

论算法的法律规制

丁 晓 东

摘 要：算法崛起对法律规制提出了挑战，它可能挑战人类的知情权、个体的隐私与自由以及平等保护。作为人机交互的算法决策机制，算法并非价值中立，具备可规制性。算法公开、个人数据赋权、反算法歧视是传统的算法规制方式，但机械适用这些方式面临可行性与可欲性难题。算法公开面临技术不可行、公开无意义、用户算计与侵犯知识产权等难题，个人数据赋权面临个体难以行使数据权利、过度个人数据赋权导致大数据与算法难以有效运转等难题，反算法歧视面临非机器算法歧视、身份不可能完全中立、社会平等难以实现等难题。传统算法规制路径面临困境的根本原因在于忽视算法的场景性，算法可能因为运用算法主体的不同、针对对象的不同以及涉及问题的不同而具有不同的性质。因此，算法规制应采取场景化的规制路径，根据不同场景类型对算法采取不同的规制方式，以实现负责任的算法为目标。在算法场景化规制原则的指引下，可以构建算法公开、数据赋权与反算法歧视等算法规制的具体制度。

关键词：人工智能 算法 算法公开 数据权利 算法歧视 场景化规制

作者丁晓东，中国人民大学法学院、未来法治研究院副教授（北京 100872）。

一、问题的提出：算法崛起的法律挑战

随着大数据与人工智能时代的到来，算法开始呈现越来越大的影响，日益成为社会关注的问题。以往，算法更多是数学家或程序员所关注的对象，算法主要在数学运算或实验室的场景下发生作用。到了大数据与人工智能时代，算法开始在越来越多的应用场景中被用于决策或辅助决策。^① 随着未来大数据与人工智能更深度地

① 例如在外卖送餐行业，算法被美团、饿了么等企业广为应用，用于提高送餐效率和压缩送餐时间；在新闻资讯与娱乐领域，抖音、快手、今日头条等利用算法进行个性化推荐与分发，以提高新闻与娱乐资讯的传播效率；在电商领域，淘宝、京东等购物网站利用算法对个体进行个性化推荐，以大幅促进销量；在搜索领域，百度等搜索引擎广泛运用以算法为核心的信息检索；在评级网站中，豆瓣、大众点评等利用算法实现对

运用，未来算法的应用场景将更为广泛，在自动驾驶、公共管理、司法等领域与场景中，算法都将发挥举足轻重甚至是决定性的作用。算法在社会中的广泛运用带来很多正面效应，它可以大幅提高决策效率，为消费者或用户提供更精准的服务。同时，算法崛起也带来很多挑战。^①2020年引起社会广泛关注的外卖算法系统，一些互联网平台利用算法设置外卖骑手的配送时间，送餐时间被压缩得越来越短，对外卖骑手的生命健康造成严重威胁。^②而且，这个算法系统采用自动化的机器决策，骑手很难理解和提出抗议。

从法律的角度看，算法从几个方面挑战了法律的一些基本原则。首先，算法黑箱可能挑战人类决策的知情权与自主决策。一般认为，在所有重要事务中，做出最终决策的主体应当是人，“人类选择是私人与公共生活的不可分割与根本性的一部分”。^③但是在算法社会中，很多时候不透明的算法——而非人——成为决策主体。如果不加检验地以机器决策替代人类决策，人类的自主性可能面临严峻考验。

其次，算法可能威胁个体的隐私与自由。算法常常建立在个人数据的收集之上，通过结合大数据运算与个人数据进行个性化推送。但这种对个体偏好的预测与迎合可能损害公民个体的自主性，因为这种个性化推荐可能使个体困于信息茧房（information cocoons）。个体受限于算法的控制，能接受到的信息只是算法根据个体偏好而筛选出来的信息，而不是那些更加中立、可能促使个体反思自身前见的信息，甚至不是随机性的信息。长期如此，个体真正的自由可能受到威胁。

最后，算法可能导致歧视与偏见。平等是一个社会的基本价值，算法的技术特征使得有人认为，算法有助于解决歧视与偏见问题。但算法也可能常常暗含歧视与偏见，甚至放大人类的偏见。^④当人们设计算法与输入数据时，此类算法或数据可能就不具有代表性。例如一种进行人类脸部识别的算法，如果算法所使用的数据都是白人男性的数据，那么就可能无法识别黑人、亚裔或女性，对黑人、亚裔或女性形成歧视。^⑤算

相关主体与对象的评级，为消费者提供引导。

① 参见张文显：《构建智能社会的法律秩序》，《东方法学》2020年第5期。

② 参见赖祐萱：《外卖骑手，困在系统里》，《人物》2020年第8期。

③ Michal S. Gal, “Algorithmic Challenges to Autonomous Choice,” *Michigan Technology Law Review*, vol. 25, no. 1, 2017, p. 60.

④ Jeremy Kun, *Big Data Algorithms Can Discriminate, and It's Not Clear What to Do About It*, *The Conversation*, Aug. 13, 2015, <http://theconversation.com/big-data-algorithms-can-discriminate-and-its-not-clear-what-to-do-about-it-45849>, June 6, 2018; Ramona Pringle, *When Technology Discriminates: How Algorithmic Bias Can Make an Impact*, *CBC*, Aug. 10, 2017, <http://www.cbc.ca/news/technology/algorithms-hiring-bias-ramona-pringle-1.4241031>, June 16, 2018.

⑤ Clare Garvie and Jonathan Frankle, *Facial Recognition Software Might Have a Racial Bias Problem*, *The Atlantic*, April 7, 2016, <https://www.theatlantic.com/technology/archive/>

法可能会固化歧视与偏见，使其更难被发现、更难以矫正。

针对算法崛起所带来的法律挑战，传统法律规制主要采取三种方式加以应对：算法公开、个人数据赋权与反算法歧视。其中算法公开的方式认为，算法崛起带来的最大挑战在于算法的不透明性，人们常常感到它是一个黑箱，无法理解它的逻辑或其决策机制。因此，应当对算法进行公开，使得算法能够为人们所知晓。个人数据赋权的方式认为，影响个体的算法都是建立在对个人数据的收集与应用基础上的，因此，应当对算法所依赖的对象——数据——进行法律规制，通过赋予个体以相关数据权利来规制算法。最后，反算法歧视的方式认为，算法中常常隐含了很多对个体的身份性歧视，因此应当消除算法中的身份歧视，实现身份中立化的算法决策。

如何看待算法崛起对法律规制的挑战？本文力图对这一问题进行较为全面的分析。首先，本文对算法进行界定，指出算法的本质在于人机交互决策，因此算法不同于纯粹的科学或工具，具备法律上的可规制性。其次，本文对算法公开、个人数据赋权与反算法歧视这三种传统的算法法律规制方式进行分析，指出机械地使用这三种方式可能导致可行性与可欲性问题。再次，本文分析算法规制的基本原理，指出算法常常因为场景的变化而具有不同属性。为此，算法规制必须采取场景化的规制路径，根据算法运用的不同主体、算法针对的不同对象、算法涉及的不同领域而进行不同类型的规制，以形成可信赖和负责任的算法。最后，本文对算法公开、个人数据赋权与反算法歧视的制度进行初步建构。

二、算法的界定与可规制性

在分析算法规制之前，需要先对算法进行界定。算法可作狭义界定，也可作广义或中义界定。^①从狭义角度看，算法源于数学与计算科学，用于表述解决数学与计算科学难题的一系列规则，例如数据结构算法、数论与代数算法、计算几何算法、图论算法、动态规划以及数值分析、加密算法、排序算法、检索算法、随机化算法、并行算法、厄米变形模型、随机森林算法等。因此，狭义的算法可被视为纯粹的科学或技术。^②根据这种理解，有的学者将算法界定为一系列“已被编码的程序”或

2016/04/the-underlying-bias-of-facial-recognition-systems/476991/, June 18, 2018.

① Jean-Luc Chabert ed., *A History of Algorithms: From the Pebble to the Microchip*, New York: Springer, 1999, p.1.

② 以科技的视角看待算法，可参见 Christopher W. Clifton, Dierdre K. Mulligan and Raghu Ramakrishnan, “Data Mining and Privacy: An Overview,” in Katherine J. Strandburg and Daniela Stan Raicu eds., *Privacy and Technologies of Identity: A Cross-Disciplinary Conversation*, New York: Springer, 2006, pp.191-208.

者“为了快速实现某个目标对一组数据进行处理逻辑步骤”。^①

算法也可作广义界定。随着社会的发展与科技的广泛应用，算法的概念不仅被应用于数学与计算科学领域，也被应用于很多社会科学领域。在这些语境下，算法被宽泛地界定为所有决策程序或步骤，而不仅是与机器相关的自动化决策。^②从广义的算法概念出发，有学者认为算法可被视为一种建构社会秩序的特殊理性形式。^③还有学者提出，算法可以被界定为“为实现某一目标而明确设定的一系列步骤”。^④

本文采取介于狭义与广义之间的中义算法定义。这种算法定义将算法界定为人类和机器交互的决策，即人类通过代码设置、数据运算与机器自动化判断进行决策的一套机制。^⑤在当前算法大规模介入人类生活决策的背景下，采取这一界定更符合本文所要描述与分析的对象。本文所要分析的是人们利用机器来进行自动化决策或辅助决策的算法，这种算法并非数学或计算机科学意义上的算法，也并非纯粹关于人类行为的决策算法，^⑥这一过程既有人类决策，也有机器的自动化判断。

从人机交互的角度分析算法，可以深入理解算法的非中立性与法律上的可规制性。一种观点认为，算法是一种科学技术或工具，在价值上是完全中立的。法律只需对算法产生的后果进行应对，而不需要对算法本身进行法律规制。就像数学公式或手机一样，当犯罪分子利用数学公式运算而盗取比特币，或者利用手机进行诈骗，法律并不将数学公式或手机纳入规制范围。但现代社会中的算法并非实验室里的算

① Tarleton Gillespie, "The Relevance of Algorithms," in Tarleton Gillespie, Pablo J. Boczkowski and Kirsten A. Foot eds., *Media Technologies: Essays on Communication, Materiality, and Society*, Cambridge: The MIT Press, 2014, p.167.

② Danielle Keats Citron, "Technological Due Process," *Washington University Law Review*, vol. 85, no. 6, 2007, pp.1249-1313; Oscar H. Gandy, Jr., "Engaging Rational Discrimination: Exploring Reasons for Placing Regulatory Constraints on Decision Support Systems," *Ethics and Information Technology*, vol.12, no.1, 2010, pp. 29-42.

③ David Beer, "Power Through the Algorithm? Participatory Web Cultures and the Technological Unconscious," *New Media & Society*, vol. 11, no. 6, 2009, pp. 985-1002; David Lyon, "Surveillance as Social Sorting: Computer Codes and Mobile Bodies," in David Lyon ed., *Surveillance as Social Sorting: Privacy, Risk, and Digital Discrimination*, New York: Routledge, 2003, pp.13-30.

④ Joshua A. Kroll, et al., "Accountable Algorithms," *University of Pennsylvania Law Review*, vol.165, no.3, 2017, pp. 633, 640.

⑤ 不同人设想的算法所包括的代码、数据或生态系统常有差异。参见 *Algorithmic Accountability: Applying the Concept to Different Country Contexts*, July, 2017, p.5, http://webfoundation.org/docs/2017/07/Algorithms_Report_WF.pdf, May 3, 2018.

⑥ 因此，算法在法律上常被认为具有多重性质。相关探讨参见陈景辉：《算法的法律性质：言论、商业秘密还是正当程序？》，《比较法研究》2020年第2期；左亦鲁：《算法与言论——美国的理论与实践》，《环球法律评论》2018年第5期。

法，而是已经深度介入社会生活的方方面面。在很多情形中，算法已经成为社会价值判断的一部分。以今日头条为例，当其宣称自身算法中立时，其实不过是采取了另一种价值立场：以商业价值与经济利益作为新闻媒体的最大价值，利用“算法实现最大推送量，获得最高点击率”，“追求利益的最大化。”^①

总而言之，算法与作为犯罪工具的数学公式与手机非常不同。因为就数学公式或手机与犯罪行为之间的联系而言，数学公式或手机并没有嵌入价值判断，^②但作为决策机制或辅助决策机制的算法却深深地嵌入了价值判断。如果此时仍然坚持算法中立性或技术中立性的立场，对价值与伦理问题视而不见，就很可能忽视算法对人类价值伦理所带来的挑战。算法或算法系统并不是价值中立的，相反，算法隐含的价值立场使得有必要对算法进行规制。

三、算法规制的传统方式及其困境

在界定算法的定义并阐述算法的可规制性后，可在此基础上分析算法规制的三种传统方式：算法公开、个人数据赋权与反算法歧视。深入分析这三种方式，会发现机械运用传统规制方式造成的困境。

（一）算法公开

算法公开的方式认为，算法崛起带来的最大挑战在于算法的不透明性，因此，应当公开算法，使得算法能够为人们所知晓。例如，丹妮尔·西特鲁恩（Danielle Keats Citron）和弗兰克·帕斯奎尔（Frank Pasquale）通过对美国征信行业算法黑箱的分析，指出征信行业的算法缺乏透明性。^③他们主张，负责保护消费者权益的美国联邦贸易委员会应强化监管，实现算法透明化。联邦贸易委员会对相关评级机构不仅可以监督评级机构的数据组，也可对其源代码进行审查，以探明评级机构的算法是否违反反歧视法的相关规定，是否存在程序员的偏见或机器学习的偏见。^④

与算法公开类似的是算法的可解释性。相比算法公开，算法的可解释性具有更

① 宣言：《不能让算法决定内容》，《人民日报》2017年10月5日，第4版。

② 手机也存在价值立场问题，例如手机是否考虑了对盲人等残障人士的需求。因此从广义上看，任何技术都不是中立的，都存在价值伦理的问题。相关讨论参见 Jack Balkin, “The Path of Robotics Law,” *California Law Review Circuit*, vol. 16, no. 45, 2015, pp. 45-60; 郑玉双：《破解技术中立难题——法律与科技之关系的法理学再思》，《华东政法大学学报》2018年第1期。

③ Danielle Keats Citron and Frank Pasquale, “The Scored Society: Due Process for Automated Predictions,” *Washington Law Review*, vol. 89, no. 1, 2014, p. 1.

④ Danielle Keats Citron and Frank Pasquale, “The Scored Society: Due Process for Automated Predictions,” p. 25.

高的要求，因为前者主要强调算法运算数据的公开以及源代码的公开，^①而后者不仅强调公开，而且强调算法必须为数据主体或终端用户所理解。在欧洲，这种对算法解释权的要求已被很多研究者所提倡。“可解释的人工智能”被冠以“XAI”（Explainable Artificial Intelligence）的专门术语，成为众多专家关注的领域。^②

1. 算法公开的可行性

算法公开首先会面临可行性的难题。算法公开假定，算法是一个黑箱，只要打开这个黑箱，算法就会暴露在阳光下，为人们所知晓。但现实是，算法黑箱的原理与国家机密或商业秘密的原理并不相同，算法黑箱是由算法的技术性特征造成的，而非人为刻意保持造成的。^③在大数据与人工智能时代，为了提高算法的准确性，算法的复杂性往往会加强，一个企业或网站的算法往往由数十上百甚至上千的工程师写作完成，同时机器学习中的算法是经过训练数据集而不断进行调整优化而产生的，并非完全按照工程师编写的代码而产生的。在这样的背景下，公开算法的源代码和架构并无太多意义，因为公开并不能提供有效的对算法的说明。^④

基于这一原理，要求人工智能中的算法实现透明性，这“听上去很好，但实际上可能没什么帮助，而且可能有害”。^⑤实践中的算法公开也印证了这一点。社交新闻网站 Reddit 曾对一部分专家公开了其网站的排名算法，但研究者发现，专家对算法到底如何运转常常存在分歧，对于算法如何真正运行，专家其实也很难完全理解。^⑥专家尚且如此，对于普通人而言，算法公开更无实质意义。

至于算法的可解释性，除面临和算法公开同样的困境之外，还面临着大数据带来的因果关系难题。传统上人们主要通过因果关系了解世界，通过把握世界中的因

① 在这个意义上，算法公开更多是程序性要求，以正当程序原则规制算法的分析，参见 Danielle Keats Citron, “Technological Due Process,” p.1256.

② David Gunning, *Explainable Artificial Intelligence (XAI)*, U. S. Defense Advanced Research Projects Agency, [https://www.cc.gatech.edu/~alanwags/DLAI2016/\(Gunning\)%20IJCAI-16%20DLAI%20WS.pdf](https://www.cc.gatech.edu/~alanwags/DLAI2016/(Gunning)%20IJCAI-16%20DLAI%20WS.pdf), Aug.6, 2018.

③ 参见沈伟伟：《算法透明原则的迷思——算法规制理论的批判》，《环球法律评论》2019年第6期。

④ Max Kuhn and Kjell Johnson, *Applied Predictive Modeling*, New York: Springer-Verlag, 2013, p.50.

⑤ Curt Levey and Ryan Hagemann, “Algorithms with Minds of Their Own,” *The Wall Street Journal*, November 12, 2017, <https://www.wsj.com/articles/algorithms-with-minds-of-their-own-1510521093>, July 30, 2018.

⑥ Christian Sandvig et al., “Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms,” <http://www-personal.umich.edu/~csandvig/research/Auditing%20Algorithms%20-%20Sandvig%20-%20ICA%202014%20Data%20and%20Discrimination%20Preconference.pdf>, July 30, 2018.

果关系，决策主体就能理解世界发展变化的逻辑，为未来决策提供借鉴和依据。但对于大数据，很多专家都指出，大数据所力图发现的并不是因果关系，而是相关关系。^①正如维克托·迈尔-舍恩伯格所说：“当我们说人类是通过因果关系了解世界时，我们指的是我们在理解和解释世界各种现象时使用的两种基本方法：一种是通过快速、虚幻的因果关系，还有一种就是通过缓慢、有条不紊的因果关系。大数据会改变这两种基本方法在我们认识世界时所扮演的角色。”^②在这个意义上，要求所有算法都必须满足可解释性的要求，实际上是要求相关主体完成一项不可能的任务，因为基于大数据的算法与可解释性所要求的因果关系阐释具有完全不同的逻辑。^③

2. 算法公开的可欲性

在有些情形中，算法的透明性与可解释性可以实现或部分实现，但算法的透明性与可解释性仍可能存在可欲性问题。一旦算法被公开或解释给相关主体，算法就可能面临被相关主体钻空子或者算法知识产权被侵犯的风险。

算法公开首先可能导致算计（gaming）的问题。算法公开的初衷在于防止算法黑箱带来的滥用，通过公开与解释算法来监督算法，但算法一旦公开，相关主体就有可能利用和篡改算法，通过设置相应的参数和制造数据达成自己的目的，从而损害其他主体的正当权益。例如在搜索算法中，谷歌（Google）曾经依赖于一种叫作 PageRank 的算法确定搜索排序，这种排序方法主要根据 META 标签、关键字等参数进行排序。当谷歌公开这一算法之后，很多网站就开始利用此类算法，在自己的网页内嵌套符合 PageRank 算法的具有隐藏内容的网页，以此达到提高网站在谷歌搜索结果页面排名靠前的目的。经过此类设计后，一些与搜索内容并不相关的网页也被谷歌搜索结果搜索并排在前面。^④出于防止篡改算法的考虑，如今谷歌采取考虑上百种参数而综合判断的搜索算法，维持了搜索算法的秘密性。^⑤

即使算法不公开，对于算法的算计也非常普遍。各类评级网站充斥着“水军”，

① 参见维克托·迈尔-舍恩伯格、肯尼思·库克耶：《大数据时代：生活、工作与思维的大变革》，盛杨燕、周涛译，杭州：浙江人民出版社，2013年，第67—94页。

② 维克托·迈尔-舍恩伯格、肯尼思·库克耶：《大数据时代：生活、工作与思维的大变革》，第84页。

③ 因果关系是一个非常复杂的哲学问题，人类的许多认知过程其实也受相关关系的影响，参见 Amos Tversky and Daniel Kahneman, “Availability: A Heuristic for Judging Frequency and Probability,” *Cognitive Psychology*, vol. 5, no. 2, 1973, pp. 207-232.

④ John Faber, “How to Future-Proof Your Search Ranking,” Chapter Three, April 2, 2018, <https://www.chapterthree.com/blog/how-to-future-proof-your-search-ranking>, July 31, 2018.

⑤ Danny Sullivan, “Google Uses RankBrain for Every Search, Impacts Rankings of ‘Lots’ of Them,” Search Engine Land, June 23, 2016, <https://searchengineland.com/google-loves-rankbrain-uses-for-every-search-252526>, July 31, 2018.

他们通过人工或机器的方式不正当地改变某些评分，以此影响消费者的消费行为。为防止此类行为，网站常常在相关算法中嵌入打击刷分行为的参数或设计。^① 如果法律对算法公开与算法可解释性进行强制性规定，要求企业公开或解释此类参数或设计，那么“水军”就可以更快地调整刷分策略，更精准地利用网站的算法漏洞。

除了算计的问题，算法公开也可能导致知识产权侵权或算法被抄袭。在当前的算法监管中，很多商业机构都以商业秘密的名义拒绝监管，一些学者对商业机构的这种做法感到愤怒，因为它实际上完全将算法置于法律监管的空白地带。^② 此种担忧有一定道理，尤其当企业算法具有一定的公共属性时，不宜将算法都视为商业秘密。但从商业秘密的定义以及知识产权原理看，并不能完全否认算法作为商业秘密的特征。从商业秘密的定义看，很多算法的确满足了秘密性、经济性等要件。^③ 而从原理层面看，无论是基于劳动价值论，还是基于促进投资的功利主义原理，将算法视为商业秘密予以知识产权保护，都具有一定的正当性基础。^④

(二) 个人数据赋权

从个人数据赋权的角度应对算法问题，与从算法公开与算法可解释性的角度监管算法具有重叠之处。但个人数据赋权的相关法律更多依赖于个人对数据的控制，更试图从算法所依赖的对象——数据——的角度切入对算法进行法律规制。^⑤

个人数据赋权的相关法律首先赋予个体一系列的数据权利，强化个人对个人数据的知情与控制。例如欧美的很多个人数据立法都赋予个人数据收集时的知情选择权、数据访问权、数据更正权、数据删除权、反对自动化处理的权利等一系列权利。个人数据赋权对数据控制者与处理者施加责任，要求数据控制者与处理者满足个人的一系列数据权利，承担维护个人数据安全与数据质量等责任。^⑥

① 例如利用贝叶斯模型识别和剔除水军。参见张艳梅等：《基于贝叶斯模型的微博网络水军识别算法研究》，《通信学报》2017年第1期。

② 参见 Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information*, Cambridge, MA: Harvard University Press, 2015, p.10.

③ 根据 TRIPS 协议第 39 条，当一项信息满足三项条件即为商业秘密：（1）该信息不为公众所知；（2）因为该信息不为公众所知，为其所有者带来了经济利益；（3）该信息所有者为了保持其秘密性需要做出一定的努力。

④ 参见梁志文：《论算法排他权：破除算法偏见的路径选择》，《政治与法律》2020年第8期；狄晓斐：《人工智能算法可专利性探析——从知识生产角度区分抽象概念与具体应用》，《知识产权》2020年第6期。

⑤ 参见汪庆华：《人工智能的法律规制路径：一个框架性讨论》，《现代法学》2019年第2期；孙建丽：《算法自动化决策风险的法律规制研究》，《法治研究》2019年第4期。

⑥ 参见丁晓东：《论个人信息法律保护的思想渊源与基本原理——基于“公平信息实践”的分析》，《现代法学》2019年第3期。

我国的相关法律、法规和行业标准也日益注重这一方式。我国《民法典》的人格权编除规定隐私权益受保护外，还规定了个人对自身个人信息的查阅、复制、更正等权利。^①《网络安全法》的一些条文和一些行业标准规定了系列个人信息保护制度。^②《个人信息保护法》对个人信息进行综合性的立法与保护。

1. 可行性

在可行性层面，个人数据赋权并不一定能很好回应算法决策或算法辅助决策带来的问题。

首先，个人常常无法在数据收集时作出合理判断。在个人数据保护的现有法律制度中，一项重要的制度设计就是所谓的“告知—选择（notice-choice）”框架：企业等相关主体在网站或产品上告知相关群体其隐私政策，再由公民个体选择是否同意。但相关研究表明，个体对于隐私风险的认知往往非常有限，^③而且企业的网站和隐私政策非常复杂和冗长，^④一般读者很难理解，^⑤要阅读所有网站的隐私政策更是要花费海量时间。^⑥因此，虽然现有各种数据隐私的立法大都赋予个体知情选择权，但人们往往很少或几乎不阅读相关的隐私公告。^⑦

其次，在数据处理时，个人也很难有效行使反对权。欧盟《一般数据保护条例》

① 参见《民法典》第 1035—1039 条。对我国民法典的个人信息保护制度的解读，参见张新宝：《〈民法总则〉个人信息保护条文研究》，《中外法学》2019 年第 1 期；王成：《个人信息民法保护的 mode 选择》，《中国社会科学》2019 年第 6 期。

② 参见《网络安全法》第 41—45 条；全国信息安全标准化技术委员会：《信息安全技术个人信息安全规范》，2018 年 1 月 24 日，<https://www.tc260.org.cn/front/postDetail.html?id=20180124211617>，2018 年 7 月 24 日。

③ Richard H. Thaler and Cass R. Sunstein, *Nudge: Improving Decisions About Health, Wealth, and Happiness*, New Haven, CT: Yale University Press, 2008, p.9.

④ 以 Facebook 的隐私政策为例，其篇幅甚至长于美国宪法。参见 Bianca Bosker, “Facebook Privacy Policy Explained: It’s Longer than the Constitution,” *Huffington Post*, April 8, 2014, https://www.huffingtonpost.com/2010/05/12/facebook-privacy-policy-s_un__574389.html, Aug. 5, 2018.

⑤ 有的法律注意到这一点，例如《一般数据保护条例》第 7 条规定，隐私政策“应当使用一种容易理解的形式，使用清晰和平白的语言”。但即使这一规定也不可能完全简化相关主体的隐私政策，因为后者本身就是非常复杂的。

⑥ 一项研究估算，一个美国公民，如果要阅读所有访问网站的隐私公告，那么他一年可能需要花费 244 小时。参见 Lorrie Faith Cranor, “Necessary But Not Sufficient: Standardized Mechanisms for Privacy Notice and Choice,” *Journal on Telecommunications & High Technology Law*, vol.10, no.2, 2012, p.274.

⑦ 一项调查发现，在美国的相关互联网使用者中，只有 20% 的人们在“大多数情况下会阅读隐私公告”。参见 Helen Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life*, Stanford, CA: Stanford University Press, 2009, p.105.

赋予个体随时反对“因为直接营销目的而处理个人数据”或完全“依靠自动化处理”而“对数据主体做出具有法律影响或类似严重影响的决策”，但这种赋予个体的权利事实上很难被个体行使。面对算法黑箱，个体很难知晓其个人数据是否被处理，其个人数据何时被处理，被哪个数据处理者处理。

最后，个人很难对侵犯个人数据权利的行为寻求救济。在信息社会，个人数据权利所面临的威胁涉及多个主体，侵权过程难以辨识。例如，个人可能只在某个社交网站上公开部分数据，其他公司可能通过爬虫技术获取此类数据，后者将此类数据部分匿名化后出售给另外一家大数据公司。对于此类侵权行为，个人即使知晓其中存在侵犯隐私权益的行为，也难以向法院提起侵权之诉或向有关主体申诉。无论是个人的诉讼成本、提供证据还是证明因果关系，都可能面临很多难题。^①

2. 可欲性

在可欲性层面，从个人数据赋权的角度回应算法也存在一些需要重新思考的问题。首先，就自主性来说，基于算法的自动决策或辅助决策可能威胁个体的自主性或自主选择，但在其他很多情形，算法也可能帮助人们作出更好的选择，从而更好地实现人的主体性。^② 如果没有算法推荐或辅助决策，人类可能会作出很多错误判断，所谓人的自主性也只是空洞的口号。正如有些学者所说，如果说算法与科技让我们“放弃自主性，那仅仅是此类自主性：做错误决策、和后来发现我们不喜欢的人去差劲餐厅、购买无聊小说、听糟糕音乐、参加付出多余收益活动。比起那个具有错误观念、有限信息和情绪化干扰的自己，其实我们已经做出了更好的选择，因为我们的真正自我已经得到了强大和有效的科技的帮助。”^③

在这个意义上，不能简单认定，基于算法的决策或算法辅助决策不利于人的自主性。如果运用合理，算法反而可能成为人类自主性的助推者，帮助人类个体或集体作出更有效和更好的选择。^④ 因此，不能简单将隐私或个人数据保护视同保护个

① James Q. Whitman, “The Two Western Cultures of Privacy: Dignity Versus Liberty,” *Yale Law Journal*, vol. 113, no. 6, 2004, p. 1204; Lawrence M. Friedman, “Name Robbers: Privacy, Blackmail, and Assorted Matters in Legal History,” *Hofstra Law Review*, vol. 30, no. 4, 2002, p. 1125.

② 以主体性为基础对算法进行规制的思路，参见苏宇：《算法规制的谱系》，《中国法学》2020年第3期；陈姿含：《人工智能算法中的法律主体性危机》，《法律科学》2019年第4期；李文静、栾群：《人工智能时代算法的法律规制：现实、理论与进路》，《福建师范大学学报》2020年第4期。

③ Richard T. Ford, “Save the Robots: Cyber Profiling and Your So-Called Life,” *Stanford Law Review*, vol. 52, no. 5, 2000, pp. 1573-1584.

④ 由外界助推而提高个人有效选择和自主性的理念，可参见 Richard H. Thaler and Cass R. Sunstein, *Nudge: Improving Decisions About Health, Wealth, and Happiness*, pp. 6-8.

人或群体的自主性。

同样，从信息的角度看，个人数据赋权也不能等同于自由。基于算法的个性化推荐可能会让人困于信息茧房，威胁个体和集体的自由。但在没有算法推荐的情形下，个体或集体也可能限于信息茧房。毕竟，信息的传播从来都不是在真空中进行的，在算法的个体性推荐以及算法辅助决策崛起之前，整个社会就存在很多的信息筛选机制，记者、编辑、出版社以及各类传播机构对于信息的采集、编辑与传播存在很多控制。^①而就信息的接收者来说，个体在没有算法介入的情形下也同样偏向于选择自己愿意听到的信息，集体也可能出现所谓的群体极化的情形。因此，不能简单认定基于算法的信息推荐会比没有算法的世界更有利于人的自由。

（三）反算法歧视

从反歧视的角度规制算法，消除算法中可能存在的身份歧视与偏见，也是目前众多研究和报告提出的重要手段。

首先，研究者指出，应当禁止利用算法对属于某些种族、信仰、性别的群体进行区别对待，以防止恶意歧视。大数据的预测功能已为广泛验证，但研究者指出，大数据的预测功能不仅被用于预测疾病、风险与事故，也容易被用于预测人的身份。例如，研究表明，通过对大数据的算法分析，人们的饮食、音乐等各种偏好可以很容易被用来预测个体的种族属性。如果算法的掌控者是一位种族主义者，很可能会利用此种算法进行恶意歧视，例如因对方是黑人而拒绝发放贷款。

其次，研究者也指出，应尽可能消除数据中的身份偏见，以反对非恶意歧视或无意识歧视。根据这种理解，歧视并非都是有意为之，更多的歧视是基于无意识或错误而产生的。在基于数据的运算中，这种情况尤其明显。大数据与算法的结合很可能产生错误与偏见。一种算法模型可能“会发现关联，并根据网上搜索做出不公平与歧视性的数据推论，对产品、银行贷款、医疗保障产生影响”。^②

1. 可行性

在可行性方面，以身份中立为目标的反算法歧视会遭遇若干难题。首先，在算法中禁止运用身份因素，不一定就能实现身份平等。离开了算法决策或辅助决策，在现实社会中这些群体可能仍然遭受差别性对待与歧视，而且此类歧视可能要比利用算法进行决策或辅助决策更为普遍。以美国网约车的算法歧视为例，Uber 的算法曾被很多人批评为隐含歧视，因为 Uber 的算法一度很容易让司机猜测乘客的种族，这导致黑人乘客很难叫到车。但在没有 Uber 算法的情况下，一般的私家车在网络

^① Randall Bezanson, "The Developing Law of Editorial Judgment," *Nebraska Law Review*, vol. 78, no. 4, 1999, p. 756.

^② Steve Lohr, "The Age of Big Data," *The New York Times*, February 11, 2012, <https://wolfweb.unr.edu/homepage/ania/NYTFeb12.pdf>, August 6, 2018.

约车或搭载乘客时更容易拒载黑人等群体。究其原因，美国社会中的歧视问题根深蒂固，既包括有意识的偏见性歧视，也包括广泛存在的无意识歧视。人们的认知过程普遍存在着类型化的心理机制，^① 即使在不存在明确动机的情况下，人的认知也会潜移默化 and 不知不觉地将人群归类，并且在此过程中形成偏见和无意识的歧视。^② 因此，即使相关数据与算法隐含了对某些群体的歧视，也不能简单地认为禁止在算法中运用身份因素可以更好地消除偏见与歧视。没有机器算法决策或机器算法辅助决策，相关群体可能遭受法律与社会规范中隐藏得更为普遍和严重的歧视。^③

其次，期待在算法中实现彻底的身份中立，这本身就是一个不可能实现的目标。现实社会中的种种经验已经验证了这一点，很多“看上去中立、不关注个人的制度性不作为、已经确立的结构和社会政治规范”，尽管看上去是理性的，但综合考虑却“制造和强化了种族隔离和不平等”。^④ 以美国社会中的黑人与白人的关系为例，“黑人更少进入大学，被大学录取的更少毕业。黑人在专业领域、学术领域和联邦政府中更少被代表。”^⑤ 在这样的背景下，当美国社会仍然沿用历史上的规则，此类规则就构成了对黑人的身份性歧视。另一方面，如果改变此类规则，在相关就业招生等政策中对黑人进行优待或行使平权行动（affirmative action），很多人又会认为此类政策存在对白人的逆向歧视。^⑥ 可以预见的是，无论如何设计，一种算法必然会更有利于拥有某些身份的群体，不利于另一些群体。

① Linda Hamilton Krieger, “The Content of Our Categories: A Cognitive Bias Approach to Discrimination and Equal Employment Opportunity,” *Stanford Law Review*, vol. 47, no. 6, 1995, p. 1164.

② Charles R. Lawrence III, “The Id, the Ego, and Equal Protection: Reckoning with Unconscious Racism,” *Stanford Law Review*, vol. 39, no. 2, 1987, p. 317.

③ 在这个意义上，法律可被视为另一种非机器的人为算法，参见 John O. McGinnis and Steven Wasick, “Law’s Algorithm,” *Florida Law Review*, vol. 66, no. 3, 2014, p. 991; 郑戈:《算法的法律与法律的算法》，《中国法律评论》2018年第2期；丁晓东:《算法与歧视？——从美国教育平权案看算法伦理与法律解释》，《中外法学》2018年第3期；蒋舸:《作为算法的法律》，《清华法学》2019年第1期。

④ Erica Frankenberg and Chinh Q. Le, “The Post-Parents Involved Challenge: Confronting Extralegal Obstacles to Integration,” *Ohio State Law Journal*, vol. 69, no. 5, 2008, p. 1016. 对结构性的种族不平等的分析，参见 Richard Thompson Ford, “The Boundaries of Race: Political Geography in Legal Analysis,” *Harvard Law Review*, vol. 107, no. 8, 1994, p. 1843.

⑤ T. Alexander Aleinikoff, “A Case for Race-Consciousness,” *Columbia Law Review*, vol. 91, no. 5, 1991, p. 1060.

⑥ James E. Jones, Jr., “Reverse Discrimination in Employment: Judicial Treatment of Affirmative Action Programs in the United States,” *Howard Law Journal*, vol. 25, no. 2, 1982, pp. 218-223.

2. 可欲性

从可欲性的层面看，以身份中立为目标的反算法歧视首先会遭到信息与统计理论的质疑。反算法歧视的观点认为，很多算法中隐含着歧视，必须通过法律予以矫正，但在有些研究者看来，很多类型的“歧视”仅是一种信息与统计区分，是社会克服信息不对称的必要手段。以劳动力市场中存在的歧视为例，市场中广泛存在的歧视和对身份的要求是信息匮乏条件下的一种理性筛选。^①对前来应聘的个人，雇主不可能在短时间内进行全面考察，只能依赖一些群体特征对个体进行预测。^②

根据群体身份辨别个人能力的行为当然会造成不公平的现象，例如造成对某些群体的刻板印象（stereotype）。^③但从信息与统计的视角看，这种群体区分对市场正常运行与社会规范建构具有一定意义。群体印象是信息不对称条件下无法避免的行为。社会总会尽可能地基于某个群体的平均表现对其进行评价，这种评价虽然可能无法完全公正反映这个群体中个体的情况，但是比起信息不对称产生的抓瞎式筛选所带来的不公，数据和统计至少为相关决策者提供了参考。而且，社会自身也会对不合理的数据与统计进行修正，不断消除数据与算法中可能存在的歧视。^④以身份中立为目标的反算法歧视还可能面临正义理论与伦理价值的质疑。如果反算法歧视的目标仅仅是追求和实现身份性的中立，那么此类反算法歧视只不过接受或默认了现存社会制度的合理性，放弃了利用算法来改善社会不公的机会。

四、算法规制的场景化原理

在算法规制中，算法公开、个人数据赋权、算法反歧视的方式面临困境最根本的原因在于，机械运用这些方式都不符合算法规制的原理。算法可能会因为使用主体、针对对象、所涉问题的不同而有很大差异。一旦场景不同，算法的性质就会非常不同，对其所采用的规制原则也应当不同。机械性地采用任何一种方法，其实都将算法看作相对孤立的规制对象，都没有结合具体场景进行思考。

就算法主体而言，如果算法的使用者是公权力机构，则算法可能成为一种公共决策机制。特别是如果相关算法决策是由代表性不足的某些规制机构所做出，而此类决策又对公民权益具有重大影响时，则此时算法公开就具有更多的必要性，因为

① Edmund S. Phelps, "The Statistical Theory of Racism and Sexism," *The American Economic Review*, vol.62, no.4, 1972, p.659.

② Dennis J. Aigner and Glen G. Cain, "Statistical Theories of Discrimination in Labor Markets," *Industrial and Labor Relations Review*, vol.30, no.2, 1977, p.175.

③ Regents of University of Calif v. Bakke, 438 U. S. 265(1978).

④ Gerald S. Oettinger, "Statistical Discrimination and the Early Career Evolution of the Black-White Wage Gap," *Journal of Labor Economics*, vol.14, no.1, 1996, p.52.

此时算法更具有公共性，更接近于一种正当程序。^① 如果算法的使用者是一般企业，则算法可能成为企业的内部决策程序，此时算法就更类似企业的自主经营权；当算法具有商业价值与保密性特征时，此时算法还可能成为企业的商业秘密。在此类情形中，算法的公开就不应成为一种强制性法律要求。

此外，很多算法的运用者兼具公共性特征与私主体特征。在大型企业与网络平台企业兴起的今天，很多机构的公共性属性与商业性属性已经变得很难区分。对于类似 Google、Amazon、Facebook、阿里、腾讯、百度这样的超级网络平台来说，仍然以纯粹的市场私主体看待它们，已经与实际情况不符。从数据、资源、算法到服务，这些企业已经日益嵌入人们的日常生活，具有准公权力的性质或者公权力的一些属性。^② 对于此类主体，算法既具有公共决策的性质，也具有企业自主决策权的性质。对于此类情形的算法公开，应当进一步结合算法的公共属性而进行判断。例如运用算法配备外卖骑手，由于关涉外卖人员的人身安全风险，此时应当更多以公共属性看待企业的算法，要求企业承担相应的安全保障义务。

就算法针对的人群而言，如果算法针对的是具有高度可识别性的个体，其数据的收集与算法的运用都是以识别特定个体为目标。那么在这种情形下，算法的性质就与个人权利密切相连，从个人数据权利立法的角度规制算法更为合理。例如《一般数据保护条例》第 15 条规定，当“存在自动化的决策”，数据主体有权知晓“在此类情形下，对于相关逻辑、包括此类处理对于数据主体的预期后果的有效信息”。第 21 条规定，当数据控制者或处理者“因为直接营销目的而处理个人数据，数据主体有权随时反对为了此类营销而处理相关个人数据，包括反对和此类直接营销相关的用户画像”。对于高度追踪和针对个体的算法，此类个人数据赋权能够提供个人权益的合理保护。

但在其他情形中，如果数据的收集主要是为了分析某个群体或不能直接识别个体的对象提供服务，则此类情形中的算法和个人数据权利的关系并不密切。在此类情形中，强化个体对于数据的各种权利，可能会影响数据发挥流通性价值与公共性价值。同时，离开了数据的汇集与共享，此类个人数据赋权还可能影响算法本身的有效运行。

如果算法涉及的是纯粹商业化的决策，则此类情形中的算法就更接近于统计区分，算法就更多是一种信息匮乏手段下的信息甄别机制。但在其他很多涉及弱势群体保护的例子中，算法可能演化为可能加剧社会不公的助推器。在此类情形中，理应对算法进行更多的干预，将扶助弱势群体的伦理嵌入算法。

① 将算法视为正当程序的论证，参见陈景辉：《算法的法律性质：言论、商业秘密还是正当程序？》，《比较法研究》2020年第2期。

② 参见齐延平：《论人工智能时代法律场景的变迁》，《法律科学》2018年第4期；刘权：《网络平台的公共性及其实现——以电商平台的法律规制为视角》，《法学研究》2020年第2期。

总之，就算法规制的一般原理而言，算法规制应当建立在场景化思维的基础上。算法并不像一般的有形物或某些无形物，后者具有相对稳定的法律属性，因而适用统一的法律框架。例如对于动产和不动产，法律一般对其适用统一的物权、合同或侵权法的框架。算法并不是一种标准化的物，而是一种人机交互的决策，因此，算法的法律属性会因为具体场景的不同而有所不同，算法法律规制的原理必须建立在场景化的基础上。

以分类场景的原则规制算法，与当前算法规制的前沿研究具有内在一致性。例如海伦·尼森鲍姆（Helen Nissenbaum）发展出场景公正（contextual integrity）理论，将其首先使用在个人数据的收集问题上，^① 提出对个人数据的收集要考虑场景的类型、行为者的身份、数据的类别以及个人数据传输原则等因素，不同的场景具有不同的个人数据保护规则。^② 近年来，一些学者又将场景理论使用在算法规制问题上，认为算法规制也应当结合不同场景设置不同规则。^③

此外，还有一些前沿研究提出负责任的算法（accountable algorithm）、^④ 可信的算法（trustable algorithm）的主张，使得算法决策能够赢得个体或群体的信任。^⑤ 尽管这些主张使用的概念各异，但它们都认为，算法规制不能机械性地采取算法公开、个人数据赋权或反算法歧视的方式，而是应当根据具体场景，综合运用透明性原则、正当程序原则、市场机制与反歧视框架进行判断。算法规制的具体手段应当帮助个体或群体作出更为正当合理的决策。

五、算法规制的制度建构

通过分析传统算法规制方式的困境以及算法规制的场景化原理，可以对算法规制的制度进行建构。通过合理的制度建构，算法公开、个人数据赋权与反算法歧视可以重新发挥其合理的功能。

① Helen Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life*, pp. 140-160.

② Helen Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life*, p. 141.

③ Richard Warner and Robert H. Sloans, "The Ethics of the Algorithm: Autonomous Systems and the Wrapper of Human Control," *Cumberland Law Review*, vol. 48, no. 1, 2017, pp. 37-66; Doaa Abu-Elyounes, "Contextual Fairness: A Legal and Policy Analysis of Algorithmic Fairness," *Journal of Law, Technology & Policy*, vol. 2020, no. 1, 2020, pp. 1-54.

④ Joshua A. Kroll et al., "Accountable Algorithms," pp. 633, 640.

⑤ Deven R. Desai and Joshua A. Kroll, "Trust But Verify: A Guide to Algorithms and the Law," *Harvard Journal of Law & Technology*, vol. 31, no. 1, 2017, p. 1.

（一）算法公开的制度建构

就算法公开而言，算法公开制度首先应当是有意义和有特定指向的决策体系的公开，而非一般性的算法架构或源代码的公开与解释。如果机械地对算法架构或源代码进行公开，或者机械地对算法进行“解释”，那么此种公开或解释不但难以做到，而且即使做到了，相关解释也可能没什么意义，因为此类解释无法给相关主体提供有意义的决策参考。

对于此种有针对性的决策解释与一般性算法解释的区别，研究者曾经在一定程度上涉及。在讨论算法黑箱与算法解释权时，爱德华兹（Lilian Edwards）和维勒（Michael Veale）分析了什么样的解释才是真正有意义的。^① 他们区分了以模型为中心的解释和以主体为中心的解释两种解释方式。所谓以模型为中心的解释，就是对算法的整体进行理解，包括理解算法所设定的信息、用来训练算法的数据组、算法模型的性能指标、算法架构设置的全局逻辑、被处理信息等。而所谓以主体为中心的解释，指的是某些相关算法与数据对主体会带来哪些影响，例如有关数据主体的相关数据的变化会对其决策产生哪些方向性变化，哪些被用于训练的数据记录与数据主体的数据最为相似，数据主体被归纳为某种类别的主要特征是什么，系统对数据主体进行分类的信任区间多大。^② 两位学者指出，应当抛弃前一种算法解释方式，因为它不但可能威胁商业主体的知识产权或商业秘密，也无法为相应的主体提供有意义的解释。相较而言，后一种解释方式不但不会威胁知识产权，也为相关主体提供了有意义的决策参考。^③

其次，算法公开应当区分不同的算法拥有者，对不同主体施加不同责任。对于公权力机构所使用的算法，应当以公开为原则，以非公开为例外，因为公权力机构具有权力垄断性质，而且其所使用的算法可能对不特定的个体产生重大影响。如果基于算法或算法辅助决策的公权力决策体系维持黑箱性质，那么算法就可能带来本文所涉的不公、歧视、偏见等种种问题，甚至带来腐败，而公民个体一旦遭遇到算法不公与腐败，也很难知晓或很难进行救济。^④ 就这一点而言，商业机构

① Lilian Edwards and Michael Veale, “Slave to the Algorithm? Why a ‘Right to an Explanation’ is Probably Not the Remedy You are Looking For,” *Duke Law & Technology Review*, vol. 16, no. 1, 2017, p. 18.

② Lilian Edwards and Michael Veale, “Slave to the Algorithm? Why a ‘Right to an Explanation’ is Probably Not the Remedy You are Looking For,” pp. 55-56, 58.

③ Lilian Edwards and Michael Veale, “Slave to the Algorithm? Why a ‘Right to an Explanation’ is Probably Not the Remedy You are Looking For,” pp. 58-59.

④ 参见 Cary Coglianese, “Regulating by Robot: Administrative Decision Making in the Machine-Learning Era,” *Georgetown Law Journal*, vol. 105, no. 5, 2017; 查云飞:《人工智能时代全自动具体行政行为研究》,《比较法研究》2018年第5期; 马颜昕:《自动化行政方式下的行政处罚:挑战与回应》,《政治与法律》2020年第4期。

所使用的算法规制原则和非商业机构非常不同。对于纯粹商业性的非垄断机构所使用的算法，算法公开一般不应当成为强制要求。除了上文提到的计算算法和知识产权问题之外，另一个原因在于，非垄断的商业机构往往有动力优化自身的算法，算法的合理性可以通过市场竞争的方式改善。^① 例如，消费者发现某家网站的评分机制非常不合理不公正，他们就可能会选择另外评分较为公正合理的同类网站，而该网站也会有很大的动力来改进自身的算法。如果市场能够拥有更好的机制来促进商业性算法优化，那么此类商业性算法就无需过多的法律介入。^②

最后，算法公开应当根据不同情形采取不同范围不同程度的公开。在有些情形下，算法应当尽可能地公开。当公权力主体使用算法进行决策或辅助决策，而公开算法又不存在侵犯知识产权与算计算法问题时，决策主体应当尽量采取普通人可理解的方式公开与解释算法。因为在此种情形下，算法的公开与透明可以构建良好的法律议论方式，^③ 构建可视正义，^④ 从而更好地监督算法与改进算法。

在有些情形下，算法应当部分公开或小范围公开。当算法黑箱问题已经引起相关主体疑虑，而公开算法又存在侵犯知识产权与算计算法问题时，此时可以考虑算法的部分公开或小范围公开。所谓部分公开，指的是有关主体可以对引起疑虑的算法决策体系进行解释和说明，以消除有关主体的疑虑。^⑤ 所谓小范围公开，指的是算法可以对决策者内部或外部的部分人员公开算法。在决策者内部，可以鼓励企业等其他主体设立关于算法的法律与伦理委员会，通过决策主体内部的专业委员会来保证算法的正当性与合理性。^⑥ 在决策者外部，可以通过设置专门机构或同行评议等方

① 参见 Joshua New and Daniel Castro, *How Policymakers Can Foster Algorithmic Accountability*, Center for Data Innovation, May 21, 2018, pp. 17-19, <http://www2.datainnovation.org/2018-algorithmic-accountability.pdf>, July 31, 2018; 张吉豫:《智能社会法律的算法实施及其规制的法理基础——以著作权领域在线内容分享平台的自动侵权检测为例》,《法制与社会发展》2019 年第 6 期。

② 对于平台评分机制的市场意义与治理意义,参见胡凌:《数字社会权力的来源:评分、算法与规范的再生产》,《交大法学》2019 年第 1 期。

③ 参见季卫东:《人工智能时代的法律议论》,《法学研究》2019 年第 6 期。

④ 参见马长山:《司法人工智能的重塑效应及其限度》,《法学研究》2020 年第 4 期。

⑤ 一个范例是对算法的反设事实解释 (unconditional counterfactual explanation), 即对在解释算法时反设事实, 告知相关主体在反设事实情况下会得出什么样的不同结论。反设事实解释不仅避免了对复杂专业的算法的完全公开, 而且便于数据主体理解自动化决策、对决策进行抗辩, 以及改变未来的做法以达到预期目的。参见 Sandra Wachter, Brent Mittelstadt and Chris Russell, "Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR," *Harvard Journal of Law & Technology*, vol. 31, no. 2, 2018, pp. 1-52.

⑥ 参见《互联网巨头应设立伦理委员会》, 2018 年 1 月 12 日, <http://opinion.caixin.com/2018-01-12/101196908.html>, 2018 年 8 月 2 日。

法审查与评估算法。^① 上文提到的外卖算法系统，执法机构可以对美团、饿了么等企业算法进行专家内部审查，检查这些企业的算法是否满足了对劳动者的安全保障义务。

在有些情形下，算法不仅不应公开，而且应当善于通过算法黑箱实现有关社会政策。在一些情形下，信息公开会带来种种问题，利用具有黑箱性质的算法决策体系完成相关任务，反而会产生意想不到的效果。^② 一个例证是利用算法黑箱进行扶贫助学。在当前高校的扶贫助学政策中，很多高校对贫困生评定与资助采取公开评议等方式，但公开评议对贫困生的心理造成伤害，甚至使得很多真正的贫困生不愿申请资助。对此，中国科学技术大学创设了基于算法决策的贫困生隐形资助政策。通过相关算法设计，当某位同学的校园卡在食堂的月消费金额低于某个金额，学校就会自动向卡内打入生活补助。^③ 相比传统的扶贫助学政策，这项具有黑箱性质的算法实际上设计了一种更好的决策体制。

（二）个人数据赋权的制度建构

个人数据赋权制度首先要区分场景与对象，赋权的类型与强度要因场景与对象的不同而不同。如果相关个人数据的收集与使用有利于相关个体或者有利于公共利益，就应当更多允许相关主体收集和处理个人数据。比如，在扶贫助学或者精准扶贫实践中，个人数据的合理收集与使用应当为法律所允许，相关扶贫主体甚至应当积极运用个人数据与算法来实现精准扶贫。再如，在纯粹商业性的活动中，法律也应当允许消费者选择个性化推荐，因为此类推荐可以节省消费者的搜寻成本，有利于消费者福利和有效决策。^④

相反，当个人数据的使用不是为了促进个人或公共利益，而是被用于支配个体，那么此时法律应当严格规制个人数据的收集与处理。比如，某些网络平台不断地向个体推送各类低俗信息，以增强平台的用户黏性，对于此类行为，应当更严格地进行规制，以避免个体陷入算法的信息宰制。再如，有些网络平台利用个性化推荐进行“杀熟”或差异化定价，针对不同个体的需求设置不同价格，对于此类行为，法律应设置更高的个人数据赋权标准。在此情形下，商家的行为已经超越了传统的商

① 参见张恩典：《人工智能算法决策对行政法治的挑战及制度因应》，《行政法学研究》2020年第4期。

② 参见季卫东：《人工智能开发的理念、法律以及政策》，《东方法学》2019年第5期。

③ 参见《暖心高校偷偷给贫困生打钱 中科大用“黑科技”实现隐形资助》，2017年7月13日，<http://www.takefoto.cn/viewnews-1205834.html>，2018年8月2日。

④ 学者还指出，互联网带来的搜寻成本的降低会促进商品价格的降低和价格的差别定价。参见 Pedro Pereira，“Do Lower Search Costs Reduce Prices and Price Dispersion?”，*Information Economics and Policy*, vol.17, no.1, 2005, pp. 61-72.

业行为，构成了对消费者的“一级价格歧视”。面对此类行为，消费者的知情权与选择权已受到很大威胁，面临商家的定向支配的风险。

其次，个人数据赋权应当更注重数据的动态规制与伦理化规制，为算法的有效运转提供可能。在过去几十年，算法本身其实并没有任何突破，算法准确率的提升依赖于数据的海量积累。^①而大数据的本质特征恰是数据的全体性、混杂性和相关性。因此，有的数据隐私法所规定的原则，例如目的限制原则（即对个人数据的处理不应当违反初始目的）、限期储存原则（即对于能够识别数据主体的个人数据，其储存时间不得超过实现其处理目的所必需的时间），并不符合大数据时代算法运行的一般原理。在大数据时代，通过算法跨界处理数据特别是沉淀数据，为消费者提供不同目的的服务，已经非常普遍，^②例如相关的物流类数据可能被用于电商与制造，曾经的餐饮娱乐类的沉淀类数据可能被用于网约车服务。对于企业收集与处理此类个人数据，应当注重此类数据处理是否符合相关伦理与人格保护，而非一般性禁止数据的流通共享。^③

最后，个人数据赋权应当更注重数据与算法的合作治理，而非个人控制。^④正如上文所述，面对大数据时代的算法，个体很难有效行使自身的若干数据权利。在这种背景下，法律应当更多对数据控制者与处理者施加治理责任，而非仅仅依赖个人的数据控制。单纯依赖个人对其数据的控制，个体可能陷入没有时间、兴趣和能力维护自身合法权益的悖论，数据收集者与处理者也没有动力对数据与算法进行有效治理。只有将责任伦理嵌入数据收集、流通与处理的每一个环节，以此倒逼算法治理，通过个人数据实现算法治理才能真正起到实效。

（三）反算法歧视的制度建构

就反算法歧视制度而言，应首先警醒算法中隐藏的歧视、偏见与不公。大数据

-
- ① 对于隐私保护与大数据有效运行之间的张力，参见季卫东：《数据、隐私以及人工智能时代的宪法创新》，《南大法学》2020年第1期。
- ② 正如舍恩伯格所言：“大数据的价值不再单纯来源于它的基本用途，而更多源于它的二次利用”，维克托·迈尔-舍恩伯格、肯尼思·库克耶：《大数据时代：生活、工作与思维的大变革》，第197页。
- ③ 对于人格权保护与数据流通共享的分析，参见王利明：《数据共享与个人信息保护》，《现代法学》2019年第1期；程啸：《论大数据时代的个人数据权利》，《中国社会科学》2018年第3期。
- ④ 参见 Margot Kaminski and Binary Governance, “Binary Governance: Lessons from the GDPR’s Approach to Algorithmic Accountability,” *South California Law Review*, vol. 92, no.6, 2019, p.1529; 高富平：《个人信息保护：从个人控制到社会控制》，《法学研究》2018年第3期；梅夏英：《数据的法律属性及其民法定位》，《中国社会科学》2016年第9期；丁晓东：《个人信息私法保护的困境与出路》，《法学研究》2018年第6期。

时代的算法早已不是数学意义上的算法或实验室中的算法，对基于算法的决策或辅助决策，不能盲目偏信其科学性或价值中立性。^①相反，必须时刻警惕算法中可能存在的歧视与不公，警惕某些主体利用算法来实现其不正当的目的。^②对于算法中可能存在的此类恶意偏见与不公，应当结合算法公开与个人数据赋权消除此类歧视与偏见。一方面，通过合理的算法公开，公众或内部专家可以对算法本身与算法决策结果进行监督，防止算法黑箱中隐藏的歧视与机器自我学习带来的歧视。^③另一方面，通过合理的个人数据赋权，法律可以防止相关主体对用户的个人数据进行数据处理和用户画像，避免对个体进行歧视性对待。^④

其次，反算法歧视应当超越身份中立，根据身份与弱势群体之间的关系规制和利用算法中的身份因素。反歧视的最终目的是实现更为公正合理的社会目标，而不是形式上对各类身份平等对待。因此，在算法中进行反歧视，应当深刻理解哪些群体在历史上与社会中遭受不公平对待，然后在相关算法中应用平权行动和差别性影响等方式扶助这些群体。例如就差别性影响来说，应尽可能避免某项算法对某些身份的群体造成更多伤害，当某项算法对某类弱势身份性群体造成显著负面影响，就应对此类算法进行反歧视审查。^⑤就算法平权行动来说，算法的参数设置应考虑对具有某些身份的群体倾斜，保证此类群体有更多的平等机会。^⑥

在中国语境下，反算法歧视尤其应当注意这一点。对于反歧视的目标与价值导

① 参见 Batya Friedman and Helen Nissenbaum, "Bias in Computer Systems," *ACM Transactions on Information Systems*, vol. 14, no. 3, 1996, p. 330.

② 参见郑智航、徐昭曦：《大数据时代算法歧视的法律规制与司法审查——以美国法律实践为例》，《比较法研究》2019年第4期；刘友华：《算法偏见及其规制路径研究》，《法学杂志》2019年第6期；徐琳：《人工智能推算技术中的平等权问题之探讨》，《法学评论》2019年第3期。

③ 在算法导致的歧视中，有的歧视可能人为有意嵌入算法，例如有的搜索引擎和评级标准的调整。参见 Eric Goldman, "Search Engine Bias and the Demise of Search Engine Utopianism," *Yale Journal of Law & Technology*, vol. 8, no. 1, 2006, pp. 188-200. 有的歧视可能是无意的，例如机器学习中出现歧视，参见 Nicholas Diakopoulos, "Algorithmic Accountability: Journalistic Investigation of Computational Power Structures," *Digital Journalism*, vol. 3, no. 3, 2015, pp. 398-415.

④ 用户画像是产生算法歧视的重要途径，参见 Tal Z. Zarsky, "Transparent Predictions," *University of Illinois Law Review*, vol. 2013, no. 4, 2013, p. 1503; Mireille Hildebrandt and Bert-Jaap Koops, "The Challenges of Ambient Law and Legal Protection in the Profiling Era," *The Modern Law Review*, vol. 73, no. 3, 2010, pp. 428-460.

⑤ Solon Barocas and Andrew D. Selbst, "Big Data's Disparate Impact," *California Law Review*, vol. 104, no. 3, 2016, p. 671.

⑥ Anupam Chander, "The Racist Algorithm?," *Michigan Law Review*, vol. 115, no. 6, 2017, pp. 1035-1045.

向，中国和美国等西方国家有区别。在美国，反歧视法律制度主要是针对身份性的歧视与不公，特别是针对黑人的种族歧视。从国情看，这种制度设计有其合理之处，因为在美国，从建国后到南北战争废除奴隶制，到布朗案废除种族隔离，种族问题一直是美国社会最大的不平等。如今，种族不平等的现状并未得到根本性改变。在美国的大部分地区，黑人区依然和白人区在事实上隔离，黑人在社会各个方面都处于落后的状态。^① 但在中国，身份并非导致歧视的直接原因，很多歧视其实源自市场的区分行为，^② 或者是市场以某些不太合理的个人特征提高企业的竞争力。^③ 在这样的背景下，如果反算法歧视过多关注身份性的歧视与不公，而不注重对弱势阶层的保护，那就只关注到次要矛盾，忽略了主要矛盾。^④ 要使反算法歧视在中国发挥更重要的作用，应当将反算法歧视与精准扶贫等扶助弱势群体结合起来。在人机交互决策中，应当尽可能将扶助弱势群体的伦理嵌入各类算法决策机制。

结 论

在大数据时代，算法对于人类生活的影响已经越来越明显，面对越来越多的自动化决策或辅助化决策，人工智能法学与未来法治领域的研究者越来越关注由算法崛起所带来的挑战，并试图从算法公开、个人数据赋权、反算法歧视等方式来规制算法。本文指出，现代社会算法的本质是一种人机交互决策，应当戳穿算法的面纱与价值中立性，对算法进行规制。

在界定算法定义与阐述其可规制性的基础上，本文对几种传统算法规制方式进行了反思。无论是算法公开、个人数据赋权还是反算法歧视，机械地采取任何一种方式可能都面临可行性与可欲性的难题。从深层原理看，机械地适用几种算法规制的方式之所以存在问题，主要是它们没有采取场景化的规制思路，仍然将算法视为孤立不变的对象。但事实上，作为人机交互决策，算法可能因为运用算法主体的不同、算法针对对象的不同以及算法所涉及问题的不同而具有非常不同的性质。^⑤ 因

① Owen M. Fiss, *A Way Out: America's Ghettos and the Legacy of Racism*, Princeton, NJ: Princeton University Press, 2003, p. 3.

② 例如当前就业市场中普遍存在的学历要求，用人单位之所以对学历进行要求，很大程度上是为了满足职业择优录取的需求，从而提高企业的效率。

③ 例如很多服务型行业对身高、长相的歧视，在很大程度上是因为员工的相貌形象虽然与某项具体工作没有直接相关的关系，但可通过这种手段取悦和招揽顾客，或者提高企业形象。这种形式是否合理，仍然值得探讨。

④ 没有此类平等伦理的介入，算法可能加剧社会的不平等。参见於兴中：《算法社会与人的秉性》，《中国法律评论》2018年第2期。

⑤ 以“过程—结果”对人工智能进行双重规制，参见马长山：《人工智能的社会风险及其法律规制》，《法律科学》2018年第6期。

此，算法规制应当采取场景化的规制方式，根据不同场景类型而对算法采取不同的规制方式。

从场景化规制的原理出发，可以对算法规制进行进一步的制度建构。对于算法公开，算法公开应当是有意义和有特定指向的决策体系的公开，而非一般性的算法架构或源代码的公开与解释；算法公开应对公权力使用的算法与市场主体的算法施加不同责任；算法应当根据不同情况而决定完全公开、小范围公开或不公开。对于个人数据赋权，个人数据赋权的程度应当根据是否有利于个人或公共利益而进行不同程度保护；个人数据赋权应当注重数据的动态化与伦理化规制，而非数据的形式主义规制。对于反算法歧视，应当警惕算法中的歧视与偏见，但应当超越身份中立，根据具体场景在相关算法中合理地运用身份性因素；在中国，反算法歧视更应注重弱势群体保护，注重身份因素与弱势群体保护之间的关联。

总之，算法作为一种新型的人机交互决策，常常被不同的主体运用、内嵌于不同的场景和处理不同的问题。对于人工智能与未来法治研究而言，应当准确把握算法规制的场景化特征与原理，根据不同场景与情形对算法进行规制，以实现可信赖与负责任的算法决策。^① 无论是算法公开、个人数据赋权还是反算法歧视，算法规制的具体制度建构都必须以此为指引，超越形式主义的法律规制。唯此，算法才能避免成为异化的决策体系，^② 算法才有可能真正造福社会。

〔责任编辑：刘 鹏〕

① Joshua A. Kroll et al. , “Accountable Algorithms ,” pp. 695-705.

② 参见季卫东：《人工智能时代的司法权之变》，《东方法学》2018年第1期；张欣：《算法解释权与算法治理路径研究》，《中外法学》2019年第6期；周辉：《算法权力及其规制》，《法制与社会发展》2019年第6期；张凌寒：《算法权力的兴起、异化及法律规制》，《法商研究》2019年第4期。

into its own opposites and develops into a new external alien force. As the basic technical support of the entire society, intelligent technology entails the overt or covert domination of human beings, who are becoming the “vassals” and “slaves” of this high-speed intelligent social system. Various intelligent systems are constantly replacing human work, so that the “digital poor” gradually lose the opportunities and values offered by labor and hence are excluded by the global economic and social system, rendering their existence empty and absurd. The rapid development of intelligent robots has blurred the boundary between humans and machines and had a strong impact on the nature of man and his position as a conscious agent, making “What is man?” and the human-machine relationship prominent issues for our times, challenging the commonplaces of philosophy. We must face up to the existing or imminent risk of alienation, expand our theoretical horizons, innovate theories of alienation in the era of intelligence, take constructive action in terms of the construction of an ideal society and the evolution of man himself, build an ecological system for the joint evolution and growth of human beings and intelligent machines, and achieve liberty of man and the all-round and free development.

(8) On the Legal Regulation of Algorithms

Ding Xiaodong • 138 •

The rise of algorithms poses challenges to their legal regulation as they may challenge people’s right to know and to have individual privacy and freedom and equal protection. As algorithmic decision-making mechanisms involving human-computer interaction, algorithms are not value-neutral and can be regulated. Algorithm disclosure, personal data empowerment, and anti-algorithmic discrimination are traditional methods of regulating algorithms, but the mechanical application of these methods encounters difficulties with feasibility and desirability. Algorithm disclosure faces problems such as technical non-feasibility, meaningless disclosure, user calculations and infringement of intellectual property rights; personal data empowerment faces the problems of individuals being unable to exercise data rights and of excessive personal data empowerment which impedes the effective operation of big data and algorithms; and anti-algorithmic discrimination faces the problems of non-machine algorithmic discrimination, the fact that one’s status cannot be completely neutral, the difficulty in achieving social equality, etc. The fundamental reason for the dilemma of traditional pathways for the regulation of algorithms lies in ignoring their context. Depending on their subjects, objects, and the different problems they involve, algorithms’ properties may vary. Therefore, algorithm regulation should adopt a scenario-based regulatory path and should employ varying regulatory methods depending on the type of scenario, with the aim of achieving responsible algorithms. The principle of scenario-based algorithm regulation can guide construction of specific regulatory systems such as algorithm disclosure, data empowerment and anti-algorithmic discrimination.