

自动化行政算法决策的特殊风险与规制路径

徐 芳

天津大学法学院, 天津

收稿日期: 2023年6月14日; 录用日期: 2023年6月28日; 发布日期: 2023年9月7日

摘 要

自动化算法决策的应用给行政公正原则、正当程序原则、权力与责任相一致原则带来了巨大挑战。这些挑战表现为行政决策违法或不当、决策违反法定程序、行政责任主体不明、行政机关能力弱化等风险。客观层面上, 数据流转偏差、算法不透明与技术权力的侵蚀促成此类特殊风险; 主观层面上, 行政主体对算法的依赖扩张了风险的不可控程度。为有效控制此类新型风险、使算法行政决策更好地造福人类, 可以从地位厘定、义务强化、程序控制等角度切入, 继续坚持行政机关在算法决策中的第一责任人地位、坚守算法的工具主义立场, 场景化、类别化规制算法决策应用程序, 在强化相关主体披露义务的同时, 创新式贯彻正当程序控制, 完善相对人的权利保障与救济体系。

关键词

自动化行政, 算法决策, 行政正当程序, 算法透明化, 行政自由裁量

Specific Risks and Regulatory Paths of Algorithmic Decision-Making in Automated Administration

Fang Xu

Law School of Tianjin University, Tianjin

Received: Jun. 14th, 2023; accepted: Jun. 28th, 2023; published: Sep. 7th, 2023

Abstract

The application of automated algorithm decision-making has brought huge challenges to the principle of administrative justice, due process, and the principle of consistency of power and responsibility, including illegal or improper administrative decision-making, violation of legal proce-

dures, unclear administrative responsibility, and weakening of the ability of administrative agencies. On the objective respect, data flow deviation, algorithm opacity and the erosion of technical power contribute to such special risks; on the subjective ones, the administrative body's dependence on algorithms expands the uncontrollability of risks. In order to effectively control this new type of risk and make algorithmic administrative decision-making better benefit mankind, we can cut from the perspectives of status determination, obligation enhancement, and procedural control, and continue to adhere to the position of the administrative organ as the first responsible person in algorithmic decision-making and stick to the algorithm instrumentalism, and scenario-based and categorized regulatory algorithm decision-making applications, while strengthening the disclosure obligations of relevant subjects, innovatively implement due process control, and improve the rights protection and relief system of the counterparty.

Keywords

Automated Administration, Algorithm Decision-Making, Administrative Due Process, Algorithm Transparency, Administrative Discretion

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. 引言

人类已经进入智能时代，数据、算法、人工智能正以令人难以想象的广度和深度飞速改变现代人类生活，正如学者所言，传统的法律世界正经历着数据、算法、人工智能等信息技术发起的“破窗性挑战”和“创造性破坏”，法学理论、规范制度和司法实践都需进行升级和转型[1]。行政执法领域并非存在于真空，自然也面临变革。

最典型的例子莫过于新冠疫情爆发时，政府为高效防控疫情，迅速开发并向全国推广个人健康码，在全国范围内进行数据收集与算法决策工作，对每个人的疫情风险进行等级化评估；除此之外，行政裁量基准自动化、行政审批流程智能化、车辆违规处理自动化等执法领域的变革也都在表明自动化行政算法决策对传统行政执法领域的渗透与变革；对摇号购房、秒批落户、线上高校贫困生审批等词汇，我们也耳熟能详，自动化行政算法决策与我们的生活密切相关。

然而，目前看来，自动化行政算法决策并非完全排除人工干预，在很多执法领域，自动化行政仍然是人工办公辅助工具，有学者按照行政决策中自动化程度的不同，将自动化行政划分为无裁量能力的完全自动化行政、有裁量能力的完全自动化行政和部分自动化行政三种类别，并对高度自动化行政下的算法决策表示担忧[2]。在提高行政工作效率、节省行政资源等方面，算法决策系统确实有无可比拟的优势，但是它也带来了违反公正原则、违背正当程序、侵蚀行政权力、造成责任主体不明等风险。因此，梳理算法决策带来的潜在的或已经显现的风险，探讨风险成因，并据此提出控制风险的措施，具有重要意义。

本文在此尝试探索三个问题：一是自动化算法决策在行政领域有何特殊风险；二是分析自动化算法决策风险背后的原因；三是提出化解此类特殊风险的措施。

2. 自动化行政算法决策的特殊风险

与传统行政执法风险不同，自动化算法决策减少了行政主体自由裁量的权力空间，改变了行政程序

的运行顺序，削弱了行政机关的权责属性。具体而言，大数据时代下，自动化行政算法决策会带来以下新风险。

2.1. 执法违反公正原则

广义上的行政公正原则包含两方面，一是执法合法，行政决策不违背既有法律规定。二是执法合理，即执法的过程和结果很好地考虑了个案因素。但应用自动化算法决策可能在以下方面违背行政公正原则：

2.1.1. 行政决策错误

算法技术运用于行政决策可能会导致行政机关作出不合法决定。比如，广州云从科技公司研发了一款人工智能系统，它能辅助公安机关根据计算一个人的脸部特征和行为举止，来判断检测对象是否具有社会威胁性，进而提示公安机关采取先发制人的行动，阻止犯罪的发生[3]，但这一系统受到媒体广泛批评。再比如，有算法按照以往违法人员所形成的各项数据，推断哪些人群具有违法犯罪可能，并将其列入高危名单中，以提示有关部门禁止其登机，从法理上来说，对一个并未违法而只是有违法风险的人作出一定限制措施是违背法律正义价值的，被采取此种不利措施的人员此前也并未收到任何通知，也无对此申辩的机会和权利。现实中甚至有因数据或系统偏差而造成飞行员被错误列入禁飞名单的现象[4]。再比如，行政机关根据人工智能技术测算道路维修基金的分配，其结果反而是位于富人区的更平整完善的道路设施，比位于贫民区的坑洼较多、较破碎的道路更需要修缮，原因在于富人区居民更愿意，也更有能力安装政府设计的、根据手机主人走路时的震动状态判断道路状况的 APP [5]。

2.1.2. 行政决策不当

如果赋予行政主体过多的自由裁量权很容易导致行政决策不公正，但过于狭窄的裁量权限就意味着行政主体作出决策时无法充分考虑社会公益、善良风俗、个案正义、公众情感等个别化因素，最终导致行政决策违背公正原则。可以说，公正目标的实现离不开一定范围内的“人”的自由裁量。但算法决策或多或少地缩减了行政机关工作人员的裁量空间。依靠指令，原本的人为裁量成了机械化的规则制定，裁量中须考量的相关因素可以通过系统事先的设定输入，不需要考量的相关因素则事先被剔除。这个复杂的微指令目录可以针对行为的小实例应用于更大的行为动作网络，实现更广泛的目标[6]。这固然保证了行政决策的精确性，但在合理性方面却受到怀疑。

也有学者从法经济学的角度指出了缺乏人为裁量空间的自动化行政处罚的不当之处。按照法律实施效果与经济惩罚措施的一般关系，损害除以违法行为被发现的概率就等于最佳的行政处罚额度。即，当违法行为被发现的可能是四分之一时，那罚款最好设置成损害成本的四倍[7]。也就是说，电子摄像头使得同一行为人的违法行为被发现的可能大大提高，几乎近似于 100%，那传统行政执法下的罚款额度至少应降低四分之三¹。智能技术使得违法行为被发现的概率大大提升，技术判断最优不等于心理价值最优，每次违法行为的处罚额度仍依据传统确定，其合理性是有待商榷的。

2.2. 执法违反正当程序原则

行政程序要求行政主体依照法定方式、步骤和顺序行使行政权力，行政程序是公民控制国家公权力、保障自身合法权益的直接工具和重要武器。尽管在简易程序中，行政程序有所缺省，但随着智能技术的深入发展，传统的正式的行政决策也可由算法完成。有学者根据作为算法决策基础的行为的发生时间，将算法分为反应型算法和预测型算法两种。反应型算法的特点是决策的极速性，在完成相应的数据输入

¹在“杜宝良”案中，灵敏的“电子眼”使得违法行为被发现的概率大大提升，因而也造成了普罗大众心理上难以接受的过重的行政处罚款额。2005年5月23日，杜宝良偶然查询得知其于2004年7月20日至2005年5月23日在驾驶小货车运菜时，在每天必经的北京市西城区真武庙头条西口被“电子眼”拍下闯禁行105次，被罚款10,500元。

后,几乎无需等待即可得出决策结果,传统的行政调查、告知与说明、陈述与申辩等环节难以被嵌入数据处理过程中。预测型算法则根据输入的过往数据进行概率推测,尽可能地预测风险,提醒行政机关做好准备,进行更高效的资源配置[4]。革命式创新的算法决策将会给传统的行政程序运行带来以下挑战:

2.2.1. 调查程序难以实现

一般来说,“除当场做出行政行为的案件外,调查是行政程序的必经程序[8]”。有关法律规定在行政调查阶段设置了诸如表明身份、告知权利、详尽说明、制作笔录并签字等法律义务。但在自动化算法决策下,行政机关工作人员应当如何表明身份、以怎样的方式才能合法恰当地告知相对人权利和义务?除此之外,算法程序是按照同一指令系统作出对相对人权利义务具有实质性影响的决策,这种机械的处理方式如何能将无法数字化的情形,比如当地风俗、情感因素等纳入算法规则系统。行政调查阶段面临的个案情况复杂多样,高度自动化的算法决策系统对传统的调查程序的运作提出了挑战。

2.2.2. 陈述、申辩环节缺乏

在反应型算法中,所有程序均可在毫秒甚至微秒内完成,几乎可以媲美人类思考的速度,除非特意加入人工审核环节。因此,行政机关的告知义务被省略,相对人的陈述、申辩权利也难以保障。而预测型算法的功能在于对未发生的行为进行概率预测,将来具体行政行为是否发生并非确定,只是基于一种可能性而被用于行政资源的合理化调配中。既然并未发生对外部相对人权利义务产生影响的具体行政行为,同时又难以被归类为抽象行政行为,那就谈不上对其进行程序控制,这很可能导致预测型算法作出行政决策前没有任何行政相对人及其他利益相关方的参与,行政相对人的陈述权、申辩权自然也难以保障。比如在江苏省,税务机关利用算法决策系统预测哪类人群税务风险较高并对其采取相应措施,使得该年内稽查次数大大减少,大幅度节省了行政资源[9]。不难看出,在某些行政执法的具体领域中,预测性算法已经发挥了核心作用,成为行政机关实施具体行政行为的主要依据,但同时行政机关没有完全履行信息公开义务,并确保相对人的知情权和申辩权得以实现。

2.2.3. 请求说明理由的权利无法保障

大数据时代下,各种人事物被高度解析,庞大的个人信息集被机器化为无数的“0”和“1”,输入算法系统并被特定的算法规则分类、计算与解析,信息透明化成为当今时代一个重要特征,然而“算法黑箱”论认为,算法规则的专业性以及不同利益集团的需求使得算法难以被普通公众认知。具体到行政领域,面对算法决策,行政相对人有权请求行政机关对具体行政行为的理由作出说明,但行政机关却可能以涉及国家安全、第三人利益等理由拒绝,即便行政机关同意作出说明,那何种程度、何种范围的说明才合法且适当?以何种方式进行说明才能让相对人充分理解算法规则?更进一步地,在自动化算法决策作出前,行政机关是否有提前公示的义务?如果有,公示的范围、公示的对象应当如何确定?

2.3. 责任主体不明与行政主体能力弱化

随着自动化技术的深入发展,在行政执法领域中,行政主体责任的独立性、行政主体执法的主动性受到质疑。高度自动化行政下的算法决策背后是技术公司在借助代码进行行政裁决吗?行政主体执法主动性和自由裁量权力会被侵蚀吗?

2.3.1. 责任主体不明

随着自动化行政的深入发展,行政决策的作出过程不再是行政机关单方面的权力运作,其权力被技术公司不断侵蚀,由此可能引发人们对“权力与责任相匹配”这一古老谚语的怀疑。算法规则的设计与算法系统的开发被技术公司垄断,而技术公司能以商业秘密为由拒绝对政府和公众公开其规则处理系统

与数据处理流程。一般情况下，技术公司可以与行政机关签订行政合同，以取得对算法系统的开发与设计资格，但行政机关缺乏相应的专业技术及专业人员，换句话说，即便技术公司向政府公开相关信息，政府充分理解算法也需大量成本，如在美国迈阿密，根据一份行政合同，一家照明公司通过在公路上安装带有摄像头和车辆牌照读取器的照明灯杆帮助市警局处理车辆违法信息，该技术公司当然享有这些数据的使用权限，但市政府却陷入对算法的“无知之幕”中[10]。从风险控制的角度来看，只有主体对风险有充分的评估和认知，方可由其承担。因此这种“算力即权力”的新型权力形态引发了人们的思考：在行政权力不断被技术公司侵蚀的背景下，单独由行政机关担当行政责任的主体是否具有正当性？如果不是，那么技术公司应该以什么比例、以何种方式承担行政责任？更进一步地，当算法决策系统进化成可以与人类媲美的“强人工智能”，那在行政领域，人工智能对其作出的决策是否可以独立承担行政责任？目前人工智能的法律地位仍不明确[5]。

2.3.2. 行政主体能力弱化

算法决策的深度应用使得人们对行政机关的治理能力产生了质疑。在人类的认知中，很容易产生一种“自动化偏见”，即因为算法程序处理更迅速，详实且全面的数据看起来也更可靠，人们在头脑中更容易将其当作一种纠错程序，用以弥补人类思考的不足，这种偏见本质上源于近现代兴起的“科学崇拜”[11]。普罗大众的这种偏见若被操控，就很容易驱动舆论对传统行政执法方式表达不满甚至反对。对行政机关工作人员来说，这种偏见很可能使他们越来越谋求便捷的算法决策以一劳永逸地解决行政问题，久而久之会弱化行政主体的能动性和治理能力。比如在江苏省某些地区已经开始使用算法决策程序对传统的行政案件的量罚标准进行智能化设计，以尽可能把行政行为的实施过程全部纳入算法的监控之中[12]。我们在关注其便捷性的同时，不能忽视科技这把双刃剑带来的负面影响。

3. 自动化行政算法决策风险成因分析

剖析自动化行政算法决策特殊风险的成因，可以从两方面入手，一是算法系统本身的特征，二是行政主体自身的发展趋势。数据、算法与算力是自动化决策系统的三大基石，数据提供燃料，算法提供运行规则，算力则起到类似发动机的作用，助推算法系统对信息的高速处理。然而，数据流转并非全程毫厘无差，任何细微的信息偏差都可能导致一个令人难以接受的结果；同时，算法黑箱使得算法规则缺乏外部监督，这增加了算法决策错误或不当的风险；除此之外，行政机关由于目前未能配备足够的专业技术设备及专业技术人员，在算力上难以与在特定科技领域深耕多年的技术公司抗衡，而这也给行政机关工作人员带来一种技术上的诱惑，使得他们或多或少地依赖技术公司提供的便捷的社会治理工具。

3.1. 数据流转偏差

科技不是上帝，算法也并非万能。在数据输入——算法处理——数据输出这一流程中，数据与算法可能会产生多种问题，既包括技术方面的，也包括伦理方面的。概括起来，自动化行政算法决策可能会产生以下五个方面的问题：

3.1.1. 原始数据瑕疵

《个人信息保护法》等法律确定了知情同意规制在个人信息处理领域的基础地位，受限于该规则，原始数据收集不能保证全面且典型。另外，基于国家安全、商业秘密、第三人同意等正当理由的考虑，原始数据本身可能存在缺失，比如在重大案件中，政府官员为防止模仿犯罪的发生，可能会拒绝公开一些至关重要的案件细节与证据，而这又使得民众对政府的不信任加重，不愿意透露自己的基因信息以方便政府办案。除此之外，即便信息样本足够全面、算法系统经过测试达到了使用标准，数据输入也可能

因时空、人为因素等出现错误，比如数据输入者的记忆错误，数据年代久远存在多种版本等。

3.1.2. 收集方式不当

考虑到收集数据的时间、人工、金钱等成本，数据收集手段很难保证能充分满足算法计算所需的数据量。而且算法设计者、制造者的设计疏忽和思维限制，也会间接导致数据收集手段不当，比如填写页面不适老，因此一些老人相关的数据样本可能会缺失，再比如数据收集需要个人配备智能手机、平板或电脑等电子设备，这就没有考虑到欠发达地区电子设备并没有普及这一现实，因此才会出现上文提到的富人区比贫民区更需要维护道路设施的荒唐结果。

3.1.3. 信息与结果不具有因果关系

即便收集到的信息真实，收集手段多样且适配算法计算需求，但数据本身可能与算法决策结果无关。比如电子眼在收集相关数据时，并未全程、全面记录整个过程，导致根据片面的画面做出相关行政处罚决定。再比如判断高校一份贫困生资格申请是否可以通过，有无苹果手机成了一条根本判断基准，但实际上，苹果系列手机并非全部都属于高奢产品，且即便手机价格较高，也不一定表明该生不符合贫困生申请条件，还需要考虑学生月均消费水平、家庭资产状况等因素。

3.1.4. 算法系统自身缺陷

除了数据流转偏差、收集手段不当、数据与结果无因果关系外，算法系统自身也会存在缺陷。比如在李健雄诉广东省交通运输厅政府信息公开一案中，被告广东省交通运输厅辩称：原告申请政府信息公开是通过广东省人民政府公众网络系统，即省政府政务外网(以下简称省外网)，而非被告的内部局域网(以下简称厅内网)；由于被告的厅内网与互联网、省外网有物理隔离，互联网、省外网数据都无法直接进入厅内网处理，需通过网闸以数据“摆渡”方式接入厅内网办理，因此被告工作人员未能立即发现原告在广东省人民政府公众网络系统中提交的申请，致使被告未能及时受理申请²。本案中，行政机关无法及时做出行政行为就是因为自动化行政系统本身存在不足。

3.1.5. 算法偏见

即便前述问题可以通过技术手段解决，对算法偏见的修正也很可能是白费功夫。已有实验证明，算法系统中的偏见主要来源于人类社会，而算法系统犹如一个放大镜，会把人性中的偏见、歧视与不公之处明显放大[13]。算法决策系统由人类编程，在技术人员将代码一行行地敲入电脑时，一方面，社会习惯或制度的偏见被工程师们无意识地写入，另一方面，法律规定的空白又为这些偏见与技术相融合提供了充足的空间。

3.2. 算法不透明

正如相关从业者所言，算法的作用只是帮助人类更快、更精准地完成目标，但并不打算告诉人类这是为什么[14]。在这样的背景下，算法被神秘化了，但人类真的无法认知算法吗？剖析算法不透明的原因，可以发现，算法黑箱的“盖子”并非完全不能打开。具体而言，算法不透明包括以下三种情况：

3.2.1. 因人为而不透明

算法如何设计、数据如何收集和输入、算法如何处理相应数据，技术公司以商业秘密为借口不愿公开，行政机关会基于国家安全的考虑不愿公开。2020年健康码实施以后，各地在具体信息搜集程序上差异明显，也存在统一规范等方面的问题。哪些个人数据被算法系统收集？不同种类的数据被以何种方式处理？风险等级的评估标准如何确定？个人申诉后算法系统如何审核和处理？这些由行政机关设定的标

²李健雄诉广东省交通运输厅政府信息公开案，《最高人民法院公报》2014年第8期。

准并未被公开。这种人为不透明的现象在世界范围内并不鲜见。

3.2.2. 因专业而不透明

不可否认的是，代码编写、算法规则设计等知识当前仍是远离一般公众的高端专家技能。面对公众的技术无知，即便基于强制性披露要求公众可以在一定时期一定范围内随时查询到算法内部指令系统和运行原理，但面对天书一般的代码内容，普通公众很难清楚理解，更不用提具体参与到算法设计、编写、运行，对其进行程序控制，并表达自身权益的过程中去了。也就是说，公众的认知能力是阻碍算法透明化的一大障壁。

3.2.3. 因认知而不透明

Alpha GO 战胜世界围棋冠军李世石一事向世界传递了一个消息：人工智能可以进化出人类无法匹敌的能力。应当承认，部分机器学习逻辑是人类目前无法认知的。比如面对飞机突然改变飞行高度，即便是专业素质过硬、经验丰富的老飞行员，也会无法理解其中原因[15]。但这只是说明了算法并非完全黑箱化，且在可以预见的未来内，不可知的部分也会随着技术的进步与人类智慧的发展而变得可知。但在现实中，第二种不透明与第三种不透明常常混合在一起，并被别有用心之人与无知之人夸大，使得算法愈加神秘化，仿佛永远不可被认知。

3.3. 行政主体的依赖与技术权力的侵蚀

智能技术对现代生活的渗透是方方面面的，行政执法领域自然也不例外。面对技术公司带来的提高效率、节省资源的“礼物”，行政主体坚持拒绝是违背历史潮流的，但这也导致了，不论行政主体乐意与否，它对技术的依赖性逐渐增强，而技术公司却借助算法系统的触角蚕食行政决策的权力。

3.3.1. 行政主体对算法的依赖增强

算法系统从设计到运行，无一不有技术公司的参与，面对愈加复杂愈加精细化的社会事务和高速发展的智能技术，不论从客观角度还是主观角度来看，行政主体对算法决策系统的依赖性逐渐增强。正如前文所述，行政机关通过行政合同委托技术公司来实现算法技术决策系统的开发与运行，因此，在自动化行政算法决策领域，缺乏专业知识与专门人才的行政机关不得不依赖技术公司进行行政执法；同时，算法决策系统的广泛应用又能大幅度地节约行政资源、提高行政效率，这又使得行政机关主观上对算法技术产生依赖性，直接或间接推动算法决策系统背后的技术公司渗入行政执法领域。同时，单一化、模式化的执法标准又削弱了行政机关工作人员的积极性和治理能力，忽视了对特殊个案的情形的考量。

3.3.2. 技术权力对行政权力的侵蚀

从技术公司的角度看，不断崛起的技术权力也在逐渐侵蚀行政权。早在第一次工业革命时期，大机器生产就催生了“机器权力”这一隐形的庞然大物，正如马克思所说，“科学、巨大的自然力、社会的群众性劳动都体现在机器体系中，并同机器体系一道构成‘主人’的权力[16]。”当机器体系发展为技术权力，其控制的范围更广、强度更大，比如在 2019 年，贵阳市多家政府部门联合推行了“数据铁笼”工程，以实现行政行为行使过程的全备案全记录[17]。由算法决策作出行政审批、行政处罚等处理决定，貌似将行政权力交给了技术系统，实则转移给了算法系统设计者，“权力与责任相重合”的行政权力被肢解，技术专家行使行政权力的怪圈正在形成。

4. 自动化行政风险的防范路径

面对来势汹涌的技术革命，人类并非无能为力，一些传统的规制仍然有效，只是应当对其进行重申或改造，聚焦到自动化行政领域中，可以从主体明定、义务厘定、程序确定三个方面入手，明确具体的

责任主体，强化相关主体的信息公示义务，并且完善相对人的权利保障与救济体系。

4.1. 明确责任主体

控制和防范自动化算法决策的特殊风险，首先需要明确各方主体责任，包括外部责任(对行政相对人承担的第一责任)和内部责任(根据风险控制理论由参与主体分配的第二责任)。具体可以分为以下三个方面：

4.1.1. 明确算法工具属性

关于算法是否具有独立的法律地位，近年来争论颇多，大体可以归纳为以下三种观点。第一种采肯定说，该说认为算法系统可以拥有独立的法律主体地位，有学者称，高度智能化的算法具有认识能力和行为能力，应赋予其“电子人”身份[18]；第二种采部分肯定说，该说认为对待算法，一般应坚持工具主义立场，但在法定特殊情况下，算法可以承担法律责任，如学者陈吉栋认为应当引入“拟制”技术，即在一般情况下并不强调算法的主体身份，只有当其决策或行为导致损害时才应赋予其法律意义上的责任主体地位[19]；第三种采否定说，认为不应当赋予算法系统独立的法律主体地位，有学者认为算法即便进化为强人工智能，也不具有独立承担责任的能力，涉及到财产责任时，也无法独立履责，允许算法成为法律责任主体是超出目前人类认知范围的[20]。

但笔者认为，至少在行政法领域，不应当赋予算法系统独立的法律主体地位。理由主要包括以下两方面：一是规避“技术专家执法”的风险。如果赋予算法独立的法律主体地位，那掌握了数据、代码与运行规则的技术公司及其工程师就间接地掌握了国家权力[21]。而且，机器学习依托的是历史数据，人类社会中的偏见和不公随着数据的输入被算法记忆、运行与学习，最终可能造就一个充满性别歧视、种族主义、地方保护主义等不良思想的算法系统，由这样的算法系统对社会进行治理会加剧歧视与社会分裂，最终可能造成无法挽回的结果[22]，因此应当坚持算法决策工具主义的定位。二是避免“技术盲区”造成的行政不公。算法决策依据的是可以被归为“0”和“1”的各项数据，人的情感因素、道德考量、习惯差异等，很难被算法决策系统公正纳入数据分析处理过程，对待这些“技术盲区”，只能依赖行政机关工作人员的公正合理的自由裁量。技术盲区的存在，同样决定了算法只能是辅助工具。

4.1.2. 明确行政主体对外第一责任人的性质

通过以上分析不难看出，自动化算法决策的背后，人类仍起主导和操控作用。鉴于当前技术仍无法将人类同理心、情感价值等因素考虑在内，技术专家也难以对其专业以外的复杂的社会事务进行公正的价值评判，因此，行政机关工作人员依然在行政决策中起着主导作用。换句话说，当具体行政行为作出后，继续强调行政主体对外第一责任人的性质，既能保持行政机关工作人员的警惕性，树立“多大的权力意味着多大的责任”的正确思想，保持兢兢业业的工作态度；又能避免技术权力借助高度自动化的算法决策系统侵蚀行政主体的自由裁量权，最终造就一个隐形而又无处不在的“算法帝国”；同时能在算法系统的辅助下，纳入行政机关工作人员的人为的道德、价值评判，以保障行政执法合法合理。但要注意的是，行政主体对相对人先行承担行政责任后，还享有追偿权，基于行政委托理论，行政主体可以在承担第一责任后，根据风险分担范围的不同对技术公司、网络平台、具体的行政机关工作人员等进行追偿。

4.1.3. 提升执法者的算法治理能力

算法治理能力包括两个方面。一是行政主体参与算法设计的能力。一项足够中立公正且完善的算法系统的诞生，不是仅靠技术专家闭门造车，而是需要融合法律从业人员、行政机关工作人员、心理学家、技术专业人员、一般公众等多种力量。行政执法者不能仅仅关心如何傻瓜式地操作一项已经成熟的算法

决策系统，而且要参与进算法程序的设计过程中，积极表达行政机关的意志，同时学习算法相关专业知
识，了解数据的收集方式，理解算法的运行规则，以更好地利用算法系统进行行政执法。二是行政主体
监督算法运行过程的能力。算法系统诞生后，不能立即用于社会治理，应当先进行试运行，并对试运
行的表现进行评估，在此期间，行政机关工作人员可以综合运用自己的法律知识、算法相关专业基础知
识、对民意的了解情况等，发现数据偏差或规则问题，并对其提出改进建议和方案。而在具体实践层面上，
应当通过加强行政机关的组织协同与行为协同提升行政执法者这两方面的能力，前者如召开不同政府部
门间的联席会议或者成立专门的领导小组，后者如相关的政府部门间进行联合执法或者行政协助^[23]。

4.2. 推动算法透明化

正如前文所述，算法黑箱是一个伪命题，算法透明化有必要实现且有可能实现。在推动算法透明化
的过程中，要注意明确“一刀切”的信息披露不可取；厘清政府公示义务；加强场景化法律规制；同时
丰富参与主体，推动软法治理。

4.2.1. 强调算法透明化的原因

为什么要强化信息披露，增强算法的透明度？首先，从正当权利的角度考虑，自动化行政算法系统
作出的决策对相对人的权利义务能够产生实质性影响，因此相对人有权要求行政主体对相关算法决策规
则进行解释，对涉及自身权益的算法运行规则进行了解属于法律所保障的知情权范畴。公民的正当权利
自然应当受到保护，这也是行政公开原则的应有之义。其次，从执法效率考虑，政府公信力也要求行政
主体对相关算法运行规则进行释明，在大数据时代，算法黑箱化无异于将历史倒退到“法不可知则威不
可测”奴隶制时代，民众对涉及自身权益的算法规则目录一无所知，可能会生出更多不满情绪，导致行
政复议、行政诉讼、信访的频率增加，破坏和谐的官民关系，增加了行政执法成本，降低了行政执法效
率。最后，从未来发展趋势来看，算法黑箱化与算法透明化是数字时代的一体两面，尽管囿于当前技术
水平限制以及知情同意规则、国家安全、商业秘密等的合理阻碍，算法运行规则还不能做到高度透明，
但是可以预见，随着科技的发展，机器学习的过程将不再神秘，公众监督的目光将会触及到科技巨头掌
控下的算法程序运作，减少算法偏见的发生。

4.2.2. 算法透明化的可行性

正如前文所述，算法不透明发生原因多样，并非完全不可公开。一方面，需要明确的是，要求对算
法进行公开和解释并非是将算法的所有源代码公开，使得任何不怀善意之人得以复制，而是将算法内
部的运行规则、相应数据的计算占比等进行公开，换句话说，是将源代码转换为法律规则后进行公开。另
一方面，针对对算法公开可能导致违法分子反向学习的担忧^[24]，有人主张对此可以成立独立的第三方算
法审计机关，由其专业人员进行算法审查，既保证了公开范围的有限性、审查的专业性，同时又能纳入
对社会公益的考量，因此，公开并非意味着任何公众任何范围都可知情。

4.2.3. 如何实现算法透明化

首先，应当明确，算法透明化至少应解决三个问题：向谁披露？披露多少？披露的时间？在对算法
进行披露时，应当全面考虑相关各方的合法权益，对披露的对象、披露的范围与披露的时限进行调整，
以实现“优质的透明度”³。其次，政府应当强化算法信息公示义务。政府部门应当公布行政决策中涉
及或运用算法的各项事务清单，并对各项事务中哪些环节运用算法，人工和算法各自占比是多少等因素列
明，以便让相对人了解自身权利义务被算法如何影响以及影响程度，同时该清单也可作为行政纠纷时的

³Frank Pasquale, *supra* note 15, at 7, 141, 218.

证据使用。再次，应当依场景公开信息。互联网、大数据、人工智能技术等的飞速发展，使得统一化、同质化的时代风潮如潮水般退去，个性化的定制服务、场景化的数据分析、细分的程序设计等等，逐渐占据社会主流[1]。即使针对同一对象，场景不同，治理规则可能就大不相同。比如，在对自动驾驶汽车进行规制时，应当考虑不同的驾驶场景下，车辆的速度、运行地段、所在路况、车载人货等不同因素，进行个性化的程序设计，如果仅仅是在平缓的、周围车辆与行人较少的路段进行物流输送，那就应当相对放松信息公示与说明的规范要求，鼓励企业进行创新，而如果是在车流量大或者路况崎岖复杂的路段进行驾驶，就应当严格规制，强制要求技术公司以及其他有关主体对算法运行规则进行公示和说明。最后，应当纳入软法治理。不断促进社会团体、行业协会等制定、更新关于算法决策领域的透明化规范内容，出台更具专业性的规定和标准，要求运营主体有针对性地对待原始数据收集、算法规则目录设置以及算法学习等问题，由政府外专业人士通过算法审计判断是否可以公开、公开对象是谁、以何种速度、何种方式公开，更具针对性和前沿性，也有利于最终实现技术助力下的法律善治。

4.3. 保障程序正义

在行政执法领域，程序和权利救济至关重要，强化程序性规范控制，对于算法决策系统的向善发展具有深远意义。如何实现自动化行政时代下的程序正义？至少应做到以下三个方面：

4.3.1. 贯彻告知与说明义务

首先，对于政府而言，在一般行政程序中，应尽到告知与说明义务，保障相对人的知情权，不能单纯依赖自动化执法机器。例如，英国在 2018 年《数据保护法》第 14 条规定，使用人工智能的公共管理部门必须在合理期限内以书面形式就自动化决策内容尽快通知被执行自动化决策的主体，数据主体可以在收到通知后一个月内，要求公共组织重新考虑对自己作出的自动化决策，或由自然人对决策进行重新评估⁴。其次，对于平台而言，在与政府合作，利用平台用户提供的个人信息辅助行政决策时，也有相应的告知义务，比如行政机关可能利用平台收集的大数据判断犯罪嫌疑人的行踪轨迹、证人证言可信程度、公务员报考者过往历史有无污点等，平台均应对所涉及个人信息及其处理方式进行提示与说明，以保障相对人的知情权与决定权[25]。最后，应当注意采用适当的告知说明方式。平台应采用足以引起相对人注意的字体、符号等，行政机关应注意面对不熟悉数字产品的中老年人、幼童，以及在认知方面有功能障碍的残疾人员，最好采取合适的图画、表格、字体等方式进行告知、解释与说明，以便其充分理解算法决策对其权利义务产生的影响。

4.3.2. 保障异议权、获得解释权和请求人工干预权

欧盟在《一般数据保护条例》中规定，在自动化决策情形中，数据主体有权请求数据控制者提供自动化决策的运行规则以及决策所可能产生的实质后果，并有权请求人工介入、对决策进行解释以及提出异议⁵。依照我国《个人信息保护法》第 24 条第 3 款，当自动化决策程序作出的决定对个人权益具有重大影响时，个人有权请求信息处理者进行说明，并有权对单纯依靠自动化决策程序作出的决定提出异议。这部分权利与行政主体的告知说明义务相辅相成，只有行政主体全面履行了告知说明义务，行政相对人才有可能获得充分的提请异议、获得算法解释以及请求人工介入、不受仅依靠自动化决策方式做出的决定影响的权利。

4.3.3. 落实听证程序

与传统听证多发生于具体行为实施后、行政决定作出前这一现象不同，自动化算法决策下的听证时

⁴Data Protection Act 2018, <https://www.legislation.gov.uk/ukpga/2018/12/contents>.

⁵General Data Protection Regulation, Recital (71) & art. 13 (2) (f).

间应当更为灵活。应增加算法系统上线前听证和行政决策作出后听证这两种形式。具体说来,在对公众具有重大影响的算法决策系统上线前,应当进行试运行,并就试运行的风险状况举行听证会,进行评估,广泛吸纳一般公众、行政机关与专业人士进行讨论分析。依照我国《数据安全法》第 22 条规定⁶,国家有关部门应建立数据安全的风险评估机制,加强信息获取、分析、研判以及预警工作,因此,行政机关通过行政合同购买的算法决策系统,在被正式投入运行前,应当进行试运行,经过公众参与的算法评估程序,并就试运行的结果举行听证会,在相关各方充分表达意见后对其进行改进或删除。

除此之外,在算法决策系统正式运行后,鉴于算法决策具有极速性的特征,将听证程序嵌入算法决策作出之前相当困难,因此,应当在算法系统做出决策后、行政机关真正作出生效决定前,举行听证程序,即在大数据时代,事后听证更为普遍,也能保障相对人获得充分而及时的救济。另外,与落实事后听证程序紧密关联的一点是,自动化算法决策系统应当让数据在一段时间内留痕,以便在需要时能够向上级机关、法院等提供跟踪记录。这种跟踪记录除了可以作为证据使用外,还可以作为算法的学习材料。此外,事后听证也有利于纠正行政机关工作人员将算法作为纠错程序的偏见。

5. 结语

近两百年来,科技以人类无法想象的速度改变着世界的面貌,大数据、算法、人工智能等词语逐渐深入人心。面对技术革新给法律制度带来的种种变化,审慎分析各项潜在风险,并提出可能的应对策略,才能长远造福于人类。具体在自动化行政领域,应当继续坚持行政主体第一责任人的地位,同时坚定算法工具主义立场,分场景规制算法决策在自动化行政中的应用,广泛开展与外部专业技术人员的合作,丰富治理主体,推动软法规制。并且,在坚持正当程序的前提下,对具体行政程序的内容和形式进行改造,保障新时代下行政相对人享有的各项正当权利。

参考文献

- [1] 马长山. 智能互联网时代的法律变革[J]. 法学研究, 2018, 40(4): 20-38.
- [2] 陈颢, 裴亚楠. 论自动化行政中算法决策应用风险及其防范路径[J]. 西南民族大学学报(人文社会科学版), 2021, 42(1): 74-81.
- [3] 中国安防展览网. 人工智能时代 伦理与创新之间该如何平衡[EB/OL]. <http://www.afzhan.com/news/detail/63997.html>, 2022-11-20.
- [4] 张凌寒. 算法自动化决策与行政正当程序制度的冲突与调和[J]. 东方法学, 2020(6): 4-17.
- [5] 李帅. 人工智能的风险预测与行政法规制——一个功能论与本体论相结合的视角[J]. 行政管理改革, 2019, 10(10): 50-58.
- [6] 胡敏洁. 论自动化行政中的瑕疵指令及其救济[J]. 北京行政学院学报, 2021(4): 81-88.
- [7] 罗伯特·考特, 托马斯·尤伦. 法和经济学[M]. 上海: 上海人民出版社, 2012: 456.
- [8] 莫于川. 行政法与行政诉讼法[M]. 北京: 科学出版社, 2008: 187.
- [9] 宋华琳, 孟李冕. 人工智能在行政治理中的作用及其法律控制[J]. 湖南科技大学学报(社会科学版), 2018, 21(6): 82-90.
- [10] Rivero, D. (2019) Miami Could Let Company Put Surveillance Poles on Public Property for Free. WLRN. <https://www.wlrn.org/post/miami-could-let-company-put-surveillance-poles-public-property-free#stream/0>
- [11] Carr, N. (2015) The Glass Cage: Where Automation Is Taking Us. Random House, New York, 94-95. <https://doi.org/10.1049/et.2015.0234>
- [12] 张兆端. “智慧警务”: 大数据时代的警务模式[J]. 公安研究, 2014(6): 19-26.
- [13] Aylin, C., Bryson, J.J. and Narayanan, A. (2017) Semantics Derived Automatically from Language Corpora Contain

⁶ 《中华人民共和国数据安全法》第 22 条规定, 国家建立集中统一、高效权威的数据安全风险评估、报告、信息共享、监测预警机制。国家数据安全工作协调机制统筹协调有关部门加强数据安全风险信息的获取、分析、研判、预警工作。

-
- Human-Like Biases. *Science*, **356**, 183-186. <https://doi.org/10.1126/science.aal4230>
- [14] 李彦宏, 等. 智能革命: 迎接人工智能时代的社会、经济与文化变革[M]. 北京: 中信出版社, 2017: 228.
- [15] Surden, H. and Williams, M.A. (2016) Technological Opacity, Predictability, and Self-Driving Cars. *Cardozo Law Review*, **38**, 121-181. <https://doi.org/10.2139/ssrn.2747491>
- [16] 马克思. 机器、自然力和科学的应用[M]. 北京: 人民出版社, 1978: 26.
- [17] 李玲. “数据铁笼”工程让权力运行处处留痕[N]. 南方都市报, 2019-03-10(A10).
- [18] 郭少飞. “电子人”法律主体论[J]. 东方法学, 2018, 63(3): 38-49.
- [19] 陈吉栋. 论机器人的法律人格——基于法释义学的讨论[J]. 上海大学学报(社会科学版), 2018, 35(3): 78-89.
- [20] 郑戈. 人工智能与法律的未来[J]. 公民与法(综合版), 2017, 466(12): 11-15.
- [21] Veblen, T. (2010) *The Engineers and the Price System*. New York University Press, New York, 135-137.
- [22] 张玉宏, 秦志光, 肖乐. 大数据算法的歧视本质[J]. 自然辩证法研究, 2017, 33(5): 81-86.
- [23] 孔祥稳. 面向人工智能风险的行政规制革新——以自动驾驶汽车的行政规制为中心而展开[J]. 行政法学研究, 2020(4): 18-33.
- [24] Sandvig, C., et al. (2014) Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms. *Annual Meeting of the International Communication Association*, Seattle, 22 May 2014, 1-23.
- [25] Gelhorn, E. (1973) Adverse Publicity by Administrative Agencies. *Harvard Law Review*, **86**, 1380-1441. <https://doi.org/10.2307/1340032>