

风险等级化视角下行政算法的差序化公开

王昊禄

(中央财经大学 法学院, 北京 100098)

摘要: 当从行政法秉持的行政公开性立场对行政算法予以关注并对其透明度规制进行制度性回应时, 与其追求一个难以普遍适用的单一路径, 不如采用“差序化”监管的理念, 顺应自动化行政中算法的不同特点, 在方法论上放弃统一适用的公开路径, 转而透过对实践的总结和归类, 以算法对权利的实质影响程度为标准, 抽象出算法介入自动化行政的不同风险等级, 并对其适配相应的公开要求, 明确其公开程度与具体的公开方式, 最大程度上规避自动化行政中算法实质风险的不确定性与法律规范所追求的普遍适用性之间的矛盾。

关键词: 行政透明; 政府信息公开; 算法风险等级; 差序化公开; 分类分级监管

中图分类号: D922.1; D922.1

文献标识码: A

DOI: 10.19358/j.issn.2097-1788.2025.07.010

引用格式: 王昊禄. 风险等级化视角下行政算法的差序化公开 [J]. 网络安全与数据治理, 2025, 44(7): 66-73.

Differential ordering disclosure of administrative algorithms from the perspective of algorithm risk classification

Wang Haolu

(Law School, Central University of Finance and Economics, Beijing 100098, China)

Abstract: When paying attention to algorithms in automated administration from the standpoint of administrative openness upheld by administrative law and institutionally responding to its transparency regulation, rather than pursuing a single path that is difficult to be universally applied, it is better to adopt the concept of "differential" supervision, conform to the different characteristics of algorithms in automated administration, and abandon the uniformly applicable open path in methodology. Instead, through the summary and classification of practice, the paper abstracts the different risk levels of the algorithm's intervention in automated administration based on the actual impact of the algorithm on the right, and adapates the corresponding disclosure requirements to it, clarifying its disclosure degree and specific disclosure methods, to avoid the contradiction between the uncertainty of the substantial risk of the algorithm in automated administration and the universal applicability pursued by legal norms.

Key words: administrative transparency; disclosure of government information; algorithm risk level; differential ordering disclosure; classified and hierarchical supervision

0 引言

在数字化浪潮的推动下, 技术领域迎来了颠覆式的革命, 这使得算法逐渐嵌入到各行各业的发展中。随着数字技术的迅猛发展, 人工智能、大数据和云计算等技术的应用逐渐从私法领域扩展到公法领域, “算法决策在私领域的日趋成熟为其应用于公领域打下坚实的技术基础和心理基础, 通过算法技术辅助行政决策成为行政领域发展的可视化趋向——自动化行政。”^[1]

自动化行政已成为现代政府治理和公共服务领域的一项重要趋势。在人工智能时代, 行政部门越来越多地通过使用智能算法来代替原本由人类参与的行政过程^[2]。

这种技术驱动的行政管理模式在带来便利的同时, 也引发了一系列法律、伦理和社会上的问题。由于算法的复杂性和专业性, 公众往往难以知晓其工作原理和决策过程, 这会导致公众对自动化行政产生质疑和不信任, 给行政透明度和公开性带来挑战。

行政透明要求行政机关在行使公权力过程中, 向社会公开其决策、执行和监督的过程, 接受社会的监督, 保障公民的知情权和参与权, 防止权力滥用。而算法因其本身的不透明性、自动性、高度的复杂性和模糊性等特征, 使行政机关在运用算法进行决策时, 难以向公众公开其决策依据和过程, 这在一定程度上背离传统的行

政程序正当性原则，导致法律赋予相对人的陈述、申辩、说明理由、听证等程序性权利被悬置，从而影响行政透明的实现。相对于那些隐藏在算法模型背后，未实现信息公开的“数字利维坦”而言，行政相对人在未获知算法决策的作出过程和依据的情况下，只能被动接受由算法模型自动生成的决策方案。这意味着，处于弱势地位的行政相对人在一定程度上已经沦为一个被高度透析的“微粒人”或高度透明的“数字化主体”^[3]。由此可见，在自动化行政中算法的制定者与相对人之间横亘着一道深深的数字鸿沟，这甚至可能加剧两者之间的信息不对称格局^[4]。在此背景下，如何平衡效率与透明度，确保公众能够理解和信任自动化行政中的算法决策，成为了一个亟待解决的问题。

在自动化行政中，自动化系统运行算法的过程相当于法律运行的过程，算法作为自动化行政中的技术手段，受实践中实质影响程度、风险等级与应用场景差异等影响，如何确定其公开方式与公开程度是当下算法规制所面临的复杂课题^[5]。本文着眼于算法在行政法体系中的定位，尝试从政府信息视角出发将算法纳入行政公开框架；其次，以现有的算法公开方案为基础，提出并讨论“差序化”的等级化监管理念；再次，以自动化行政中算法的实质影响程度为标准对其进行分级；最后，在等级化的基础上具体分析自动化行政中算法宜采取的公开程度。本文的基本立场是，为维护原有的法治秩序与行政透明性，自动化行政中算法的规制应当从行政公开原理出发，通过自动化行政中算法的风险等级化、场景的类型化适配不同的公开程度，以实现行政算法透明。通过构建算法公开制度，制定算法透明度标准，实现在自动化行政中算法公开领域的差序化监管。

1 行政算法的行政法属性

关于算法的法律属性，目前学界主要形成了“规范论”与“技术论”两种观点，前者以约耳·芮登博格（Joel Reidenberg）提出的“代码即法律”为代表，将算法视为规范本身^[6]。后者则以格罗斯（Groß）提出的算法即“承载法规范的计算程序”（normativ aufgeladene Programme）为代表，将算法视为法律规制的客体，即实现法律规范的工具^[7]。为了摆脱如上关于“算法即规范”与“算法即客体”的认知模式争论，避免陷入算法定性或“技术”或“规范”的绝对视角，在理解算法时不应过分拘泥于“统一性”。美国行政法学者科利亚尼斯（Coglianese）将行政算法区分为“自动化规则制定”（robotic rule making）和“算法裁决”（algorithmic adjudication）^[8]。带着这种“区分性”的视角去审查算法，便会

区分出静态算法与动态算法。前者即算法规则本身，对其进行定性应更多关注制定算法的价值与意义，关注其本身的抽象特性，阐明算法本身的属性；后者即算法运行，对其进行定性应更多关注算法应用的实时性与决策影响，阐明算法应用的属性。

1.1 静态行政算法实为行政规范

关于行政法视域下算法规范以及与算法有关的行为规范的属性，有观点认为：算法可以通过具体的应用，直接对特定的行为人、特定的事项施加影响，故其应属于具体行政行为。如前文所述，行政算法不仅包括动态的算法运行，还包括静态的算法规则制定，这种“具体行为说”仅目及动态的算法运行，而忽略了静态算法规则的定性。实际上，算法规则本身具有复杂的特性，其作为自动化行政的行为依据，具有不可忽略的研究价值。与之相对，“行政规定制定说”则认为：行政机关在作出行政行为过程中所使用的这种带有抽象性、执行性、基准性的技术规则具有行政规定的性质^[9]。换言之，行政机关为做出行政行为而设计、使用的能考虑各种事实或规范因素的技术规则属于行政规定^[10]。

静态的算法规则通常是指那些可供自动化行政使用的、稳定且不随时间变化而改变的计算准则与决策逻辑，这些规则构建起解决问题的技术机制，为后续的行政决策设定了框架和依据。具体而言，行政机关通过技术手段制作算法，设定解释数据、做出判断和执行任务的规则，辅助行政机关行使职权，进而提升行政效率。故算法在行政法体系中具有类似法律规范的性质^[11]。首先，算法规则的实质制定者是行政机关^[7]。其次，算法规则的制定依据是行政规则^[12]。再次，算法规则的制定目的在于优化、细化行政规则的功能。最后，算法规则具有规范属性。静态的算法规则作为自动化行政的基础，会对公民的生活和权益产生影响。

关于行政规范的制定，在法学上往往具有不同的含义。有的学者曾从所制定的法律规范的性质来界定，认为凡是制定行政法规范的行为，不论制定主体的性质如何，都属于行政法^[13]。有的学者则认为，行政立法既应当从机关性质界定，又应当从所制定法律规范的性质来界定，即只有行政机关制定行政法规范的活动才是行政立法^[14]。行政算法规则虽然不具备传统形式上的文本规范性，但其通过电子数据化的形式嵌入到行政过程中，以自动化系统为载体，可以对相对人产生实质上的约束力，在事实上具有法律规范的功能，可以满足行政规定的要式，实为一种特殊载体形式的行政规范。同时，行政算法规则是一种由行政机关制定的具有普遍约束力的规则，这在一定程度上也与我国行政诉讼法与行政复议

法中的相关规定相契合。综上,在区分静态的算法规则与动态的算法应用的基础上,行政算法规则实为一种特殊形式的行政非制定法源,整体上可以被纳入行政规范的范围^[9]。

1.2 动态行政算法实为政府信息

随着政府信息公开研究的不断升温,世界各国和地区都在逐步建立政府信息公开制度。从20世纪80年代的“村务公开”到90年代的“政务公开”,再到21世纪初对政务公开的体制探索,我国的政务公开工作得到了稳定发展并取得丰硕成果^[15]。2008年《中华人民共和国政府信息公开条例》(以下简称《政府信息公开条例》)出台,首次从国家立法的层面确立了信息公开原则,具有极其重大的立法意义。

《政府信息公开条例》中所界定的“政府信息”,是指行政机关在履行职责过程中制作或者获取的,以一定形式记录、保存的信息。也有学者将“政府信息”定义为“资料”^[16]。2019年新修订的《政府信息公开条例》进一步将“履行职责的过程”明确为“履行行政管理职能过程”,这在一定程度上扩大了“政府信息”的涵盖范围。随着自动化行政的快速发展,《政府信息公开条例》所界定的政府信息已无法全面适应社会的发展和法律的进步。动态算法运行中的信息,作为自动化行政过程中产生的数据和决策,应当属于政府信息。

首先,算法具有决策性。在自动化行政中,算法通常被用于辅助或直接做出决策。其次,算法具有公共性。提高行政效率和服务质量是自动化行政的重要目的,而算法之所以得以嵌入行政过程,也是行政机关为了实现这一目的所做出的选择和设计。最后,算法具有数据性。行政算法往往会涉及大量政府收集或产生的数据的处理,算法将这些数据进行处理并产生输出结果,这一过程实质上就是对政府所获取的数据的利用和加工。故动态算法运行中所依赖的数据及其处理结果均应被视为行政机关在履行职责过程中获取的信息。同时,需要强调的一点是,虽然算法运行具有数据性,但“数据性”是其作为电子数据化的政府信息的属性之一,仍对应“政府信息公开”,而非“政府数据开放”^[17]。

根据《政府信息公开条例》,政府信息包括行政机关在履行职责过程中制作或获取的信息。行政算法运行虽然并未在《政府信息公开条例》中明确体现,但其是行政机关为提高行政效率所采取的一种以自动化系统为载体,以电子数据记录储存的信息,可以对公共利益产生实质上的影响,具有政府信息属性,实为一种特殊载体形式的政府信息。综上,在区分静态的算法规则与动态的算法应用的基础上,行政算法运行实为一种特殊形式

的政府信息,整体上可以被纳入政府信息的范围。

2 行政算法的可公开性证成

在区分静态的算法规则与动态的算法应用的基础上,静态的算法规则可归属于行政规范,动态的算法运行可纳入政府信息的范围。关于行政法体系中对算法公开的讨论,应继续回到行政规范公开的原理与制度框架中,探寻自动化行政中算法公开的可行性与可欲性。

2.1 行政算法公开具备法律正当性

自动化行政中算法的可公开性是由其本质所决定的。《政府信息公开条例》第5条规定了政府信息应当坚持以公开为常态、不公开为例外,并遵循公正、公平、合法、便民的原则。同时,《政府信息公开条例》第19条、20条说明了法定主动公开的内容。一方面,第20条第1项正面列举了“行政法规、规章和规范性文件”。而静态的算法规则虽然不具备文本形式的行政法规、规章和规范性文件载体,但从立法目的上来看,算法规则作为一种特殊载体的行政规范,也应属于主动公开的政府信息^[11]。另一方面,结合第20条其余规定与反面兜底规定,动态的算法运行可能涉及各个行政管理领域,对于自动化行政中涉及公众利益调整、需要公众广泛知晓或者需要公众参与决策的算法运行信息^[9],行政机关应当主动公开。

此外,针对算法技术本身,我国也陆续在立法层面提出了算法公开要求。首先,针对是否应当公开,《电子商务法》和《个人信息保护法》在明确算法自动化决策的自然结果展示义务的同时,也对自动化决策的算法提出了透明度要求。其次,针对如何公开,《算法综合治理指导意见》与《网络数据安全管理条例(征求意见稿)》等从宏观层面强调了算法应用应“透明可释”,并提出要构建算法决策披露制度。最后,针对构建何种算法公开制度,《互联网信息服务算法推荐管理规定》重点规定了算法推荐服务提供者的披露义务和责任,并在第23条明确指出“网信部门会同电信、公安、市场监管等有关部门建立算法分级分类安全管理制度”。总而言之,关于算法“应该公开”的问题,上述规定已经给出政策性认同。

2.2 行政算法公开具备技术可行性

行政算法因其电子数据的形式与自动化行政的载体,而有别于传统的行政规范与政府信息,那么行政算法的公开是否因其技术特征而存在技术障碍?算法公开的技术障碍主要是“算法黑箱”问题。在理论层面上,算法黑箱是指算法制定主体以外的人对算法的模型结构、调用数据、算法测试等内容无从获悉,因无法观察或者理解算法,算法制定主体以外的人处于完全的被决定者地位,沦为被算法决定的客体^[9]。在实践层面上,在行政

算法公开的实践中，行政机关可能会以行政行为是由算法自动作出的为由，借算法黑箱而逃避公开^[18]。

对于“算法黑箱”是否存在，其决定权与主导权究竟属于技术特性还是制定主体，学界存在不同看法。“算法黑箱肯定说”认为，算法黑箱由算法的技术特性所造成。与此相对，“算法黑箱否定说”则认为，算法黑箱的主导者是算法的制定者，而非技术本身^[19]。着眼于自动化行政领域，对于行政机关基于既有的行政规范而设计、制定、采用的算法，其“行政属性”在一定程度上可以破解“算法黑箱”问题。

首先，从自动化行政中算法的制定角度来看，制定主体与制定依据是既定的，而是否选用算法、选用何种算法都是行政机关的主动选择^[10]。其次，从自动化行政中算法的本质属性角度来看，行政算法是技术规则与行政规则的结合物，技术规则是其外相，即其表现形式，而其内里的行政规则才为其本质，故无论具体的算法以何种技术规则形式存在，都不能否定由其行政规则本质所决定的可公开性^[11]。最后，从自动化行政中算法的发展现状角度来看，现阶段的算法技术并未完全脱离人类的控制，也就是说，当下的机器学习尚未达到强智能阶段，也就不存在绝对的算法黑箱。

行政算法公开，其主要目的是为了“限制公权力”，所以在讨论自动化行政领域的“算法黑箱”时，不能忽略行政主体、行政规范等“行政性”要素。行政算法有别于一般的商业算法应用，其属于行政机关应主动公开的政府信息，在不涉及国家秘密的前提下，其算法应当充分透明^[20]。为了刺破这种披着算法技术外衣的权力实质表达，首先应该坚定地进行“算法公开”，而对于公开后可能出现的问题与暴露的缺陷，则可以通过调整公开范围、公开程度、公开手段等予以规避，而不应因此而动摇算法的可公开性。

3 行政算法公开的传统方案及困境

在我国现阶段的自动化行政立法与实践中，对于自动化行政中算法所依据的法律规则，政府应遵循“公开原则”对其进行公开，而对于自动化行政中因算法技术而产生的算法规则到底如何进行算法公开，尚无制度规范。一方面，在算法公开的立法层面，目前仅在私主体领域对自动化决策算法的透明度提出了算法结果公平公正的原则性要求，仍缺乏对算法公开范围、公开程度的具体规定。另一方面，在算法公开的实践层面，我国尚无统一的监管与规则制度，这在一定程度上导致了行政算法公开与《政府信息公开条例》脱节的问题。目前，针对行政算法公开，国内外学者对于是否全部公开形成

“支持”与“反对”两种观点，同时，因对技术公开的要求不同与对算法公开范围的要求不同，形成了三种主要的公开方案，如表 1 所示。

表 1 行政算法公开的传统路径

公开路径	文本规则	技术规则	
		范围	方式
说理式公开	全透明公开	无限制	算法解释
定向性公开	全透明公开	限定开源	算法公开
全透明公开	全透明公开	无限制	算法公开

3.1 行政算法的说理式公开：不利于行政效率

说理式公开是自动化行政中算法的最小范围公开，即对算法所依据的法律规则进行公开，而对算法运行的技术规则仅进行“释明”即可。此为反对全部公开者所持观点，他们认为算法公开的最佳方式应为说理式公开。说理式公开中的“说理”，即为公开的一种方式，使用说理的手段，以达到公开的目的。行政机关对于其所制定、使用的算法规则，仅需对算法嵌入行政所涉及的部分关键要素进行说明，而无需对整个规则、流程进行公开^[21]。

然而，说理式公开的方式既难以与我国的政府信息公开制度相契合，不具备制度基础，亦会加重行政机关的负担，不利于提升行政效率。一方面，我国政府信息公开制度的核心在于“公开”，而说理式公开的核心则在于“说理”。“说理”实际上是一种“解释”，在公开的基础上如果对行政机关附加“说理”的义务，则可能在一定程度上超越政府信息公开制度的框架。另一方面，政府信息公开在方便公众周知的时候，不应过高增加公开成本和妨碍行政效率^[22]。

3.2 行政算法的定向性公开：有悖于制度理念

定向性公开是自动化行政中算法的中等范围公开，即对算法所依据的法律规则进行公开，而对算法运行的技术规则进行“定向公开”。所谓的“定向公开”，实际上就是在全透明公开的支持与反对之间找寻了一个折中路径，通过附条件在全透明公开与否中选取了一个平衡点，即向无利害关系的第三方进行全透明公开——“定向开源”。目前关于“定向开源”的具体方式，学界尚有不同观点，有学者主张限定公开面向，即仅对“政府治理决策者以及对行政决定不信任态度的直接行政相对人”实施有限的公开^[5]。有学者则主张限定公开类型，即以算法类型为基础来区分公开面向，尤其是对于具有技术壁垒的学习型算法，应将其公开面向限制于“专业且中立的第三方技术人员”^[23]。此外，英国下议院科技委员会也要求政府公布和印发中央政府系统中重要算法

引擎的使用列表,向无利害关系的第三方“定向开源”。

然而,定向性公开的路径既不具有制度基础,亦与政府信息公开制度的基本价值相背离。具体而言,无论是限制公开类型、缩小公开范围或是其他形式的“定向公开”,其在本质上都是给政府信息公开增加了额外的限定条件,限制了政府信息的公开面向,这有悖我国政府信息公开制度的“扩大化”趋势。

3.3 行政算法的全透明公开:受限于单一方案

全透明公开是自动化行政中算法的最大范围公开,即不仅对算法所依据的法律规则进行公开,而且对算法运行的技术规则也进行公开。此为支持全部公开者所持观点。全透明公开即意味着行政机关需要对行政法规规则进行全部公开^[9],除非算法因特定治理需要经法定程序被认定为应当保密,否则对于一般公众可理解、不存在技术隔阂的自动化行政算法,均应当无条件公开。实践中,上海市公共数据开放平台通过提供2000多个数据接口的方式向公众提供了代码信息访问渠道,以实现最大范围的算法公开^[23]。

然而,全透明公开路径虽然契合我国政府信息公开制度的原则和精神,但这种过于激进的“大一统”公开路径在一定程度上忽略了制度与实践的复杂性。一方面,全透明公开路径忽略了《政府信息公开条例》第14条规定的不予公开事由。另一方面,受实践中技术介入程度、实质影响程度与应用场景差异的影响,行政算法外延不断泛化。无论是何种透明路径,于纷繁复杂且快速发展的算法技术而言,始终都难以实现一体化适用^[24]。

4 行政算法差序化公开的制度构建

随着算法技术的不断发展,行政算法样态日益复杂,但制度上却依旧缺乏明确的规制。本文认为当从行政法秉持的行政公开性立场对行政算法予以关注并对其公开规制进行制度性回应时,与其追求一个难以普遍性适用的单一路径,不如顺应自动化行政中算法的不同特点,透过对实践的总结和归类,在方法论上放弃统一适用的公开路径,转而从等级化的角度,以算法介入自动化行政对权利的实质影响程度为标准,抽象出算法介入自动化行政的风险等级,规避技术介入程度与风险差异的不确定性与法律规范所追求的普遍适用性之间的矛盾,为从制度上实现行政算法公开提供前提。

4.1 行政算法公开的衡量标准:风险等级化

以行政算法对实质权利的影响程度为标准对自动化行政进行研究,首先应该考虑的就是价值衡量。面对算法公开可能面临的利益冲突,应该明确行政公开性、公民知情权与商业秘密、个人隐私之间的法益衡量,寻求

自动化行政中算法的公开与商业秘密、个人隐私之间的平衡路径。

行政公开性所涉及的法益往往具有公共性与重要性。首先,自动化行政作为一种新型的行政治理模式,算法在其中发挥了关键的工具作用。公开自动化行政中算法意味着打开公众参与政治的入口,有利于保障公民的政治权益。其次,部分行政算法涉及公共采购、机构运营、资金配置等财政领域,对其进行公开有利于保障公民的监督权。最后,对于具体应用于食品、药品、税务、司法、治安以及社会保障等领域的自动化行政,算法的嵌入可能涉及公民的生命健康、人身自由及生活保障等,因此对其公开有利于保障公民的人身及财产权利。与之对应,商业秘密与个人隐私涉及的法益往往更具经济性与私人性。一方面,商业秘密的核心价值就是经济利益,商业秘密的公开往往会对私主体的财产利益、竞争利益、创新利益造成影响,更严重的可能对多个私主体造成影响进而影响经济市场。另一方面,个人隐私则往往涉及单一或特定主体的人格尊严与自由。个人隐私的公开可能对信息主体的隐私权益造成影响。由上分析可知,在行政公开性与商业秘密、个人隐私之间的法益衡量中,行政公开所涉及的公民的基本权利以及社会秩序是主要法益,而商业秘密与个人隐私处于相对次要的位置。因此,分析行政算法对实质权利的影响程度时,应以保障算法公开与公民基本权利为第一要务,兼顾商业秘密和个人隐私保护^[25],充分考虑行政算法所涉价值目标的多元性与所涉利益范围的广泛性^[26]。

在社会治理的实践中,根据项目对象负外部性程度的大小对其进行分类管理的“影响评估”制度与本文所研究的算法公开制度具有一定的适配度^[27]。自动化技术治理领域的“影响评估”制度源于20世纪90年代的隐私影响评估制度和数据处理活动评估制度,2019年美国国会发布的《算法问责法案》也提出了“自动化决策系统影响评估”,进一步,2019年加拿大颁布的《自动化决策指令》提出从算法决策可能引发的风险以及风险的可控性出发,结合具体场景对算法适用分级保护和监管。具体而言,《自动化决策指令》以自动化行政决定所影响的当事人的权利类型与自动化行政决定所产生的风险的持续时间与可逆性为双重标准,根据自动化行政决定对个人、群体或者社会的影响,将自动化行政分为四个等级:Ⅰ级影响,即行政决定通常会对上述因素产生可逆且短暂的影响;Ⅱ级影响,即行政决定通常会对上述因素产生可逆的短期影响;Ⅲ级影响,即行政决定通常会对上述因素产生难以逆转的持续性影响;Ⅳ级影响,即行政决定可能会对上述因素产生不可逆的永久性影响^[28]。

这种根据算法应用的风险程度及其可控性为标准实现技术的差序化治理的思路,同样可在实现自动化行政中算法透明时被纳入^[29]。

兼顾全面性与适应性的前提下,为了对自动化行政中算法可能造成的实质权利影响进行精准识别、预防和救济,自动化行政中算法的实质权利影响评估制度可参照上述制度理论和立法实践采取分级管理。如表2所示,零级影响主要指行政算法应用并不会对公民个人、群体或者社会造成负面影响,其并不会影响公民权利,或仅仅具有授益性而不产生义务。一级影响主要指行政算法应用会对公民个人、群体或者社会造成负面影响,且其对公民权利造成的这种影响是短暂且可逆的,换言之,在这种影响结束之后,所受影响的权利是可恢复的。二级影响主要指行政算法应用会对公民个人、群体或者社会造成负面影响,且其对公民权利造成的这种影响是可逆的。同时,虽然这种影响具有持续性,但持续时间较短。三级影响主要指行政算法应用会对公民个人、群体或者社会造成负面影响,且其对公民权利造成的这种影响难以通过其他手段进行逆转并恢复原状。同时,这种影响具有较长时间的持续性。四级影响主要指行政算法应用会对公民个人、群体或者社会造成负面影响,且其对公民权利造成的这种影响不可逆转,同时,这是一种永久性的影响。

表2 行政算法的影响等级

影响等级	是否影响公民权利	影响程度	
		影响持续时间	影响是否可逆
零级影响	无影响	/	/
一级影响	影响	短暂	可逆
二级影响	影响	短期	可逆
三级影响	影响	长期	难以逆转
四级影响	影响	永久	不可逆

4.2 行政算法公开的方式划分:程度层次化

确定行政算法公开程度的目的在于尽可能使公众能够获得和信任算法的决策过程,从而提高算法的可接受性和合法性。行政算法公开和透明是行政公开的原则性要求,算法的具体公开事项则是算法公开程度的层级问题。以自动化行政中算法公开程度为视角,可通过分级认定精准识别具体算法应用应该如何公开。

第一,基本公开主要指按照政府信息公开的基本要求,增强自动化行政中算法的基本可见性。这是最低程度的算法公开。基本公开对行政机关提出的公开要求较低,仅要求其使用简单的手段对行政算法应用进行直接

公开,比如我国的《个人信息保护法》与《算法治理指导意见》,前者宏观性地提出算法公开透明、结果公平公正,后者原则性地提出算法透明可释。二者均未涉及算法公开的具体事项,可视为算法的基本公开。

第二,中等程度公开主要是指要对行政算法“从始至终”的应用情况作出全流程的公开。中等程度公开对行政机关提出的公开要求较高,在基本公开的基础上还要求其对算法应用之“始”即算法参与行政决策的强度和领域与算法应用之“终”即算法对决策结果的影响程度也进行公开^[30]。比如,欧盟在GDPR要求公开算法的存在、算法运行逻辑和算法结果的基础上,对算法的透明度提出了更高的要求,在《平台与商户间公平性和透明度法》中要求公布影响算法自动化决策的重要参数,可视为算法的中等程度公开^[31]。

第三,深度公开主要是指对自动化行政中算法的设计原理、工作机制、源代码、数据集等深层次信息进行公开。深度公开对行政机关提出的公开要求极高,在中等程度从始至终公开的基础上,深度公开还要求其对算法应用的相关信息进行深入挖掘。比如,除了《过滤气泡透明度法案》中要求的算法透明,公开算法运行逻辑、结果和重要参数外,纽约州的议案甚至还要求政府机构披露所使用算法的源代码,可视为算法的深度公开^[31]。

4.3 行政算法公开的方案适配:公开差序化

从自动化行政中算法的公开程度角度入手,以前文对自动化行政中算法风险等级的划定为基础,按照实质权利影响决定公开程度的逻辑,可以对自动化行政中算法的差序化公开做如下分析:

第一,自动化行政中的零级影响算法无需强制公开。自动化行政中的零级影响算法并不会对公民个人、群体或者社会造成负面影响。故自动化行政中的零级影响算法的透明性,并不是行政透明、算法透明的重点所在,也无需对其提出强制公开的要求。

第二,自动化行政中的一级影响算法适配基本公开。因为其在实质上会对公民权利造成影响,故需对其提出公开要求,同时,其造成的这种影响是短暂且可逆的,所以对其公开要求可以较低。因为这种影响往往可以在非常短的时限内被发现并修正,所以在确保其基本可见性的同时也要避免对其提出严苛的公开要求而浪费行政与算法资源。换言之,可以以《政府信息公开条例》为基本依据,宏观性地确保其算法公开即可。比如对于已经及时识别与修正错误的民政部门分发福利物品的算法系统,仅要求民政部门对其所用的算法进行直接公开,明确算法的存在、公开算法的运行逻辑和运行结果^[31],即可满足对自动化行政中的一级影响算法的公开性要求。

第三,自动化行政中的二级影响算法适配中度公开。首先,因为其在实质上会对公民权利造成影响,故需对其提出公开要求。其次,因为其造成的这种影响具有持续性,所以对其公开要求可以适当提高。最后,这种影响的持续时间较短,且最终可以被修正逆转,所以对其公开要求也不宜过高。应该在确保其全程透明性的同时避免对其提出过分严苛的公开要求而浪费行政与算法资源。比如对于识别与修正错误具有一定延迟性的民政部门分发福利物品的算法系统,可以要求民政部门对其所用的算法进行全流程公开,明确算法参与行政决策的强度和领域与算法应用对决策结果的影响程度等信息,以满足对自动化行政中的二级影响算法的公开性要求。

第四,自动化行政中的三级影响算法适配深度公开。一方面,因为其在实质上会对公民权利造成影响,故需对其提出公开要求。另一方面,因为其造成的这种影响具有较长时间的持续性,且对这种影响进行修正逆转需要耗费大量行政资源,换言之,其对公民权利造成的负面影响的“纠错成本”较高,所以应提高对其的公开要求。为节约行政资源与算法资源,并尽量减少这种算法错误,敦促行政机关加强监督,应确保其从始至终、由表及里的透明性。比如对于辅助刑事司法证据审查的算法系统,应该要求司法部门深入其使用算法的内部,对算法的设计原理、工作机制、源代码、数据集等深层次信息进行公开,只有这样才能寻找到算法系统出错的根本原因并促进其优化与监督,避免之后出现类似的问题,以满足对自动化行政中的三级影响算法的公开性要求^[31]。

第五,自动化行政中禁止授权使用四级影响算法。自动化行政中的四级影响算法应用会对公民个人、群体或者社会造成永久且无法逆转的负面影响。对于四级影响的算法,应延续法律保留的思路,把行政权力紧紧握在人的手里,从源头上消灭算法技术可能出现的误差。故对于四级影响的算法,其并不具备自动化行政适用的法律正当性,更谈不上讨论其公开问题。

行政算法的差序化公开如表3所示。

表3 行政算法的差序化公开

影响等级	影响程度	公开程度
零级影响	无影响	非强制性公开
一级影响	短暂+可逆	基本公开
二级影响	短期+可逆	中度公开
三级影响	长期+难以逆转	深度公开
四级影响	永久+不可逆	禁止授权

综上,通过前文对行政算法公开的分析,可以初步得出如下行政算法差序化公开的制度模型。首先,通过衡量算法介入自动化行政对权利的影响程度,以权利影响从低到高可以将行政算法划分为零级影响(无影响)、一级影响(短暂+可逆)、二级影响(短期+可逆)、三级影响(长期+难以逆转)与四级影响(永久+不可逆)。其次,从公开程度角度,以算法介入自动化行政对权利的影响程度为基础,排除因算法介入自动化行政对权利无影响而无需对算法公开问题进行讨论的零级影响算法,以及因算法介入自动化行政对权利影响极大而停留于可否适用性层面尚未触及算法公开问题的四级影响算法,结合自动化行政中算法影响的其他分类,从低到高可以将特定的行政算法的公开程度确定为:基本公开、中度公开与深度公开。最后,针对具体应用场景,结合前述分析所得公开范围与公开程度结果,以公开要求从宽松到严格匹配具体的公开方式,例如:发布文字说明、流程图、结构图以及技术报告和白皮书等。

5 结束语

在算法技术迅猛发展的今天,算法的介入正在重塑行政行为模式与流程,传统单一的算法公开路径已难以应对行政算法应用对行政正当程序带来的挑战。为保持行政法在算法时代的生命力,行政法律体系应当对这一新趋势做出回应。算法技术具有局限性,行政权力具有扩张性,站在技术主义与法治主义的双重立场上,算法技术的应用也需内在契合行政法的基本价值诉求和理论体系,应对其提出相应的透明度要求,避免算法黑箱、法规空白和行政机关的消极态度等阻碍自动化行政中的算法公开。同时,不同应用场景中,行政算法对权利的影响程度亦有不同,因此制定一个合理的差序化机制,对不同级别的算法应用施以相应的公开要求和规则,是实现行政公开资源优化配置的关键。鉴于此,提出行政算法差序化公开制度显得尤为重要,这不仅关乎传统行政正当程序的原则要求,更是保障公民知情权、监督行政权力与算法权力的重要手段。

在算法治理与治理算法交织的今天,亟需超越统一算法公开的狭隘视角,在探索算法技术应用与行政行为过程的统一可能性中,将算法应用的不同样态有机串联起来,聚焦不同自动化行政中不同算法的盲点并提出行之有效的公开方案,真正找到技术发展、行政效能与社会可接受程度之间的平衡路径。

参考文献

- [1] 郭琪. 自动化行政中基本权利保护探讨[J]. 上海政法学院学报(法治论丛), 2021, 36(2): 88-96.
- [2] (德) 里斯多夫·库克里克. 微粒社会: 数字化时代的社会

- 模式 [M]. 黄昆, 夏柯, 译. 北京: 中信出版社, 2018.
- [3] SOLOVE D J. The digital person: technology and privacy in the information age [M]. New York: New York University Press, 2004.
- [4] 张恩典. 人工智能算法决策对行政法治的挑战及制度因应 [J]. 行政法学研究, 2020 (4): 34-45.
- [5] 任蓉. 算法嵌入政府治理的风险及其防控 [J]. 电子政务, 2021 (7): 31-41.
- [6] (美) 劳伦斯·莱斯格. 代码 2.0 网络空间中的法律 [M]. 李旭, 沈伟伟, 译. 北京: 清华大学出版社, 2009.
- [7] 韩春晖. 自动化行政中算法目的失范的公法治理 [J]. 比较法研究, 2024 (2): 17-31.
- [8] (美) 卡里·科利亚亚斯. 自动化国家的行政法 [M]. 苏苗罕, 王梦菲, 译. 北京: 《法治社会》编辑部, 2022.
- [9] 查云飞. 算法的行政法属性及其规范 [J]. 法制与社会发展, 2023, 29 (6): 168-185.
- [10] 陈敏. 行政法总论 [M]. 台北: 新学林出版股份有限公司, 2019.
- [11] 于一帆. 自动化行政中算法的可公开性及其范围 [J]. 行政法学研究, 2024 (1): 159-167.
- [12] 周佑勇. 行政法基本原则研究 (第2版) [M]. 北京: 法律出版社, 2019.
- [13] 应松年, 朱维究. 行政法学总论 [M]. 北京: 工人出版社, 1985.
- [14] 姜明安. 行政法与行政诉讼法 (第7版) [M]. 北京: 北京大学出版社, 2019.
- [15] 康峻琿. 新形势下我国政府信息公开的实践探究 [J]. 管理观察, 2017 (4): 88-90.
- [16] 胡仙芝. 我国政务公开的现状、挑战与进路 [J]. 国家治理, 2021 (9): 40-44.
- [17] 胡小明. 从政府信息公开到政府数据开放 [J]. 电子政务, 2015 (1): 67-72.
- [18] 王贵. 算法行政的兴起、挑战及法治化调适 [J]. 电子政务, 2021 (7): 2-14.
- [19] 刘东亮. 技术性正当程序: 人工智能时代程序法和算法的双重变奏 [J]. 比较法研究, 2020 (5): 64-79.
- [20] 苏宇. 数字时代的技术性正当程序: 理论检视与制度构建 [J]. 法学研究, 2023, 45 (1): 91-107.
- [21] EDWARDS L, VEALE M. Enslaving the algorithm: from a "right to an explanation" to a "right to better decisions"? [J]. IEEE Security and Privacy, 2018, 16 (3): 46-54.
- [22] 王万华. 知情权与政府信息公开制度研究 [M]. 北京: 中国政法大学出版社, 2013.
- [23] 王锡铎. 数治与法治: 数字行政的法治约束 [J]. 中国人民大学学报, 2022, 36 (6): 17-34.
- [24] 展鹏贺. 数字化行政方式的权力正当性检视 [J]. 中国法学, 2021 (3): 114-138.
- [25] 王怀勇, 邓若翰. 算法行政: 现实挑战与法律应对 [J]. 行政法学研究, 2022 (4): 104-118.
- [26] 张凌寒. 算法评估制度如何在平台问责中发挥作用 [J]. 上海政法学院学报 (法治论丛), 2021, 36 (3): 46-49.
- [27] 李安. 算法影响评价: 算法规制的制度创新 [J]. 情报杂志, 2021, 40 (3): 146-152.
- [28] 郑智航. 平衡论视角下个人免受自动化决策的法律保护 [J]. 政法论坛, 2022 (4): 94-105.
- [29] 赵宏. 公共决策适用算法技术的规范分析与实体边界 [J]. 比较法研究, 2023 (2): 1-16.
- [30] 朱瑞. 论算法行政的技术性正当程序 [J]. 财经法学, 2023 (4): 103-117.
- [31] 靳雨露. 算法披露的域外经验与启示 [J]. 情报杂志, 2022, 41 (7): 91-99.

(收稿日期: 2024-11-19)

作者简介:

王昊禄 (1998-), 女, 硕士研究生, 主要研究方向: 行政法、数字法。

版权声明

凡《网络安全与数据治理》录用的文章，如作者没有关于汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权等版权的特殊声明，即视作该文章署名作者同意将该文章的汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权授予本刊，本刊有权授权本刊合作数据库、合作媒体等合作伙伴使用。同时，本刊支付的稿酬已包含上述使用的费用，特此声明。

《网络安全与数据治理》编辑部

www.pcachina.com