破解当前监管机制不完善，算法责任界定不明等诸多局限，需要实现人工智能算法可解释、可评估、可监管等可信特征，有效防范人工智能算法风险。从人工智能算法侵害风险类型化的角度出发，提出对与人工智能算法发展有关的传统的法律制度进行适用性改进，并从立法方法、立法模式、立法路线三个方面提出规制人工智能算法的法律路径。

关键词

人工智能算法，侵害风险类型，立法规制

Research on the Legislative Mechanism of Artificial Intelligence Algorithm Infringement

Siyu Chen

School of Law and Public Affairs, Nanjing University of Information Science and Technology, Nanjing Jiangsu

Received: Aug. 30th, 2023; accepted: Sep. 13th, 2023; published: Nov. 10th, 2023

Abstract

Artificial intelligence algorithm is the product of the development of social productivity and scientific and technological progress. Its wide application has greatly promoted the improvement of social productivity and the transformation of production relations, but it has also spawned a series of new risks, including algorithm discrimination and algorithm bias at the individual level, algorithm collusion and algorithm monopoly at the market level, algorithm black box and algorithm hegemony at the national level, etc. Legislative regulation is urgently needed. In order to solve the

文章引用: 陈思羽. 人工智能算法侵害的立法规制机制研究[J]. 法学, 2023, 11(6): 5185-5191.
DOI: 10.12677/ojls.2023.116741

current many other limitations such as the imperfect regulatory mechanism and the unclear definition of the algorithm responsibility, it is necessary to realize credible characteristics of artificial intelligence algorithm that can be explained, evaluated and regulated, effectively prevent the risks of artificial intelligence algorithm. From the perspective of AI algorithm's infringement risk type, this paper proposes to improve the applicability of the traditional legal system related to the development of AI algorithm, and proposes the legal path to regulate AI algorithm from three aspects: legislative method, legislative model and legislative route.

Keywords

Artificial Intelligence Algorithm, Infringement Risk Type, Legislative System

1. 问题的提出

在信息技术背景下，伴随着人工智能和大数据的蓬勃发展，智能算法在社会各个领域得到广泛应用。算法技术不仅为政府、企业和公众带来了良好的生活体验，而且大大提升了经济生产效率和效益。然而，智能算法的广泛应用也带来了一系列社会治理问题。算法正日益形成一种介于公私权利之间的“准公权力”，即“算法权力”。由于其强大的算力优势和深度自主学习的能力，再加上算法“黑箱”提供的可操作化空间，算法正日益摆脱其“工具性”的角色，扩大其部署社会资源的能力。

在数字社会中，伦理非常重要，我们非常担心隐私泄露、国家安全问题等。伴随着大数据分析技术的升级与智能技术的普及，算法带来的风险不容忽视。从传播学视角看，算法的直接影响是产生信息茧房、信息过滤泡等，人们的视界被熟悉的内容和思维所限制并进一步固化，甚而这种固有之见会被推向极端形成群体极化。从社会学和政治经济学角度看，算法使数字鸿沟更加显著，数字劳动被进一步商品化和资本化、隐私主体丧失控制权。从经济学的角度看，算法应用有助于网络平台通过收集个人或企业的基础信息及行为趋势、产品结构以及各类参与方的行为特征，令共谋行为更易达成，此外，还可以针对个人用户的偏好实施动态的价格歧视或服务歧视，剥夺了更多的消费者剩余。

在此背景下，为了适应技术发展与社会需要，国家政府需要通过立法和标准制定对算法应用的相关领域进行治理与监管。目前，国内外法学学者已对算法权力的形成以及其对社会资源和数据的控制作用进行了分析。同时，他们也对算法权力在商业领域和公权力程序中嵌入所带来的异化风险提出了警示，并从立法的角度考虑治理的架构与措施。但目前国内对人工智能的研究主要集中在相关的纲领性文件中，具有系统性、战略性的特点，其内容相对概括、规制对象较为模糊。

2. 人工智能算法规制的理论基础

从概念上，算法并非一个新奇的事物，它与现代信息技术同步产生，距今已有超过半个世纪的历史。人工智能算法作为广义上计算机软件算法的一种特殊类型，也属于为了解决某个特定问题或者达到某个特定目的所要采取的一系列步骤[1]。在“人工智能”的概念还没有诞生之前，控制论正在流行。“控制论之父”诺伯特·维纳认为不仅在人类和人类社会中，在其他生物群体乃至无生命的机械世界中，都存在着同样的信息、通信、控制和反馈机制[2]。维纳呼吁科学家和技术人员承担起更大的道德和社会责任，

以迎接这个正在蓬勃兴起，建设性和破坏性能量并存的时代。如今，为有效应对算法歧视，算法霸权等风险，需将法律的知识体系、价值体系与算法的技术体系结合，从而达到对算法歧视的法律规范，以及网络信息服务算法的安全管理，让算法技术能够在可信的特征下为人类社会服务。建立起由科技理性(技术良序)和法律理性(法律良序)共同支撑的、人与人之间相互协作配合的，以追求社会成员的共同利益为社会主要目标的算法时代的“良序社会”。

我国的算法治理实践已初具规模，但仍处于碎片化状态。《“十四五”国家信息化规划》提出，要“建设技术规则治理体系”“开展技术算法规制”，推进算法“安全评估审查”“标准制定”“伦理论证”。算法立法研究具有必要性和紧迫性。同时，理论研究主要集中在算法偏误、算法透明、算法解释等方面[3]，或集中于算法应用领域，如电子商务，社交媒体等领域，算法治理研究相对薄弱。目前对于算法规制研究的问题在于，单纯地对策性研究太多，而没有回归到基础原理的角度去考虑问题。本文将人工智能算法侵害的风险类型化，在此基础上提出人工智能算法风险预防的核心思路与规制路径。

3. 人工智能算法侵害风险的类型及防范思路

3.1. 人工智能算法侵害风险的类型

由于我们离不开算法技术，因此，我们对于算法的监管，是对算法的应用场景和服务进行规制，而不是对技术本身进行监管，算法的治理需要与算法的服务结合在一起。我们对于算法治理，需要根据其产生的风险进行分类，再通过跨法域的治理规制那些具有多重风险属性的算法侵害，这样既能回应具体问题，又能形成整体架构。基于人工智能算法侵犯对象的特点，可将人工智能算法侵害风险划分为以下三类。

第一，人工智能算法侵害私人权益。算法强化着信息茧房，带来人的认知窄化风险。所谓知识窄化，是指人们对知识的认知、情感或思维意识向某一方面或某一方向高度集中，知识的领域日益狭窄。所谓“信息茧房”，就是通过算法，准确了解用户的阅读内容和浏览路径，给用户创建标签，继而进行准确地匹配，将相关内容推荐给用户。使用户如同生活在茧房中，知识信息逐渐单一化、自我中心化，侵害用户的信息自我决定权，造成社群沟通鸿沟。信息茧房不仅操纵个人信息接受者，继而影响个人的价值观与政治观点[4]，它还导致了社会差距，甚至是社会分裂，阻碍公共议题的协商一致，将会产生深远的负面影响。

算法技术平台的研发人员和操控者，利用其庞大的技术优势，在不同的领域中，广泛而持久地发挥着自己的行为支配，引导、支配着人类的思维模式和行为，并在宏观层面上支配着人类的行为。这就是控制论在人类行为上的支配力。商业数字化极大地改变了企业在个人层面影响消费者的能力，一套特定的新兴技术和技术将企业能够发现和利用每个消费者追求自身利益能力的极限。企业将越来越有能力引发消费者的非理性或脆弱性，导致挑战消费者保护法限制的实际和感知的伤害。

第二，人工智能算法侵害市场利益。算法遵循一种商业逻辑，构成了监视资本主义的危险，算法“黑箱”的出现，会使个人利益在主观上攫取公众利益，从而使资金在主观上回避了公权力的制约。同时，算法“黑箱”在客观上掩盖了算法本身的不足，并且有可能引发安全隐患，使得监管部门很难对算法信息进行审核。由于管制算法在内容和方法上都有局限性，很难对算法问题进行及时的问责和纠正。算法霸权带来的骚扰电话、电信金融诈骗、广告轰炸、数据泄露、网络黑产等现象，已经严重侵害了社会不特定多数人的利益，影响了社会秩序的稳定。

企业不仅有参与市场操纵的能力，而且还有经济激励：如果一些市场参与者利用偏见，那些没有利用偏见的人可能会被挤出市场。算法监控传播的主要原因是商业利润的最大化。对公司来说，数字化为

创造新的商业世界提供了基本条件。一方面，他们整合现有数据，建立用户档案，确保准确地营销，另一方面，通过收集所有利益相关者的反馈，他们可以及时根据用户需求调整营销策略，实现利润最大化。这就是。当技术与资本融合时，它不可避免地被社会关系和资本所代表的逻辑所驱动，算法监控的应用是公司利润动机的反映，它面向的是无限的利润增长。在这个过程中，数字数据是元素，算法监控是手段，资本积累是最终目标。

第三，人工智能算法侵害国家安全。算法具有被某些利益集团用来进行社会控制和政治权力再生产的政治风险[5]。传统意义上的国家安全主要包含经济安全、环境安全、文化安全、科技安全、信息安全、人类安全等在内的综合安全。由于国家安全领域的安全问题涉及范围广、内容复杂，安全要素相互累加形成的系统效应(System Effects)已远远超出人们的理解范围，这就为算法的介入与辅助决策提供了条件，而这种依赖性同时也为算法安全带来了隐患[6]。

由于人工智能技术的高效率和狭窄的门槛，使得技术巨头们对大数据的使用有着绝对的优势，特别是在数据收集和利用、算法更新和迭代、深度学习等方面，政治精英和智力精英的结合越来越有必要，也越来越有可能，他们所组成的政治同盟，使得他们与一般民众之间存在着“数字鸿沟”，由于技术、政治资源的匮乏，一般民众获得政治权利的可能性越来越小，其对政治力量的影响也越来越小，政治力量的寡头倾向或将越来越明显。

3.2. 人工智能算法侵害风险的防范思路

传统法律法规规则在人工智能场景下存在可用性、有效性，但某些与人工智能发展有关的传统法律制度需要适应性改进，个别法律制度需要创新。例如《中华人民共和国个人信息保护法》第 24 条第 2 款规定：“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。”该条将个性化推荐算法纳入法律的规范射程内，强调了个人的自主决定权。不过，该条的规定过于简单，无力全面规范个性化推荐活动。实际上，这一方法具有很强的应急性质，而且缺乏系统性，很难在电子商务以外的人工智能应用中使用。应对算法侵害的风险可以从以下三个方面入手。

首先，在个体层面上，应强化个人被算法侵害的保护力度。社会个体在算法面前逐渐被演化成数据信息的集合体，成为算法的“囚徒”，隐私泄露与数据杀熟的风险加大，个人自由难以实现，强化个人被算法侵害的保护力度需坚持以下两个立法原则：一是在人工智能产品风险评估领域创设“风险预防”立法原则。二是在人工智能产品风险评估领域创设“法律与技术相结合”立法原则。一些带有歧视性的结论不是根据个人的主观偏好，而是通过计算机进行编辑，形成了一种带有歧视性的语言，比如，黑人的照片会因为肤色和一些特点而被误以为是“黑猩猩”。算法歧视与外部因素相比造成的算法歧视，更多地归因于算法中的技术因素。即由于算法技术本身的局限性。相较于基于外部原因的算法歧视，克服这类算法歧视则更加困难，需要在算法技术的升级中寻找突破。在设计人工智能算法时，考虑算法技术在发展过程中可能会出现无法预测、可能会造成巨大损失等情形，应及时出台相应机制保障算法风险防控与治理。对于人工智能算法所引发的法律纠纷问题要根据司法实践特点有针对性地加以解决。

其次，在市场层面上，在人工智能市场监管上进行针对性立法建设。要建立和完善平台经济管理制度，明晰规则、划定底线、强化市场监督、规范市场秩序，促进公平竞争，反对资本垄断，防止资本野蛮生长。平台封禁行为有其正当性，规制平台实施封禁行为也有必要的限度，应加强平台封禁行为的系统治理、确立保护数字人权的思想理念、探索平台数字法治发展道路。在维护数字市场交易秩序的同时促进平台经济的创新发展。在算法共谋的行为中，数据驱动下的算法能够迅速监控竞争对手的价格，并统一地调整价格[7]，因此经营者可以通过算法实时地协调价格、限制产量。从传统的反垄断规制理论体

系当中汲取合理内核，结合实践探索出一套合理的人工智能算法反垄断路径，塑造公平合理的网络市场，充分维护消费者的合法权益。

最后，在国家层面上，国家应重点研发提升防止可信人工智能算法侵害风险立法的立法技术。现有法律体系对此并没有充分有效的应对措施，在制定法律规则时仍然存在不平衡不充分问题。要应对算法风险，应当首先从国家层面完善相关法律法规，在此基础上进一步明确相关原则和要求，使法律在规范算法行为方面具有稳定性、可预测性以及可操作性，避免算法滥用造成的后果。目前《互联网信息算法推荐服务管理规定》等部分法律文件的出台表明了我国正式回应了算法挑战，只是立法偏原则性，具有可操作性的具体条文仍需补充。国家应当重点研发提升防止可信人工智能算法侵害风险立法的意见综合分析技术、立法知识库构建技术、基于语义意图理解的算法歧视预测预警技术、产品风险智能识别应用技术等立法技术。加快智能算法领域的立法进程。

4. 防范人工智能算法风险的规制路径

由于算法、数据、信息立法的目的、对象、内容各不相同，算法规制的单独立法亟待得到推进。算法立法的本质不是对技术本身的规制，而是在传统领域数字化改造过程中，避免算法运行对已有价值或秩序造成损害，保障算法决策下的竞争机制公平、用户权益保护、公共秩序维护。解放和发展生产力，防止算法脱离社会控制而影响算力，是算法规制的根本目的。为了搭建更为全面的算法治理框架，2021年12月31日，国家互联网信息办公室等四部门联合发布《互联网信息服务算法推荐管理规定》(以下简称《算法推荐管理规定》)，该规定已于2022年3月1日生效。《算法推荐管理规定》意图加强对算法风险的全流程监管，创设算法规范的中国方案。然而，该规定创设的算法规范方案以行政机关的直接管制为主，大量规范缺乏执行保障，一些具体规范的设计也有进一步权衡的必要。算法规范的设计离不开对算法风险的理性认识和对规制框架的体系性思考。如果没有一个清晰的规范思路，我们只能回应性地设立规则，不同算法规范之间的冲突将难以避免。

4.1. 防范人工智能算法风险的制度改进

算法、数据、信息、网络密不可分，算法立法必然与已有立法相联系，聚合数据信息立法、电子商务法等多项法律价值。不过，通过已有立法进行算法规制并不具有周延性。在把握算法立法与数据信息等立法的异同基础上，还需处理算法立法与已有法律制度的衔接问题，确保算法制度设计能够在实践中真正落实。人工智能算法要素渗入到传统法律制度后，需要对传统法律程序制度、法律责任制度、法律裁量制度和法律监督制度等进行适应性改进。

首先，在法律程序制度中，需增设过程公平规则。我国算法公平侧重于结果公平，过程公平规则寥寥。可事实上，没有过程公平，就不可能有结果公平，因为“过程”有着塑造结果的能力。算法的过程公平即个人能够平等且有意义地参与到算法活动全过程之中，其体现在：(1) 决策算法选择的公平性：算法使用者应说明为何采取算法，并简要说明优化目标；(2) 算法输入数据的公平性：如数据特征选择的个人自主性、数据特征可信度、数据特征的相关性等；(3) 在自动化行政的场景下，算法结果的有效性以通知与申诉的履行程序为前提，允许民众提出质疑，并有权在专业审计人员的协助下审查算法并及时纠错。

其次，在法律责任制度中，一是算法责任治理应当全面涵盖人、算法、社会三要素，包括对“人”的治理、对算法的治理和对社会的治理。二是算法责任治理需要综合运用算法责任的多种实现机制，包括基础层次的社会责任融入、中间层次的负责任研究与创新、高阶层次的敏捷治理。三是归因矫正的综合性。算法责任治理应当囊括对人为型算法失当、功能型算法失当、触发型算法失当的治理，针对它们所产生的各类成因进行矫正。

再次，在法律裁量制度中，为保证机器计算结果的准确度和合理性，应侧重于完善传统的裁量基准并对其进行精准的代码转译。在决策权限保留方面，自动化处罚裁量的功能在于规范裁量权行使，而非剥夺执法人员的个案考虑义务，执法人员也不应怠于裁量。负责个案处理的执法人员应保有处罚决策权，应有权决定是否采用机器计算得出的量罚结果。为应对机器故障等偶然因素的出现[8]，执法人员不应被机器架空。

最后，在法律监督制度中，我国在市场监管领域对算法问题的研究和监管还处于相对早期的阶段，更需加强调查研究，提升监管能力。建议采取以下四方面措施：首先，在现有法律法规基础上，进一步出台平台算法应用合规指引，细化算法原理公示等相关要求，为行政执法和企业行为树立明确边界。其次，强化我国平台经济市场竞争状况调查，进一步加强对算法歧视、算法合谋等行为的实证研究和竞争损害分析研究，以便在反垄断相关配套规定中进一步明确具体违法情形。最后，加强对算法透明、事前监管等制度的研究和影响评估，可探索运用消费者权益保护、经营者权益保护等其他监管工具，对一些涉嫌垄断的算法应用行为进行综合约束。

4.2. 防范人工智能算法风险的立法路径

算法规制不能停留于伦理约束，而须通过立法，设定“算法合规”准则，确保算法应用的正当性，推动“算法伦理”价值的落实。在保障算法合理应用的同时，防止过度规制对算法技术创新的抑制、对算法核心竞争力的损害。在保障算法合理应用的同时，防止过度规制对算法技术创新的抑制、对算法核心竞争力的损害。

算法法律关系不仅包括行政法律关系，还包括提供算法服务的企业与消费者之间的法律关系、算法服务提供者之间的竞争法律关系。防范人工智能算法风险的立法路径选择面临诸多挑战。算法立法不仅要防止算法应用对消费者权益等传统法益的损害，而且要应对诱导沉迷等新兴领域的“算法‘驯化’”问题，这对算法立法方法、立法模式、立法路线提出了更高的要求。

首先，在立法方法方面，采用适用性改进与专项突破的方法。算法立法需要区分四种算法侵害风险，设置专门条款加以规制。从行为角度看，算法应用主要包括算法匹配、算法推荐、算法决策、算法筛查四种行为。算法立法应当分类对算法应用行为作出细化规定。人工智能的三要素(数据、算法、算力)进入了传统法律领域，面对算法产生的新型风险，应该以传统法律为基础，在传统法律制度上进行适用性改进。同时，算法专项立法亟待得到推进，算法专项立法应当通过申辩权、标签删除权、算法解释权、备案及评估信息公开请求权等程序性权利，以及备案审查、责令停用或整改的权力设置，引导算法规制正当程序的确立，建立算法决策领域公益与私益的个案调和机制。

其次，在立法模式方面，采用实验性立法模式。在人工智能立法模式上采用实验性立法模式，有计划、分步骤地积极稳妥推进，在与人工智能行业相关的领域开展法治智能社会实验并探索总结规律，根据实验反馈的现实需要稳步推进立法。2022 年《互联网信息服务算法推荐管理规定》生效实施，其以部门规章形式，对使用五类算法技术的算法推荐行为作出规制，对防范算法推荐技术滥用起到至关重要的作用。但部门规章在防控算法风险过程中存在一定的局限性。算法风险的治理是一项系统、复杂的工程，不仅涉及个人信息保护，也与数据商业利用、流转和共享等数据生态链的诸多环节密切相关，需要政府部门、监管机构等多元社会主体参与合作治理。

最后，在立法路线方面，采用纵向立法与横向立法结合的立法路线。部门规章的调整对象是行政管理关系，部门规章的调整对象是行政管理关系，但很难统筹算法和法律关系的所有内容。而且容易导致针对一种违法行为制定一种规则的滞后性问题。提出遵循省市/部委+行业领域先行纵向立法试点，后制定中央层面的人工智能领域一般上位法的立法路线图，并在及时总结地方和部门立法经验的基础上，适

时出台法律、行政法规层面的人工智能单行法在纵向立法层面，先开展省市立法试点以及行业领域的纵向立法，再制定中央层面的人工智能领域一般法。简而言之，即在纵向立法层面，在中央层面搭建促进人工智能产业的立法框架。在横向立法层面，根据促进法和地方立法的探索试验，适时出台行政法规、法律层面的人工智能单行法。

5. 结语

算法的到来给人类社会带来许多便利，但这也给人类社会增加了许多风险。由于算法在设计、部署和应用过程中不受控制，因而产生算法歧视、算法垄断、算法霸权等问题，这将极大地影响到人们追求的公平、自由等核心价值观。我们需要可解释、可靠、可控的“可信算法”，仅依赖算法控制者的技术伦理和道德自觉并不能达成算法可信[9]，亟需借助法律对人工智能算法产生的风险进行规制。

法律应以一种全新的调整姿态介入防范算法侵害风险的难题中，用系统性的数字法律制度实现可信人工智能算法治理。算法风险治理需要确立“风险预防”和“法律与技术相结合”的立法原则，加强市场监管，提升立法技术。将算法侵害的风险进行类型化，同时着眼于实现机制的研究，能够深入解构算法治理的内容和分类，系统地理清算法尺度和判据。

参考文献

[1] Diakopoulos, N. (2015) Algorithmic Accountability: Journalistic Investigation of Computational Power Structures. *Digital Journalism*, **3**, 398-415. https://doi.org/10.1080/21670811.2014.976411

[2] 王攀. “人工智能之父” 66 年前的预言[J]. 经营管理者, 2016(11): 110-111.

[3] 许可. 驯服算法: 算法治理的历史展开与当代体系[J]. 华东政法大学学报, 2022, 25(1): 99-113.

[4] 王莹. 算法侵害类型化研究与法律应对——以《个人信息保护法》为基点的算法规制扩展构想[J]. 法制与社会发展, 2021, 27(6): 133-153.

[5] 孟天广, 李珍珍. 治理算法: 算法风险的伦理原则及其治理逻辑[J]. 学术论坛, 2022, 45(1): 9-20.

[6] 贾珍珍, 刘杨钺. 总体国家安全观视域下的算法安全与治理[J]. 理论与改革, 2021(2): 135-148.

[7] 李丹. 算法共谋: 边界的确定及其反垄断法规制[J]. 广东财经大学学报, 2020, 35(2): 103-112.

[8] 王正鑫. 机器何以裁量: 行政处罚裁量自动化及其风险控制[J]. 行政法学研究, 2022(2): 166-176.

[9] 袁康. 可信算法的法律规制[J]. 东方法学, 2021(3): 5-20. https://doi.org/10.3969/j.issn.1007-1466.2021.03.002