

论算法解释权的正当性与制度建构

郑秋瑶

华东政法大学国际法学院, 上海

收稿日期: 2023年10月20日; 录用日期: 2023年11月17日; 发布日期: 2023年11月24日

摘要

算法“效率优先”的价值取向与现代社会自由、平等精神相抵触, 消解了人的主体性, 与人的诸多基本权利相冲突。基于边沁的功利主义理论和利益平衡原理, 法律对算法的规制是必要且正当的。美国、欧盟对算法解释权的立法进程, 对我国算法规制的完善有所裨益。我国算法规制的建构, 应立足本国国情, 加强行业自律监管, 对算法解释权的具体制度作出更进一步的规范指引, 为算法的继续发展提供坚实的法律保障。

关键词

算法规制, 伦理基础, 算法黑箱, 算法解释权

On the Legitimacy and Institutional Construction of Algorithm Interpretation Right

Qiuyao Zheng

School of International Law, East China University of Political Science and Law, Shanghai

Received: Oct. 20th, 2023; accepted: Nov. 17th, 2023; published: Nov. 24th, 2023

Abstract

The value orientation of “efficiency first” in algorithms contradicts the spirit of freedom and equality in modern society, dispels human subjectivity, and conflicts with many basic rights of human beings. Based on Bentham’s utilitarianism theory and interest balance principle, it is necessary and legitimate to regulate algorithms. The legislative process of the algorithm interpretation right in the United States and the European Union is helpful to the improvement of the algorithm regulation in China. The construction of algorithm regulation in our country should be based on the ac-

tual conditions of our nation. The self-discipline supervision of the industry should be strengthened, and further normative guidance for the specific system of algorithm interpretation right should be made, so as to provide solid legal guarantee for the continued development of algorithm.

Keywords

Algorithm Regulation, Ethical Basis, Algorithm Black Box, Algorithm Interpretation Right

Copyright © 2023 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. 问题提出

算法诞生于数学领域，但早已超出数学范畴，广泛影响着我们的日常工作生活。在线打车、网上购物、智能司法等程序都建立在算法运行控制的基础上。而现实中因算法运用不当导致的“大数据杀熟”等事件层出不穷，这表明在制度缺失的情形下，算法逻辑正改变着人类传统的生产、生活方式，导致系列算法伦理危机的出现。如何规避“算法黑箱”，实现算法透明与算法正义，是法律制度妥当回应算法等科技发展的时代性难题。

算法的深度学习能力与无限延展能力，深刻影响科技迭代发展，进而推动法律时代变革。算法社会实质上是人运用算法工具实现治理的社会，人并非算法的附庸。寻求社会治理与算法运行之间的契合点，构建智能社会合理的算法运行秩序，法学界须作出一定的理论回应与制度设计。那么，算法给人们生活带来“智能”“精确”“便利”的同时，为现代社会的价值观带来了哪些冲击？对算法进行法律规制的正当性基础是什么？如何理性认知算法，对算法解释权的具体制度作出更进一步的规范指引？本文将对上述问题进行探讨，求教大方之家。

2. 算法规制的正当性诘问

(一) 算法规制与现代社会的价值观冲突

算法“效率优先”的价值观与现代社会正义观产生分化。算法正义通常从后果正义出发，注重决策行为绩效的正当性，忽视对个人权利、个人尊严的保护。而在古典哲学看来，人是目的，不是手段[1]。只有体现平等、自由和权利等价值的社会，才是正义的社会[2]。促进现代社会持续健康发展应始终坚持以个人正义观为核心、为人类创造福祉的价值观。可见，算法“效率优先”的技术价值观，与现代社会“个人权利本位”的正义价值观之间存在一定分化。这就要求，在算法初始程序设计到算法推荐输入阶段，应坚持工具理性与价值理性的统一，加强“以人为本”的算法正义推荐权重，回归人类本质意义[3]。

算法“利益至上”的价值观与平等精神相背离。算法因其信息整合与分析能力日趋成为一种具备控制力属性的准公权力。这种算法权力打破了“权力——权利”相互制衡的既有格局，对传统平等权保护模式产生颠覆性冲击。以算法价格歧视为例，在各大购物平台上，算法常从不同消费者的购买习惯和购买力分析出其消费规律，通过不透明的机制向不同消费者做出个性化定价，此即大数据杀熟¹ [4]。相比起传统商业活动中的差异定价，算法的个性化定价无法让消费者知悉个性化定价是否满足经济理性以及

¹比如，互联网经营者常利用大数据分析技术故意对“熟客”提高价格，使“熟客”对商品价格陷入误解而作出购买选择，涉嫌价格欺诈。

公平价值理念的实质追求，严重损害了消费者的平等权、知情权等合法权利[5]。除个性化定价以外，算法的标签化定义，也使部分弱势群体收到了不平等对待。据研究显示，黑人、女性在就业领域得到的算法评价远低于一般群体；有色人种在银行信贷领域获得的贷款机会明显比白种人少[6]。算法歧视对公民平等权造成侵犯，对现代社会平等原则的挑战，值得社会重视与反思。

(二) 算法规制中人的主体性消解

算法为促进人类福祉而诞生，却逐渐将其服务对象——人类变成了创造效益的手段，逐渐消解了人的主体性特征。

在人工智能算法运用中，用户被迫适应着算法的逻辑。在进入人工智能算法时，用户必须提供大量个人信息并同意基于代码而建立的平台规则。进入算法后，用户又不得不遵守平台设定的信息筛选机制。通过分级标签语法及分级检索方式，算法限制了用户在平台的搜索范围。

在人工智能算法规制中，用户多数情况下处于被定义、被分析、被数据化的地位。算法的精准性来自其合理的代码逻辑，更来自其庞大的数据源，而其最大的数据源，就是算法平台中的用户信息。平台将用户被迫录入的个人信息数据化后储存在云端，不断充实着云端大数据。在此过程中，用户不是具有生命体征的人，而是变成了扁平化的数据对象、分析客体。

算法在对用户数据的分析中存在标签化效应的风险。算法主要是基于信息显著特征来对大数据进行分类筛选的。虽然算法分析能力超前，但其仍可能对用户作出不恰当的“污点”判断，且基于算法强大的预测分析能力，算法通常会一直存储用户的“污点”数据并基于此数据作出更多的不利判断。即使用户确实出现了“数据瑕疵”，现代社会的正义观也要求人们应当在尊重主体平等和多元化价值观的前提下对其作出相应的理解与忍让，不能因一个错误行为便剥夺个体未来应当取得的一切权利。同时，算法对信息的筛选标准通常是非公开的，公众无法对算法进行监督，这直接导致了“算法黑箱”的诞生。算法对用户数据的这种非公开、刻板的判断，无视了人的主体性地位，背离了现代社会的正义观。

(三) 算法规制正当性基础的反思

既然算法与现代社会正义、平等的价值观有所背离，对人的主体性产生消解作用，那么算法规制的正当性依据是什么？本文认为，可将边沁的功利主义理论作为算法规制的正当性基础。功利主义认为良好的立法是指引人们获得最大幸福和最小痛苦的艺术[7]。对任何一项政府的措施、法规或命令来说，若其增大共同体福祉的倾向大于它减小这一福祉的倾向时，就可以说它是符合功利主义要求的[8]。根据功利主义原理，探究是否该规制算法，即探究算法规制为社会大多数人带来的福祉是否能大于算法规制减小的福祉。

衡量算法规制带来或减少的福祉，应当遵循利益平衡原理。利益平衡原理是指利益主体根据一定的原则和方式对利益进行选择、衡量[9]。该原理要求，在解决利益纠纷时，应首先考量具体的利益对比，使受减损的利益降到最少。如当人身权利(如隐私权、平等权)与财产权利(如算法主体对算法所拥有的处分权、收益权)冲突时，应首先考虑如何通过财产的利益补偿来使人身权免受侵犯，即首先考虑两个具体形式的冲突。当在具体情况层面无法达成二者的利益平衡，则应在抽象利益层面考虑该二者所分别代表的利益，在日常生活中何者的保护价值更大。

算法规制触及的利益主体有二，一为算法用户(即使用相关软件的普通群众)，二为算法主体(即利用算法开发系统或软件的企业等)。触及的利益则分别是算法用户的人身权利——算法技术所侵犯的用户隐私权、平等权，和算法主体的财产性权利——规制算法后算法主体将受到的经济损失。根据利益平衡原理，在具体情况层面，由于算法解释权等规制措施仅适当限制了算法的自由度，不会对算法发展带来重大阻碍，故算法规制不会造成算法主体过多的权利减损。即使在具体情况层面无法达成二者的利益平衡，在抽象价值层面，比起算法主体的财产性权利，算法用户的人身权利应当优先受到保护。毕竟“人是目

的，而不是工具”，对基本权利的保护是所有保护位阶中具有最高价值的。同时，算法用户群体远大于算法主体群体，算法规制能保护基数更大的群体。因此，根据利益平衡原理可得，算法规制为社会带来的福祉将大于算法规制减小的福祉。

综上，算法规制行为增大该领域最大多数人福祉的倾向大于减小这一福祉的倾向，故算法规制行为是具有正当性基础的。在具备必要性要求的情况下，立法应认真考虑其规制行为的现实化。

3. 算法规制的法律实践难题

算法规制的实践中存在如算法黑箱等诸多难题。本部分将通过赋予用户算法解释权的角度，对算法黑箱问题进行回应。

(一) 美国算法解释权的有关法律实践

美国作为人工智能领先发展的国家之一，算法歧视的问题出现得较早。

2013年美国威斯康星州的 *State v. Loomis* 案中，法院使用人工智能算法计算得出的“再犯风险评估内容”对被告进行量刑，引发了被告的强烈抗议。被告认为，算法得出该评估内容的程序不透明，法官参考该评估内容的行为违反了正当程序原则。被告要求检查算法以及法院公开算法过程，但法院没有采用被告的提议。案件上诉后，州最高法院仍驳回了被告的请求，其借鉴了 *Gardner and Ska* 案的处理方式，以保护算法主体商业秘密为由拒绝了公开算法计算过程^[10]。该案终审判决引发了社会对算法歧视的激烈争论。

在美国密歇根州，三位美国公民向法院起诉了该州失业保险局。他们提出，密歇根融合数据自动化系统(MIDAS)对失业补贴的审查存在着极高的滥用风险，请求公开算法过程。通过调查，该案最后发现该系统对失业补贴申请的审核中，存在着高达93%的错误率，约有四万人成为算法重大错漏的受害者^[11]。

2016年，纽约市议员 *James Vacca* 提案要求对政府机构披露算法系统源代码进行强制立法。具体而言，凡是涉及用人工智能算法计算福利分配、治安和处罚的，都必须公开算法源代码。但由于“一刀切”式地公开源代码太过激进，侵犯企业核心商业秘密，侵犯自主经营权，打击企业创造性，该提案原始版本并没有被采用。2017年年底，纽约州通过了《算法问责法案》，该法案替代了原始草案中强制披露源代码的内容，设立了一个由算法系统专家和相应公民组织代表构成的事实调查工作组，专门负责推动算法应用的透明化与公平化^[12]。在该法案之前，《信用机会平等法》以及《公平信用报告法》都规定了算法主体负有向金融消费者解释对其不利的算法评分的义务，包括说明具体原因以及救济权利等，主要都是在回应算法让民众难以理解的问题^[13]。

为了改进智能投顾领域算法的自动化决策情况，2017年2月，美国证券交易委员会(直属美国联邦的独立准司法机构，简称“SEC”)发布了《智能投顾升级指导意见》，介绍了算法披露的相关内容。此外，2017年美国计算机学会公众政策委员会也发布了关于算法治理的六项指导原则，其中就包括了解释原则，这无疑也是有利于推进美国算法透明化、可解释化的。

(二) 欧盟算法解释权的有关法律实践

1972年法国颁布了《关于信息技术、数据文件和公民自由1978年1月6日第78-17号信息自由法案》。其中第10条规定：“不能仅根据自动化处理的与公民个性或侧写相关的数据做出涉及评价个人行为的决定”。第39条规定：“基于自动化处理做出的决策对数据主体造成法律影响时，数据主体有权知道该自动化处理的逻辑信息并有权反对自动化处理”²。这是最早提及算法解释权的法案，后被1978年《法国数据保护法》采纳，并对欧盟《通用数据保护条例》产生了深刻影响。

²Commission Nationale de l'Informatique et des Libertés (CNIL): LOI INFORMATIQUE ET LIBERTES ACT N° 78-17 OF 6 JANUARY 1978 ON INFORMATION TECHNOLOGY, DATA FILES AND CIVIL LIBERTIES.

2018年欧盟《通用数据保护条例》(GDPR)出台。该条例第71条第一款明确规定,数据主体有权“获得对决定的解释并质疑该决定”³。该条被人们视为是算法解释权法律化的标志。自此,算法解释权成为了一项实证化的权利,算法解释权在欧盟成员国内陆续获得了法律化属性,受到了广泛认可[13]。

2019年,又一部旨在推动人工智能算法透明化、公正化的法律——《平台和商户间公平性和透明度法》出台。该法要求机构应当披露算法参数以促进算法决策透明化,取代了以前对披露源代码的规定。其第5条规定:“线上中介服务提供商应在其条款中列出决定排名的主要参数以及这些主要参数相对于其他参数的相对重要的理由。”其后还附上了更为具体的披露内容规定,对法律实践提供了较为明确的指导。

可见,美国及欧盟在人工智能技术发展进程中都出现了算法黑箱的乱象,二者针对算法规制问题都进行了各层级的规制,从立法到行业规范性文件,确立了算法解释权,对相关主体算法解释权利形成了多方位保护,其对算法规制的先进经验值得我国学习。

(三) 我国规制算法黑箱的有关法律实践

近年来,我国算法技术飞速发展,但在发展的进程中也出现了许多乱象。2020年,作为会员用户的胡女士发现其通过携程网购买的酒店房间价格远比线下或非会员购买的价格要高昂,遂对携程网提出了诉讼⁴。此案被称为“大数据杀熟第一案”,揭开了我国“大数据杀熟”的大幕。此后,大量用户发现在滴滴打车平台上,同样的距离与路线,老用户收费比新用户高昂,苹果手机用户收费比安卓手机用户高昂⁵。凡此种种,皆为算法对个人数据进行自动化处理,进行差异化定价,谋求商家“利益最大化”,侵犯用户的合法权益。以上滥用算法技术的行为,将算法黑箱问题带入了我国公众视野[14]。

针对算法黑箱问题,早在2017年国务院印发的《新一代人工智能发展规划》就提到,应当“建立健全公开透明的人工智能监管体系”。虽然该文件提及了人工智能监管体系的公开透明,但并未提及人工智能算法过程的公开透明,对算法黑箱问题未作出实质性回应。

2021年出台的《中华人民共和国个人信息保护法》则着重强调了自动化算法决策的透明和决策结果的公平性、合理性⁶。该法第24条第三款规定:“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。”明确了公民具有算法解释权,是我国算法解释权法律化的标志。为我国实现算法公开化、透明化提供了法律基础。

然而,对比美国的算法规制,我国在行业规范方面略显欠缺,缺少如《智能投顾升级指导意见》此类规制具体行业算法运行的规范性文件,完全依赖国家规制的现状既不利于激发行业管理主动性,也不利于减轻国家管理压力;对比欧盟算法规制,我国算法解释权规制的不足主要体现在缺乏对算法用户实际行使算法解释权、知情权的具体化规范,《互联网信息服务算法推荐管理规定》⁷的相关条款也较为笼统,这使得算法用户的具体权利难以实际实现,算法解释权、知情权基本上处于架空状态。

³Parliament and Council of the European Union. General Data Protection Regulation, Recitals 71, “the right...to obtain an explanation of the decision reached after such assessment and to challenge the decision”.

⁴胡红芳诉上海携程商务有限公司侵权责任纠纷案,浙江绍兴市中级法院,(2021)浙06民终3129号。

⁵国家互联网信息办公室:《关于下架“滴滴出行”App的通报》,2021年7月4日。

http://www.cac.gov.cn/2021-07/04/c_1627016782176163.htm

⁶《中华人民共和国个人信息保护法》,第24条第一款:“利用个人信息进行自动化决策,应当保证决策的透明度和处理结果的公平合理。”

⁷如《互联网信息服务算法推荐管理规定》第7条:“算法推荐服务提供者应当落实算法安全主体责任,建立健全……数据安全和个人信息保护……等管理制度和技术措施,制定并公开算法推荐服务相关规则,配备与算法推荐服务规模相适应的专业人员和技术支撑。”第12条:“鼓励算法推荐服务提供者综合运用内容去重、打散干预等策略,并优化检索、排序、选择、推送、展示等规则的透明度和可解释性,避免对用户产生不良影响,预防和减少争议纠纷。”第16条:“算法推荐服务提供者应当以显著方式告知用户其提供算法推荐服务的情况,并以适当方式公示算法推荐服务的基本原理、目的意图和主要运行机制等。”

4. 算法解释权的制度完善

(一) 支持算法解释权条款的落地

算法解释权能让算法运行规则变得相对透明，是解决算法黑箱问题的重要举措。算法解释权的设立有利于在算法发展进程中始终维护正义、平等的现代社会价值观，支持算法用户在面对非正义、不平等的数据处理时依法进行权利救济；有利于在工具理性的发展中始终维护人的主体性地位，防止算法将人类数据化、客体化而偏离了造福人类的初衷。然而，国内对算法解释权的确立仍存在反对声音，包括认为人类没必要理解算法内容的^[15]，认为人类理解不了该内容的^[16]，以及认为算法透明化将带来巨大负面影响^[17]等。但本文认为，算法解释权的设立符合我国现实需求，不应因为些许不足而“因噎废食”。

第一，当决策对个人利益产生重大影响时，用户确有必要知道算法的过程。有观点认为，即便用户知道了算法运行规则，用户也无法制约算法主体做出的于己不利的决策^[18]^[19]。然而，这并不能成为排斥算法解释权的充分理由：一方面，这是公民知情权的构成之一。根据《消费者权益保护法》第8条第一款，消费者享有知悉其购买、使用的商品或者接受的服务的真实情况的权利。在人工智能领域，由于算法技术的复杂性，信息不对称的局面更有可能出现，算法用户的知情权更应当得到重视；另一方面，算法解释权是公民救济途径的重要基础。行使算法解释请求权，找出算法决策问题所在——这是“查漏”。向平台或有关部门申诉，请求变更或拒绝服从决策——这是“补缺”。用户行使算法解释权，是其获得救济、实现实质公平的必要条件，缺少了算法解释权，算法用户的权利救济就无从谈起。即便无法制约决策的作出，算法解释权依然为用户提供了有效的救济途径。

第二，我国《电子商务法》、《消费者权益保护法》无力规制算法领域的现实问题。反对设立算法解释权的声主要提及《电子商务法》第17条和第24条⁸。但这两个条文只是分别提及了数据主体的知情权和信息获取权，实际上数据主体如何行使知情权，如何获取目标信息，法条并没有具体规定。关于《消费者权益保护法》，反对声音主要提及第8条和第10条⁹。但实际上，算法在商业场景中多被定位为“工具”，不符合消费者权益保护法“商品”或“服务”的要求，用户往往无法通过援引该条文来获得算法解释^[20]。

第三，算法解释权披露的是算法说明，而不是源代码。算法解释权的设立，目的在于保护算法用户的合法权益不受算法黑箱及算法错误的侵犯。故算法披露的内容，应当仅限于权利救济的内容，比如告知用户算法在某个环节对某个变量做出了误判，进而对决策产生了影响。根据算法披露的这些提示，用户可以找到相关环节的负责人主张权利，最终实现权利救济。

第四，当算法解释权披露的是算法说明时，商业秘密泄露的可能性将会大幅降低。由于算法说明实质上是平台就相关数据给算法用户的一个答复，使用的文字应为用户可读的自然语言，源代码将继续受到平台的保护，因此不存在源代码泄露导致优质算法曝光的情况。即使平台披露的算法说明涉及部分商业秘密，由于算法说明文件是平台自主撰写的，由其未履行审慎检查义务而引发商业秘密泄露的后果应当由平台自己承担。

综上，应坚决支持算法解释权条款的落地，实现算法解释权的落地，推动算法进一步透明化、公正

⁸《中华人民共和国电子商务法》第17条：“电子商务经营者应当全面、真实、准确、及时地披露商品或者服务信息，保障消费者的知情权和选择权。电子商务经营者不得以虚构交易、编造用户评价等方式进行虚假或者引人误解的商业宣传，欺骗、误导消费者。”第24条：“电子商务经营者应当明示用户信息查询、更正、删除以及用户注销的方式、程序，不得对用户信息查询、更正、删除以及用户注销设置不合理条件。电子商务经营者收到用户信息查询或者更正、删除的申请，应当在核实身份后及时提供查询或者更正、删除用户信息。用户注销的，电子商务经营者应当立即删除该用户的信息；依照法律、行政法规的规定或者双方约定保存的，依照其规定。”

⁹《中华人民共和国消费者权益保护法》第8条：“消费者享有知悉其购买、使用的商品或者接受的服务的真实情况的权利。”；第10条“消费者在购买商品或者接受服务时，有权获得质量保障、价格合理、计量正确等公平交易条件，有权拒绝经营者的强制交易行为。”

化、合理化。

(二) 加强行业自律监管

我国多从政府管控的层面作出规定，行业自律管控方面尚存在较大的实践空白。实践中出现算法歧视问题时，算法主体极少持积极干预的态度，他们并不会主动承担算法带来的伦理责任以及法律责任^[21]。在这种语境下，规制算法，实质上就是通过法律规制算法主体的行为，因此算法主体所在的相关行业协会与政府的多主体合作就显得尤为重要。强化对算法的规制，不能单靠外部监管，还要激发算法主体自主管理的意识；不能单靠自上而下的政府监管，还要兼顾行业自律监管。

应当借鉴美国在算法规制领域的实践经验，加大行业规范方面的投入，以国家新一代人工智能治理专业委员会发布的《新一代人工智能治理原则——发展负责任的人工智能》与《新一代人工智能伦理规范》等文件为指导，敦促算法相关行业制定行业自律规范，积极推动我国行业协会或学会建立算法伦理准则和监管标准^[22]，让行业形成自律监管的良好秩序与风气，推动“行业的事情行业管”，充分发挥算法主体的主观能动性。另外，还应考虑到各行业的特殊性，不可将同一规范机械地适用于所有行业。应根据不同行业的行业特点与需求，制定适应该行业具体实践的自律规则，如智能投顾领域应制定适应智能投顾行业现实需求的规范，电子智能购物领域应制定适应电子智能购物行业现实需求的规范，使行业自律规范真正做到促进行业长远发展，共同推动算法治理能力与治理体系的现代化。

(三) 对算法解释权的具体制度作出更进一步的规范指引

《个人信息保护法》仅对算法透明做出了原则性规定，规定当算法决策做出的决定对个人权益有重大影响时，个人有权要求信息处理者予以说明。而算法解释权解释的内容是什么，算法解释需要解释到什么程度，在什么时间段可以主张算法解释权，算法解释权实现的方式，权利的救济途径，以及行使算法解释权需要考虑到的相关因素等，都没有同步的立法说明。

关于解释内容，2017年2月美国SEC发布的《智能投顾升级指导意见》对算法披露的内容做出了规定。该指导意见规定，算法披露的内容应包括算法的功能介绍、算法的假设与限制、特定风险解释、利益冲突解释等¹⁰。换言之，算法披露的算法说明是自然语言，而不是源代码。算法解释提供的是个人能理解与运用的自然语言，但由于算法解释的问题较为超前，国内可供参考的司法案例较少，对算法解释权具体内容的认定仍存在困难。可以先以《个人信息保护法》第24条的原则性规定为认定基础，赋予法官自由裁量权以积累实务经验，定时进行总结归纳，颁布进一步规定算法解释内容的司法解释或适时更新法律。另外，也可以采取列举式、非要件式的立法，如《反垄断法》中对认定经营者市场支配地位的规定一样，列出需要考虑的因素，而不严格规定算法解释内容必须满足的要件^[23]。

算法的解释程度，极大取决于算法的解释时间、解释领域等相关因素。

算法解释程度与算法解释时间相关。面对算法决策之前的解释，如告知环节的算法解释，相关立法可以采取模糊性解释以减少企业负担，提高决策效率，同时还能应对侵犯商业秘密等问题；面对算法决策之后请求作出的解释，多是存在个人对算法决策结果不满且该结果极大影响个人利益的情况，此时应当要求企业作出更为具体的解释。如在算法对贷款资格、贷款额度的认定问题上，决策之前银行可以向用户告知其算法是否考量收入、年龄、性别、资产配置状况等情况，但没有必要对某个参数具体如何影响算法结果进行告知；而在决策作出后，依用户申请，可以针对该算法决策结果作出具体的说明。

算法解释程度也与算法解释领域相关。在个案影响较小、可替代性方案较多的领域，算法解释的程度应远低于对社会具有重大公共性影响、对个体造成重大影响的领域。前者如游戏皮肤的设定，后者如机器人进行的外科手术。领域不同，算法解释的必要性、迫切性不同，算法解释的程度也应作相应区别。

¹⁰SEC, Guidance Update on Robo-Advisers.

除了算法解释内容和程度以外, 算法解释权的行使方式以及救济途径也应当作出具体化规定。对于个人来说, 人工解释更有利于获得个人的信任, 故从保护个人算法解释权的角度来说, 应当以人工解释为优先选择[24]。而人工解释的成本高、效率低, 大规模人工解释显然不具有客观现实性。对于企业来说, 机器解释更有利于提高运作效率, 故从促进企业发展的角度来说, 机器解释也具有重大意义。

然而, 算法所处场景的不同, 也可能导致使用主体、针对对象等的不同, 在不同场景中机械地运用同一解释方法的做法显然也是不可取的[25]。应当结合算法领域影响力大小、用户数量、领域特征等因素, “场景化”决定解释方式, 有针对性地决定解释方式[26]。同时, 没有保护的權利就仿佛空中楼阁, 若没有相关的救济机制, 算法解释权将始终无法落到实处。应设立相关救济途径, 从企业自身规制、行业监管投诉机制、政府工商管理等部门等角度共同保护算法解释权的行使, 使个人能够真正行使权利, 维护合法权益, 实现算法规制的立法目的。

5. 结语

综上, 算法与现代社会追求正义等价值观产生了冲突, 消解了人的主体性。根据功利主义及利益平衡原理, 算法规制具有充分的正当性基础。对比美国及欧盟算法规制的先进经验, 我国在《个人信息保护法》之后, 应继续推进算法解释权条款的实际落地, 加强行业自律监管, 对算法解释权具体制度作进一步规范指引, 实现算法的全方位规制, 推动算法治理能力与治理体系的现代化, 推动经济社会的长足发展。

参考文献

- [1] 康德. 实践理性批判[M]. 韩水法, 译. 北京: 商务印书馆, 2003: 95.
- [2] 姚大志. 何谓正义: 罗尔斯与哈贝马斯[J]. 浙江学刊, 2001(4): 10-16.
- [3] 张爱军, 王首航. 算法: 一种新的权力形态[J]. 治理现代化研究, 2020, 36(1): 39-45.
- [4] 廖建凯. “大数据杀熟”法律规制的困境与出路——从消费者的权利保护到经营者算法权力治理[J]. 西南政法大学学报, 2020, 22(1): 70-82.
- [5] 贾章范. 论算法解释权不是一项法律权利——兼评《个人信息保护法(草案)》第二十五条[J]. 电子知识产权, 2020(12): 49-61.
- [6] 张恩典. 大数据时代的算法解释权: 背景、逻辑与构造[J]. 法学论坛, 2019, 34(4): 152-160.
- [7] 边沁. 政府片论[M]. 沈叔平, 等, 译. 北京: 商务印书馆, 1995: 29.
- [8] 张玉堂. 边沁功利主义分析法学研究[D]: [博士学位论文]. 上海: 华东政法大学, 2010: 55.
- [9] 冯晓青. 论利益平衡原理及其在知识产权法中的适用[J]. 江海学刊, 2007(1): 141-146.
- [10] Wisconsin Supreme Court (2016) Case: State v. Loomis | 881 N.W.2d 749. <https://law.justia.com/cases/wisconsin/supreme-court/2016/2015ap000157-cr.html>
- [11] Eltham, B. (2017) High Farce: The Turnbull Government’s Centrelink “Robo-Debt” Debacle Continues to Grow.
- [12] Powles, J. 纽约市尝试对算法问责——政策有待完善, 但行动敢为人先[EB/OL]. 姜开锋, 译. 大数据和人工智能法律研究院. <https://mp.weixin.qq.com/s/g1OJG71osEQn0hiEL4hUJg>, 2018-12-19.
- [13] 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.
- [14] 王利明. 敏感个人信息保护的基本问题——以《民法典》和《个人信息保护法》的解释为背景[J]. 当代法学, 2022, 36(1): 12.
- [15] 卢克·多梅尔. 算法时代: 新经济的引擎[M]. 胡小锐, 钟毅, 译. 北京: 中信出版集团, 2016: 123.
- [16] 张淑玲. 破解黑箱: 智媒时代的算法权力规制与透明实现机制[J]. 中国出版, 2018(7): 49-53.
- [17] 徐凤. 人工智能算法黑箱的法律规制——以智能投顾为例展开[J]. 东方法学, 2019(6): 78-86.
- [18] 徐文. 个人推论数据是如何被藏匿的? [J]. 社会科学, 2020(10): 107-118.
- [19] Wachter, S., et al. (2017) Why a Right to Explanation of Automated Decision Making Does Not Exist in the General

Data Protection Regulation. *International Data Privacy Law*, 7, 76-99. <https://doi.org/10.1093/idpl/ix005>

- [20] 张凌寒. 商业自动化决策的算法解释权研究[J]. 法律科学: 西北政法大学学报, 2018, 36(3): 65-74.
- [21] 金梦. 立法伦理与算法正义——算法主体行为的法律规制[J]. 政法论坛, 2021(1): 29.
- [22] 张吉豫. 构建多元共治的算法治理体系[J]. 法律科学(西北政法大学学报), 2022, 40(1): 115-123.
- [23] 江溯. 自动化决策、刑事司法与算法规制——由卢米斯案引发的思考[J]. 东方法学, 2020(3): 76-88.
- [24] Vladeck, D.C. (2014) Machines without Principals: Liability Rules and Artificial Intelligence. *Washington Law Review*, 89, 117-150.
- [25] 丁晓东. 论算法的法律规制[J]. 中国社会科学, 2020(12): 138-159.
- [26] 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.