



自动化行政给付的算法风险 及其法律规制

梅 帅

摘要:算法与行政给付的深度融合是行政给付数字化转型的关键因素。自动化行政给付不仅重塑了申请受理的运行方式,也实现了给付事项的智能处理。然而,自动化行政给付的过程也带来算法歧视或偏见、算法权力、信息茧房等新型风险,影响自动化行政给付的效能发挥。因此,规制自动化行政给付的算法风险显得尤为必要。在宏观上,应明晰自动化行政给付算法风险的法律规制向度,更新分配正义理念,确立风险防范原则;在微观上,应完善风险防范的法律规范,增强数字包容的技术供给,形成多维给付数据的共享机制,建立算法偏见的合作监管机制,完善技术运用的法律责任机制,以健全自动化行政给付算法风险的法律规制体系。

关键词:自动化行政给付;算法风险;分配正义;法律规制

DOI: 10.13734/j.cnki.1000-5315.2025.0302

收稿日期:2024-10-09

基金项目:本文系四川省哲学社会科学规划项目“行政机关自我纠错制度研究”(SC24FZ023)的阶段性成果。

作者简介:梅帅,男,安徽六安人,法学博士,四川大学法学院助理研究员、讲师,E-mail: meishuaiscu@163.com。

随着数字政府建设的不断推进,以“秒批”、“免申即享”为典型代表的自动化行政给付场景不断涌现。政府运用自动化行政给付技术进行治理,为提升行政给付的精准度、优化行政给付的复杂流程、避免行政给付事项效率低下等提供了契机。数字技术的发展促使传统行政给付产生变革,数字技术作为内在驱动力,与行政给付关联领域相融合,形成自动化行政给付的新行为、新方式。但是,这也带来了算法歧视或偏见、算法权力以及信息茧房等新型风险。现有研究多从自动化行政给付的正当性抑或数字技术对行政给付带来的变革等方面展开^①,较少对自动化行政给付的算法机理、风险以及法律规制等展开探讨,导致了对自动化行政给付算法风险缺乏系统、全面的理性认识。算法机理是科学认识自动化行政给付算法风险的基础,本文将在厘定算法机理的基础上,以算法风险及其特征为切入点,探究算法风险法律规制的向度,并在此基础上提出完善算法风险法律规制架构的设想,以防范算法风险,保障相对人权益。

一 自动化行政给付的算法机理

(一)自动化行政给付的缘起及内涵

自动化行政给付的生成由来已久,早在电子政务的初期,政府通过自动化设备实施给付,实现政府信息服务的有效传递。当时,自动化行政给付虽冠以“自动化”一词,但自动化更多体现的是政务的电子化,而不涉及数据的智能分析,政府通过电子设备发送给付事项信息,由申请人接收并提出申请。例如,对于最低生活保障的发放,政府有效利用信息技术,通过不同的信息服务设施,按照预先设定的程序,为低收入家庭中的

^①曹梦娇《自动化给付的正当性基础及适用展开》,《浙江学刊》2024年第4期,第129页;黎慈《人工智能嵌入行政执法的法理分析:现状、风险与应对》,《湖北社会科学》2020年第7期,第135页;张梁、董茂云《“数字法治政府”:概念认知、机理阐释、路径塑造与机制构建》,《求实》2023年第5期,第28-29页。

特殊困难人员等提供自动化的信息及相关服务,以准确、快捷地传递信息,方便困难群众。随着电子政务的不断发展,许多事务由线下办理向线上业务转型,办公自动化成为政府履行政务服务的着力点,行政给付作为社会保障等部门的职能内容,被纳入对外的政务服务之中。与此同时,信息技术的快速发展促进了部门信息的共享、政务工作流程的整合及协同机制的完善,使更多部门在政务服务的同一条业务流程上开展工作,行政给付的环节大大缩减,行政给付的范围不断扩大。近年来,自动化行政给付的智能升级更快、运用更广,不仅重塑了政府的治理方式,更为人们的生产生活带来了实实在在的便利。例如,在住房保障领域,沈阳市政务服务 App 通过对职工基本信息及支出信息等完成自动审核,确定公积金提取的额度,实现完全丧失劳动能力等 10 种情形的“掌上办、零要件、秒到账”^①;在失业保险领域,云南省智慧服务系统自动筛选失业保险稳岗返还的企业,同时对比企业信用情况,公示无异议后,将资金打入企业对公账户^②;在医疗保障领域,安徽省医保服务系统筛选生成“预享受人员名单”,在慢特病群体、困难群体的医疗救助及生育津贴发放等方面打造“免申即享”新模式,实现服务事项的“零跑路”和资料的“零提交”^③。可见,自动化行政给付有了更多智能化、数字化的内容,随着数字技术与行政给付的融合程度不断深化,自动化行政给付的内在逻辑也由政务服务的流程化逐步发展为政务服务的自动化,审核模式从线上的人工服务更多地转向机器自主审核,处理时效从低效不断向高效发展,自动化行政给付得以形塑和生成。因此,自动化行政给付指的是以大数据、人工智能等技术为支撑,按照一定算法设计,由智能系统自动地作出给付决定或建议的行为。按照自动化决策程度的差异,自动化决策又可分为全自动化和半自动化,前者由智能系统直接作出给付决定,无须行政机关意志的介入,后者由智能系统作出给付决定的建议,行政机关可对此调整、修正或确认,并作出最终的给付决定。

(二) 自动化行政给付的算法机理探析

1. 法律规则的代码化是自动化行政给付的“转译”基础

法律规则是规范行政给付行为的依据,也是衡量行政给付是否在法律框架内的关键标准。算法是计算机可识别和执行的具体指令的集合,它将人的意图“翻译”给机器,是人和机器之间的“通约装置”^④。在自动化行政给付中,算法将法律规则转化为机器可识别和读取的代码,实现了部分法律规则从规范到代码的转变。智能系统通过算法对行政给付法律规则的转化,明晰了不同情形下的给付事项及其要求。进一步讲,智能系统将算法嵌入自动化行政给付的判断过程,把法律规则的语言转译成计算机语言,在代码运算、数据分析、自动处理等机制作用下形成相对确定的给付结果,以实现法律规则的有效“翻译”。从转化的内容来看,行政给付法律规则的代码化主要是从效果裁量方面着力的。这里涉及要件裁量和效果裁量的差异,要件裁量是行政机关按照法律文本上的要件解释不确定法律概念,使其具体化;而效果裁量是行政机关在法律要件补足后作出适当选择的自由^⑤。在智能系统中,绝大多数的算法将行政给付裁量基准翻译为计算机的执行语言,并将效果裁量与裁量基准相对应,使裁量效果格次化,进而实现给付裁量的自动处理。当然,也有少部分智能系统把要件裁量纳入判断范围,把既往的行政给付实践案例作为智能系统的数据读取基础,由算法对海量数据进行处理分析,作为不确定法律概念具体化的经验支撑。

2. 算法规则的适用是自动化行政给付的“个案”基础

从规范适用来看,确定个案事实和解释法律要件后,还须实现个案在规范中的逻辑演绎,以实现个案正义。在自动化行政给付中,智能系统首先分析案件基本情况,确定案件所涉及的给付类别,搜集涉案相关的给付数据,按照代码化的规则进行比对和裁量,进而得出行政给付决定或建议。例如,在失业保险领域,企业职工既往的工作信息、收入情况及待业时间等,构筑给付事项的“数据集”,经过代码化规则的适用,确定特定企业职工的给付标准与给付数额。当然,这也涉及案件情节等法律规则未明晰的部分,算法往往对既往行政

①徐佳婷《互联网+政务服务:让百姓生活更便捷》,《沈阳日报》2022年2月10日,第6版。

②刘佳华、方敏、叶传增《免申即享,让利企便民服务精准直达》,《人民日报》2024年7月23日,第11版。

③黄振君《“免申即享”“即申即享”助医保服务更亲民》,《池州日报》2024年7月26日,第3版。

④肖峰《认知的算法阐释:人工智能对当代认识论研究的启示》,《学术界》2021年第2期,第69页。

⑤王贵松《行政裁量的构造与审查》,中国人民大学出版社2016年版,第45—46页。

给付处理决定或建议进行比较分析,形成行政给付裁量的补充内容,作为确定特定行政给付数额的支撑。此外,智能系统的定位亦对算法设计提出差异化的要求。半自动化与全自动化差别较大。半自动化基于智能系统的有限介入,主要采取程序处理和人工判断相结合的方式,在个案数据输入后,经过代码化规则的个案适用,算法按照既定代码运算得出给付建议,行政机关再考量行政给付的个案性,作出行政给付的最终决定。全自动化则不依赖人工判断的环节,算法按照既定代码运算直接得出给付决定。在行政实践中,“秒批”、“免申即享”等智能系统更多体现的是全自动化行政给付算法的适用。行政给付的算法决定开始于相对人的意愿确认,智能系统对相对人给付事项的相关数据进行搜集、分析和处理,经过代码化规则的运算以及对既往处理情况的比对,智能系统直接作出行政给付决定。值得注意的是,为防止给付范围的泛化,人们有意识地设置给付领域以及相应场景下的具体情形,以约束行政给付行为的行使,防止自动化行政给付技术的滥用。

从技术表征来看,当前行政实践中自动化行政给付的算法设计,实现了生育、医疗、住房等多领域给付的整合以及社会救助、社会福利、社会保险等给付范围的扩充。自动化行政给付算法依托智能系统而展开,尽管各地的智能系统有所不同,自动化行政给付系统存在智能化或数字化融合程度上的差异,但在流程上主要体现为“申请人意愿确认—部门兑现”的方式,即申请人在智能系统上确认其对特定给付事项的意愿,部门借助智能系统审核受理,给付数额自动发放至申请人的账户。

二 自动化行政给付的算法风险及其特征

在数字时代,算法不再是简单的数据加工,而是凭借对海量数据的学习与处理能力而输出结果。在自动化行政给付中,算法在发挥主动给付、精准给付、即时给付等优势作用的同时,更是为行政给付最佳行政目标的实现提供助力。具体来说,自动化行政给付的算法运用,既化解了行政给付海量数据的搜集、贫困人口的分佈分析及成因发现等困难,又解决了行政给付申请流程复杂、事项办理效率低下、裁量结果不一致等问题。依托算法对海量数据的快速分析和高效处理,将原本由执法者依据法律作出给付决定的过程转变为算法自动作出决策的过程^①,把复杂的行政给付申请受理过程简易化,进一步推动了行政给付数字化流程的再造,降低了公民的程序性义务负担。这不仅有助于提升政府对公民给付诉求的回应能力,更促进了行政给付敏捷治理的实现。然而,自动化行政给付在带来诸多便利的同时,也产生了新的风险。

(一) 自动化行政给付的算法风险

与传统的行政给付不同,自动化行政给付的算法风险产生于算法设计以及对给付数据的整合、处理、分析等环节。在行政给付法律规范的落实中,算法可能对公平给付、自主给付、个体给付等产生不利影响。

1. 由算法设计和数据处理流程产生的歧视或偏见风险

在智能系统中,算法偏见贯穿于算法设计、数据处理等全过程^②,对公平给付产生不利影响。公平是一个模糊的、难以量化的概念,尽管计算机程序在语言和逻辑学习方面进步很大,但仍然难以准确理解此抽象概念^③。在技术上,算法是一种无涉价值的工具。然而,在实践中,算法的主导者将其自身立场和价值嵌入代码中,使得代码对给付数据产生区分、认可或排斥等不同效果。这种自身立场和价值可能是有意识的,也可能是无意识的。例如,在数据要素的设计环节,算法设计者基于对给付主体的经济、教育、健康等个人因素形成算法偏见,导致技术运用的非中立立场,将使得有关群体受到不公正的对待^④。在数据处理的过程中,经过算法系统对海量数据的读取,社会公众的信息被精准地识别,产生“数据画像”^⑤。算法歧视来自智能系统中的内隐驱动,个体部分特征或被突出或被忽视,算法歧视结果由此而逐步生成。例如,算法在读取数据或执行指令时,对部分特征诸如贫困、低收入、老年人等的记录,逐步产生区分性的结果,可能导致数字红利分配的不均衡加剧和显现。在此背景下,贫困人口、低收入群体等特定群体具备的数字服务机会存在缩减的

① 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,第37—38页。

② 岳平、苗越《社会治理:人工智能时代算法偏见的问题与规制》,《上海大学学报(社会科学版)》2021年第6期,第2—3页。

③ 凯西·奥尼尔《算法霸权:数学杀伤性武器的威胁》,马青玲译,中信出版集团2018年版,第103页。

④ 弗吉尼亚·尤班克斯《自动不平等——高科技如何锁定、管制和惩罚穷人》,李明倩译,商务印书馆2021年版,第3—5页。

⑤ 解志勇《数字法治政府构建的四个面向及其实现》,《比较法研究》2023年第1期,第12页。

风险,数字弱势群体的权益更易受到忽视^①。

2. 算法权力干扰自主给付的运行风险

行政给付的前提是政府对给付事项的依法自主决定。随着算法在智能系统中的运用,算法的本体特征逐步显现,它可能掠夺行政机关的判断权和决定权。算法原本是作为提升效率的工具而存在,其融入智能系统能够有效降低监管成本,但是,这也带来了治理的“算法化”问题,即算法的“专断”可能造成算法预设规则的普遍化,所产生的控制力在更深层次上重塑了个体与社会的关系问题^②。在自动化行政给付中,通过给付规则的代码化,所有给付情形和事项判断几乎都由代码运算,行政相对人只有满足算法所设置的形式要求,才能被有效识别。例如,在失业保险领域,企业职工既往的工作信息、收入情况及待业时间等,是算法识别和决策的关键基础,若其中数据存在缺失或不完整,算法规则可能认定为不完全满足要求,进而将其排除在给付对象之外。行政给付对象确认给付意愿后,算法自动搜集相关数据,形成给付判断,促进了自动化行政给付算法权力的建构。伴随着算法的自我迭代更新,算法可能背离设计者的预设,对相对人的隐私权、人身自由权等造成侵害,甚至出现技术资本操纵数字政府侵犯人权的风险^③。

3. 信息茧房强化同质给付的生成风险

信息茧房来自“蚕蛹吐丝结茧”的形象化表述,指的是算法习得用户行为数据和偏好来处理数据,强化预先的数据处理习惯^④。按照辅助性原则,行政给付的个体特征,在更深意义上来自国家对个人自主选择、决定的尊重^⑤,而基于算法对部分数据的处理而产生的信息茧房,可能弱化个体特征,推动同质给付的形成。智能系统的算法决策产生对已有观点的相同选择倾向,不断筛选和确定与之相同的判断逻辑,进而避免不一致结果的产生。依照代码规则立足的信赖基础,信息茧房可被分为特定观点的信息茧房和类型观点的信息茧房,前者为基于特定观点标准而形成与之一致的信息空间,后者为按照观点类型标准而形成与之一致的信息空间。在自动化行政给付的背景下,两种信息茧房逻辑意味着给付策略和行为模式的趋向一致。特别是在给付数据单一且有限的情形下,给付数据的不充分使得个体特征在算法规则的判断中逐步消弭,可能产生行政给付结果的趋同化现象。行政给付原来由行政机关综合考量申请人的实际情况而定,但是自动化行政给付的代码化处理,使得与特定观点或类型观点不一致的要素被忽视,进而导致“机械判断”、“结论趋同”等现象的产生。

(二) 自动化行政给付算法风险的特征

1. 复杂的交织性

自动化行政给付算法风险受技术发展水平、智能系统层级、算法适配场景等诸多因素的影响。基于各要素内容建构的智能系统模型的不同,自动化行政给付算法风险的交织性呈现出不同侧重。在自动化程度较低的行政给付中,算法风险可能交织于数据的适配性以及数据搜集、处理和分析的准确性等。在自动化程度较高的行政给付中,算法风险可能交织于代码规则的解释、算力本身的优化以及算法模型适配性的提升等。从具体体现来看,算法风险表现为多种形式,加之算法风险的滞后性,算法风险生成的逻辑变得更为复杂。尤其是在智能系统自主学习的过程中,算法对新数据、新语料的读取,产生与既往不一致的新解读,伴随着多维度神经单元反馈的调整,新旧风险的可能叠加促使算法风险分析逻辑更加复杂化。而且,在算法风险与行政给付领域的特殊性相结合的情况下,算法歧视或偏见、算法权力、信息茧房等风险关联耦合,带来一系列连锁反应,加剧了算法风险构成逻辑的复杂化,导致对相应风险的识别和应对的难度不断攀升。

2. 内在的隐蔽性

在根源上,算法风险产生于“算法黑箱”。算法黑箱建构一个难以观测、无从预知和不易理解的系统空间,包裹算法决策的内核,使得算法运转成为连接数据输入和数据输出之间的“黑盒子”。传统行政给付的申

① 曹梦娇《自动化给付的正当性基础及适用展开》,《浙江学刊》2024年第4期,第131页。

② 周辉《算法权力及其规制》,《法制与社会发展》2019年第6期,第118页。

③ 刘权《数字政府建设中数字化与法治化的融合》,《当代法学》2023年第6期,第19-20页。

④ 李龙飞、张国良《算法时代“信息茧房”效应生成机理与治理路径——基于信息生态理论视角》,《电子政务》2022年第9期,第55页。

⑤ 喻少如《论行政给付中的国家辅助性原则》,《暨南学报(哲学社会科学版)》2010年第6期,第58页。

请、审查、批准等按照法定程序展开,审查标准也严格遵循法定条件^①。而在自动化行政给付的智能系统中,推导的因果关系被打破,数据比对相关性得到提升,从数据输入到数据输出,人们都无法知晓黑箱运转中间阶段的逻辑,难以了解算法按照怎样的标准和逻辑行事,这使得自动化行政给付算法风险呈现隐蔽性的特征。尽管从外围角度看,自动化行政给付的设计者和服务者一般提供数据关联、共享和确认的指引手册,但是,指引手册不等同于算法决策及给付分析逻辑的公开。在实践中,行政相对人在表达申请意愿后,即收到给付决定的答复,难以知晓其数据输出结果的原因,无从评判和监督算法决策的科学性。即便“算法黑箱”存在异常,基于算法内部推理机制的不透明,行政相对人也难以感知给付的内在变化。

3. 发生的不确定性

从发生学来看,算法风险的发展、演变与转化规律具有不确定性,加之算法决策的复杂特征,算法风险发生的规模大小、影响范围和深度难以估测。除了算法引发的不确定风险之外,算法治理的不确定风险还包括决策主体和治理情境两个方面,前者基于外部因素和人为因素的复合,无法确定由谁对算法风险负责,后者为算法在执行特定任务时真实数据偏离初始数据,产生侵犯隐私等不确定的后果^②。在自动化行政给付中,这三方面的风险在一定程度上均存在。首先,算法风险的发生规律具有不确定性。在智能系统中,算法风险的发生交织于算法决策本身的复杂运行,部分风险的发生,人们难以提前预测。其次,在风险治理中,算法风险的发生具有不确定性。在日常监管中,智能系统若出现异常并引发权益侵害,究竟是属于系统开发者、设计者、运营者还是行政机关的责任,难以准确确定。再次,在智能系统运行中,侵犯隐私等算法风险的发生具有不确定性。例如,给付事项确认中的数据侵夺、隐私暴露等现象,可能带来对个人信息权、隐私权等权益的侵犯,算法的不利影响难以预知,发生结果的规模和范围更是难以衡量。

三 自动化行政给付算法风险的法律规制向度

自动化行政给付的参与主体较多且技术性较强,对制度规范要求较高。这决定了自动化行政给付算法风险的治理,不能仅施以规则与技术的简单叠加,而是要形成法律规制体系,明晰算法风险的法律规制进路,以防范算法风险的发生。为此,应从宏观层面明确算法风险的法律规制向度,结合算法风险的特殊性,更新分配正义理念,确立风险预防原则,并以此作为自动化行政给付算法风险法律规制的指引。

(一)更新分配正义的理念

规制理念是行政给付法律规制的逻辑起点。传统行政给付理念主要关注政府职责问题,如政府对孤儿、无人照料的病人、孤寡老人等无法自立群体的保障、物质权益或与物质相关权益的赋予以及政府给付任务的实现等^③。在数字政府建设的背景下,算法作为工具,涉及多元价值的实现及其平衡问题。在自动化行政给付中,算法不仅作为政府给付行为的基础,更融入了政府、企业、社会公众等主体的价值认识,因而,算法风险的法律规制须凝聚多主体共识,秉持分配正义的理念。而自动化行政给付算法风险的特殊性,要求技术运用既要保障分配正义的实体内容,又要保障分配正义的程序内容^④。

1. 自动化行政给付的实体分配正义

在数字社会,自动化行政给付所产生的算法风险具有不确定性,所带来的危害影响范围广、权益损害多且具有隐蔽性。由于多主体参与智能系统的设计、维护等环节,政府职责的履行和给付任务的负担逐步模糊化。随着传统的线下给付受理逐步转变为线上的给付处理,数据成为融合给付利益、给付标准、给付判断为一体的集合要素。给付的分配逻辑逐步以数据为中心,实体分配正义以公众需求为导向,政府以积极主动的姿态提供全方位、精细化、个性化的行政给付服务。在此背景下,算法融入自动化行政给付,通过技术变革与服务创新,为政府职能“做加法”以提升实效^⑤。反过来,实体分配的非正义逐步体现为信息不对称、算法权

^①尹建国、余睿《公共行政给付中的裁量权治理》,《环球法律评论》2010年第4期,第59页。

^②张海柱《算法治理中的不确定风险及其应对》,《科学学研究》2024年第9期,第1802—1803页。

^③胡敏洁《给付行政范畴的中国生成》,《中国法学》2013年第2期,第36—37页。

^④分配正义理念包含两个方面:一是关心何为平等与平等什么等实体内容,二是注重平等的程序性实现等程序内容。参见:汪行福《分配正义与社会保障》,上海财经大学出版社2003年版,第7—11页。

^⑤陈可翔《数字时代公共服务行政法构建的法理及进路》,《中外法学》2024年第3期,第630页。

力、算法歧视等,对行政相对人的权益产生损害。例如,基于身份差异而产生的数字歧视,将损害数字弱势群体的社会保障权等权益^①。而源于组织区隔、技术复合和法律保护隐匿,人们难以及时发现这种非正义性,从而使得有关风险不断累积。故而,应从实体层面关注自动化行政给付涉及群体的分配正义问题。例如,综合考虑给付公正、给付信赖利益保护、政府给付义务的设置,提升算法的透明度,推进算法公开和算法解释,注重算法在给付中的科学运用,考量不同的给付模型设计,提升算法与给付场景的适配性。

2. 自动化行政给付中的程序分配正义

传统行政给付要求政府承担法定义务,积极作出给付行为,并按照法定责任制度保障给付程序的公正。在自动化行政给付中,算法融入行政给付的诸多环节,在很大程度上冲击传统行政给付法律制度中的给付程序,对程序正义产生影响。例如,在相对人表达给付意愿后,算法对相对人给付相关数据进行搜集、整合、处理和分析,自动输出给付建议或给付决定。从给付的数据输入到自动输出,相对人无从知晓算法运算的中间内容,更对给付标准及数额难以预知。就程序法治而言,传统行政给付主要通过法定程序,增强相关主体监督,保障申请人的程序权利,以实现给付的程序公正^②。而在自动化行政给付中,算法融入给付过程,带来了给付事项算法治理的新课题、新挑战。按照程序分配正义的理念,算法需要经过各环节“程序”的锻造,以确保系统的科学性、合理性和有效性。智能系统经过开发、设计、维护等多个环节,每一环节的负责主体把关给付技术要求,以实现算法的技术环节控制。然而,在行政给付法律规则未对技术规制予以规范的情况下,算法的规制权限可能被滥用,进而影响程序正义的实现。例如,自动化行政给付算法的设计者,提供智能系统的使用说明及风险防范要求,这是否确定其已尽到算法合理设置的义务仍存在一定疑问。因此,需要从自动化行政给付算法运用的过程,也即从输入到输出的过程保证程序正义的实现。在给付数据的搜集阶段,需要通过程序设置保障相对人的知情权等权利,为防止设计主体对算法的恣意设计,应加强自动化行政给付的算法备案。在给付数据的运行阶段,应加强算法透明的进路。随着技术的发展,算法运用的告知设置、报备参数、源代码公开等,使得算法透明不仅可能而且可行^③。在自动化行政给付中,算法透明的实现并非一蹴而就,而是要考量给付目标、给付场景等差别化的透明度要求。例如,在医疗保障领域,基于对慢特病群体健康权、隐私权保障的设置,可采取基本信息部分隐匿的方式,以平衡算法透明价值与其他价值。在给付数据分析的输出阶段,既要保障自动化行政给付的结果公正,也要保障相对人的异议权、人工介入权等程序权利,以防止算法权力的滥用。

(二) 确立风险预防的原则

传统行政给付基于权益侵害的实际发生,按照“行为—责任”的归责逻辑,确定有关责任主体及其责任分担,责任承担是以结果为导向展开建构的,损害的现实发生为法律规制的前提条件。由此,法律规制围绕权力制约展开,旨在保障申请人在行政给付中的权益,并通过强有力的规制措施实现有关侵害的预防。对于申请人提出的给付诉求,行政机关依法受理,审核给付相关材料的真实性,并及时作出给付决定。若给付决定存在违法或不当的情况,司法机关采取事后的法律规制措施,监督行政权力的行使。然而,在自动化行政给付中,随着算法融入智能系统,风险的不确定性应运而生。从技术变革来看,算法对传统行政给付的事后规制模式提出了挑战。一方面,事后的给付权益侵害规制,难以应对自动化行政给付的算法风险。按照事后规制的逻辑,算法风险唯有不断累积形成现实的损害,才能纳入规制框架,这使得权益保护范围过窄,相对人权益难以受到及时保护。另一方面,事后的给付权益侵害规制,无法预防算法风险的再发生。无论是从算法风险发生的频率、强度来看,还是从司法机关的技术监管、损害认定等局限来看,事后规制模式在规制算法风险上存在诸多短板。为此,2021年6月通过的《中华人民共和国数据安全法》(以下简称《数据安全法》)对数据安全、风险监测等作出规定,为数据安全的风险预防奠定了基础。同年8月通过的《中华人民共和国个人信息保护法》(以下简称《个人信息保护法》)对预防侵害个人信息权益行为作出规定,兼有损害预防与风险预防

① 宋保振《社会权视阈下“数字弱势群体”权益保障》,《法学》2024年第1期,第22页。

② 崔卓兰、周隆基《社会管理创新与行政给付新发展》,《当代法学》2013年第1期,第59页。

③ 汪庆华《算法透明的多维维度和算法问责》,《比较法研究》2020年第6期,第165页。

的意蕴^①。为应对自动化行政给付的算法风险,应确立风险预防原则,明晰其对算法风险规制的指引作用,结合行政给付的特殊性及智能系统的场景,采取预防性的规制措施,以防止算法风险对相对人权益的损害。

在自动化行政给付算法风险的应对中,应科学审视算法驱动自动给付的过程,以事前规制来预防风险。具体来说,需要结合成本收益分析方法,统筹算法发展与风险防范措施,保障算法风险规制措施的合理性。可从技术合理使用、风险预期设置、预防措施的程度设置等展开,通过比较预期危害发生的概率与预防措施的负担,运用成本收益分析方法确定预防措施的选择^②。算法融入自动化行政给付,带来行政给付领域的技术革新,是数字政府建设的必然要求。然而,算法的不确定性与行政给付的特殊性交织,引发技术运用的不确定风险。因此,应将风险发生的不确定性考虑在法律规制的范围内,并通过持续跟踪的知识信息更新,不断调整优化规制措施,从而及时实现对风险的识别、处理和规制^③。

防范自动化行政给付的算法风险,应确立风险预防原则,并发挥其指引作用。自动化行政给付算法风险存在于给付数据搜集、处理及分析等环节,要确立技术运用的标准,防止技术的非理性化发展。风险预防原则的引入,旨在发现侵犯相对人行政给付权益的现象,并予以及时处理,以调整信息不对称、算法权力等引发的技术失衡格局,促进相对人给付权益保障的实质均衡。此外,风险预防措施也须把握其运行的边界。在自动化行政给付中,采取风险预防措施时,应按照比例原则中的必要性、适当性、损害最小等要求,划定预防措施的界限,以避免风险预防措施的滥用。

四 自动化行政给付算法风险的法律规制架构

面对自动化行政给付的算法风险,有学者认为,应框定自动化行政给付的边界,实现人与自动化设备的合理定位^④。也有学者认为,数字政府的建构不能简单套用平台经济的理论,对于自动化行政给付,政府应承担起必要的义务和责任^⑤。还有学者认为,行政给付的全自动化裁量应当坚持法律保留,通过法规或规章对给付裁量自动化的适用作细则规定,防范自动化技术的风险^⑥。现有研究多聚焦于自动化行政给付算法风险法律规制的某一方面,并未形成系统、完整的法律规制架构。基于此,笔者认为,自动化行政给付算法风险的法律规制需要充分考虑法律规范、技术供给、数据共享等方面的因素,结合分配正义理念和风险预防原则,从法律规范、技术供给、数据共享机制、合作监管机制和法律责任机制等五个方面展开。

(一)完善风险预防的法律规范

当前,我国已出台《数据安全法》、《个人信息保护法》等法律法规,对算法权力运行过程中的数据、平台等要素的安全保障作出规范。然而,鉴于行政给付法律规范散见于社会保险、社会救助、社会福利等具体领域,相关法律规范与行政给付领域的结合、衔接及落实存在一定不足,自动化行政给付的算法风险规制仍有诸多短板。为应对自动化行政给付的算法风险,按照风险预防原则的要求,应强化预防型法律规则的设计,将算法权力纳入法治轨道。首先,应完善自动化行政给付算法本体规制的规则。算法在改变行政给付运行逻辑的同时,带来算法本体规制的规范空白。这使得算法的运用处于“无法可依”的状态。对此,在现有法律规范的基础上,应不断完善行政给付的法律规范,尤其是算法本体规制的法律规范,把算法权力纳入法治轨道。具体而言,可推进行政给付法律规范的“数字转型”,树立预防为主的法律理念,明确算法的安全运营要求,提升算法规制的行业自律能力,促进行政给付与算法法律规范的衔接。其次,应建立自动化行政给付算法风险的评估机制。由于算法风险发生的不确定性,对其规制应采取事前预防而非事后制止的措施。《数据安全法》第二十二条确立的数据安全评估机制,为数据安全风险信息的获取、分享和研判等工作奠定了基础。在自动化行政给付中,可借鉴这条思路,围绕算法的风险隐患,建立行政给付领域的算法风险评估机制,对行政给付算法运行展开定性和定量的风险分析,科学划分算法风险的发生层级,把握算法风险的发展趋势,把事

①张涛《风险预防原则在个人信息保护中的适用与展开》,《现代法学》2023年第5期,第54—55页。

②苏宇《风险预防原则的结构化阐释》,《法学研究》2021年第1期,第43—44页。

③金自宁《科技不确定性与风险预防原则的制度化》,《中外法学》2022年第2期,第519页。

④王敬波《数字政府的发展与行政法治的回应》,《现代法学》2023年第5期,第117—118页。

⑤马长山《数字法治政府的机制再造》,《政治与法律》2022年第11期,第21页。

⑥刘星《行政裁量中的技术控制——基于政务服务应用场景的实践观察》,《公共行政评论》2022年第1期,第61页。

故发生的概率及损害降到最低^①。最后,应完善自动化行政给付算法风险的分层处理机制。对于自动化行政给付风险的定级,可在低、中、高基本层级的基础上,结合行政给付领域的特点,实现层级的精细化,形成自动化行政给付风险的定级标准。针对自动化行政给付的具体层级,在考量算法风险发生的概率、层级性以及紧急性等基础上,采取与之相适应的处理措施,如禁止算法对给付数据的获取、限制算法的部分运行权限等,以规制算法风险。

(二)增强数字包容的技术供给

公平分配要求行政机关平等地对待行政给付事项的申请人,依法考量申请人的给付诉求,尽可能地消除歧视,保障申请人给付权益。在自动化行政给付中,公平分配的实现并不是通过对技术的结果控制,而是要在技术的研发环节对技术的供给提出要求,以防范算法歧视等风险。为此,要增强数字包容的技术供给^②,通过包容性的技术,降低给付数据采集、处理和分析过程中的歧视,为算法应用提供具体的技术指引。数字包容最初主要用来化解数字鸿沟等问题,保障数字弱势群体权益。而后,数字包容的重要性不断提升,逐步发展为数字时代社会包容的概念范畴,并与数字排斥相对应。特别是域外发达国家关于数字包容的细化实践,将数字包容进一步分为数字准备、数字就绪、数字参与的基本框架,分别对应数字公平、数字信心、数字赋能等目标,这为我国数字技术的规制与理性发展提供了经验启示^③。2019年,国际电信联盟(ITU)对数字包容作出专门界定,并将其表述为“旨在确保所有人都有平等的机会和适当的技能,从广泛数字技术和系统中受益的策略”^④。这不仅为数字鸿沟的破解提供了治理指引,也为数字时代分配正义的实现提供了具体进路。在自动化行政给付中,应推动实现数字包容,尤其是关照数字弱势群体的给付需求,探讨算法包容的技术路径与治理方案。具体而言,政府要鼓励研发企业积极开展普惠性、包容性、无障碍的智能系统研发或升级工作,为避免算法歧视等风险提供技术支撑。政府在设计智能系统时,要最大程度地尊重相对人的给付需求,尤其是数字弱势群体的诉求,并将相关诉求与算法、智能系统功能相连接,进而优化自动化行政给付算法的技术供给,形成包容性技术研发运用的内循环,提升算法运用的人性化水平。

(三)形成多维给付数据的共享机制

给付数据作为算法的依托对象,在智能系统的处理中占据基础性地位。然而,数据处理容易诱发信息茧房等算法风险,在很大程度上限制了给付数据的价值发挥,促使智能系统的给付出现趋同乃至同质化的现象。故而,规制自动化行政给付算法风险的重要一环便是形成多维给付数据的共享机制,化解信息茧房等潜在风险。一方面,应从源头上丰富给付数据的来源。基于有限的给付数据,算法在读取或处理数据时容易产生信息茧房的风险。为此,应保障相对人对给付事项的知情权、参与权等,增强公众对自动化行政给付的了解,提升给付的多元参与。具体来说,可在智能系统运用区域开展宣传、培训活动,如制定自动化行政给付智能系统的使用手册等,提升公众对给付事项的参与能力。这种保障公众参与的方式,有助于提升给付数据来源的丰富度,避免因给付数据不足而诱发信息茧房风险的发生。另一方面,应整合给付数据资源,完善多部门给付数据的共享机制。目前,行政给付的数据共享更多地基于单一部门与其他部门的联合,数据共享程度较低。为有效降低信息茧房风险的发生概率,应加强给付事项的数据利用,建立自动化行政给付事项的数据共享平台,作为算法处理给付事项的基础^⑤。具体而言,可将给付职能框定在数据库中,构建给付行政整体的公共服务职能,在厘定多种给付职能的基础上,形成给付行政的具体清单,由给付事项的负责部门提供基本数据,给付事项涉及单位提供相关数据,实现给付事项数据的共治^⑥。在数据共享中,给付数据可重复使用,并作为其他部门处理给付事项的基础。可确立统一的给付标准,对算法实施给付的数据进行分类处理,

①王素芬、付浩然《数字时代的社会救助立法检视》,《理论学刊》2023年第2期,第87页。

②数字包容是针对数字技术利维坦问题提出的新路径,也是从数字排斥现象中引申的新概念,其肇始于20世纪90年代,而后发展于21世纪初。参见:沈费伟、胡紫依《建构数字包容体系:实现乡村数字弱势群体权利保障的策略选择》,《社会主义研究》2024年第1期,第116页。

③杨巧云、梁诗露、杨丹《数字包容:发达国家的实践探索与经验借鉴》,《情报理论与实践》2022年第3期,第195—196页。

④匡亚林、蒋子恒《政府负责任创新下数字包容型老龄社会构建研究》,《电子政务》2024年第5期,第34页。

⑤匡亚林、蒋子恒、张帆《数字技术赋能社会救助:缘起、风险及治理》,《上海行政学院学报》2023年第2期,第95—96页。

⑥关保英《数字化之下的给付行政研究》,《法律科学(西北政法学报)》2022年第6期,第75页。

对不同给付事项中的基础数据、代码标准、算法程序等展开标准化的处理,形成自动化行政给付算法的技术平台,方便负责不同给付事项的部门自主利用。当然,在形成数据共享机制时,须注重数据共享的法定界限,应在严格落实《数据安全法》等法律规范的基础上,平衡技术使用与个人信息权益的保障。例如,在算法处理给付数据后,应建构相应的安全保障机制,保障给付数据的安全性,以防止算法侵害个人信息情况的发生。

(四)建立算法偏见的合作监管机制

作为自动化行政给付算法治理的理念,分配正义是规制算法风险、保障数据安全和实现技术理性发展的指引。在应对算法偏见风险时,应关注分配正义理念的制度化,建立算法风险的预防机制,结合自动化行政给付算法的实际情况,采取合作监管机制,实现数字时代行政给付的公正目标。首先,应加强对自动化行政给付算法的行政监管。面对自动化行政的算法决策风险,行政机关应进阶对算法本身进行规制,以提升算法治理的能力^①。具体而言,对于线上监管,行政机关应建立自动化行政给付算法监管平台,完善算法风险的定期检查和动态追踪机制,以应对算法应用所带来的偏见。为保证算法风险监测的常态化,行政机关应建立算法应用的监管机构,设计科学的监管规则,监督算法的实际运行。例如,在事前,政府可采取算法备案监管的方式,要求委托的算法设计或研发单位将算法的技术信息向监管机构备案^②;在事中,政府可结合算法监管平台,监测算法风险;在事后,政府可依据法律规范对算法主体的实践合规性进行审计,并围绕公平性、透明度、安全性指标对算法输出结果的风险展开审计,进而衡量算法系统的社会嵌入影响以及嵌入过程的合法性^③。其次,政府应引导企业、行业协会、科研机构等参与算法规制的过程。在行政监管之外,为促进多主体治理的协同化,应引导包含设计方、第三方在内的主体参与算法风险的协作规制,以应对算法应用带来的风险。企业作为自动化行政给付算法的研发主体或维护主体,其参与算法规制的监督过程,有助于及时了解社会公众的实际所需,理解行政机关对算法监管的实际要求,促使其增强自律意识,提升算法决策服务的质量。行业协会一般由算法、大数据、人工智能等领域的专业人士组成,由其对算法风险规制提出专业意见,有助于克服算法风险规制专业知识和应用经验不足的短板,增强对算法风险的监测能力。另外,专家对自动化行政给付算法的监督,有助于提升算法规制的专业性,补充算法风险跨领域性、前沿性发展等规制的不足,形成算法规制外部监督的有效力量。再次,应促进社会监督与政府监管的协同联动。行政给付是典型的惠益性行政方式,按照法律法规的规定,公民在失业、年老、疾病或遭受自然灾害等特殊情况下,可向行政机关提出申请,以获得一定物质性或物质性相关的利益^④。相较于商业领域的算法设计,公共事业领域的算法设计嵌入社会的程度更深,且与个人经历、内容场景等有很强的关联性^⑤,这就要求强化相对人的参与式治理,帮助监管方及时发现和化解自动化行政给付算法风险,以形成监管的合力应对^⑥。在权利设置上,应赋予相对人在自动化处理中获得有效通知和发表意见的权利,如设置必要的人工介入等,以对抗“自动化偏见”^⑦。

(五)完善技术运用的法律责任机制

法律责任的分担对于规避算法虚化、算法逃逸等风险意义重大。在数字政府的视域下,需要完善自动化行政给付的责任机制,明确政府、算法设计者、算法维护者等主体责任。对于政府而言,其责任是行政给付过程中的法定责任。自动化行政给付的算法设计由行政机关设定规则,并委托给特定企业或机构研发,算法行为的责任主体仍为行政机关。因而,应强化行政机关在算法设计和运用中的责任设置。例如,针对政务数据系统的算法黑箱、算法解释不足的情形,可分别赋予行政机关提升透明度的责任、作出相应解释的责任,以防止算法的技术表象对政府责任的掩盖^⑧。在类型上,政府责任可分为监管责任、担保责任等。监管责任要求行政机关履行监管职责,对算法模型的设计、测试和应用等各阶段进行监管,提升算法的透明度和解释性,防

①张凌寒《算法规制的迭代与革新》,《法学论坛》2019年第2期,第23页。

②张吉豫《论算法备案制度》,《东方法学》2023年第2期,第92页。

③张欣、宋雨鑫《算法审计的制度逻辑和本土化构建》,《郑州大学学报(哲学社会科学版)》2022年第6期,第35页。

④江国华《中国行政法(总论)》,武汉大学出版社2017年版,第233页。

⑤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,第1441—1442页。

⑥张欣《生成式人工智能的算法治理挑战与治理型监管》,《现代法学》2023年第3期,第119页。

⑦喻少如、雷刚《精准扶贫中个人信息的利用及其边界》,《北京理工大学学报(社会科学版)》2022年第1期,第149页。

⑧徐继敏《数字时代行政法的发展》,《行政法学研究》2024年第3期,第92页。

止申请人的数据被侵夺和隐私被泄露,保障算法给付公正性、效率性的实现。担保责任要求在算法设计者等承担部分或全部责任时,政府承担担保其实际履行的责任。在民营化的合作治理中,政府承担的担保责任主要包括挑选合格的合作伙伴、履行协议、国家赔偿以及合作难以为继的承接责任^①。算法设计者的责任来源于对政府委托设计要求的实现以及算法设计中偏差、错误等避免的合同责任。当然,对其设计的算法在实践中引发权益侵害,也应承担侵权责任。算法维护者的责任同样来源于维护协议的规定,此时算法维护者是否应当承担责任,应取决于其对算法运作是否产生实质影响。例如,在算法失灵情况下,由其参与调整算法逻辑的运行导致的侵权行为,应在行政机关委托维护协议的范围内承担责任。

在责任分担中,需要确定侵权主体、侵权行为、因果关系和责任认定。可运用算法记录留存机制和算法影响评估机制来判断相关主体责任承担及比例。算法记录留存机制通过确定算法运行的轨迹,回溯侵权行为为相关时间段的行为过程,记录申请人给付意愿确认后算法系统的行为活动,以判断在此期间是否存在算法维护者对系统的调整、参数的修改等。算法影响评估机制的运行,主要来源于第三方的算法评估报告,行政机关借此判断算法侵权行为发生后的影响范围及程度等。在侵权行为发生后,行政机关可基于侵权行为发生后相近时间点的评估报告,把握算法侵权的风险点及成因,采取及时补救措施。在责任追究方面,前述机制同样可作为责任归属和后续处理的基础。算法记录留存机制可还原侵权行为发生的情况。算法影响评估机制可作为判断侵权行为的影响及侵权扩大前后的责任范围。反过来,若相关主体违反前述义务的要求,其责任应由哪个主体承担及承担多少将更为明晰。在侵权行为与责任承担之间,还存在过错责任、无过错责任或过错推定责任的归责以及归责方案的细化,鉴于行政给付的公共属性,应以无过错责任作为行政机关、算法设计者、算法维护者等主体责任的要求,以强化其义务履行要求。当然,在智能算法系统的运行中,各主体对外承担连带责任,对内承担相应责任,并允许责任主体之间的相互追偿^②。

五 结语

随着数字技术的快速发展,算法与行政给付的深度融合不仅增强了政府服务的回应能力,也提升了行政给付的效能。但是,这也带来了算法歧视或偏见、算法权力、信息茧房等新型风险。为规制自动化行政给付算法的风险,我们需要更新分配正义理念,确立风险预防原则,并将其贯穿于自动化行政给付算法运用的全过程,厘定算法风险的法律规制向度。在规制架构上,应完善风险预防的法律规范,增强数字包容的技术供给,形成多维给付数据的共享机制,建立算法偏见的合作监管机制,完善技术运用的法律责任机制,以健全自动化行政给付算法风险的法律规制体系。自动化行政给付算法是给付法律规范实施向计算机运用的转型,本质上是实施手段的更新变化。其中,更为重要的是从算法给付向依法给付的回归。

[责任编辑:苏雪梅]

^①张敏《数字政府建设政企合作法律风险的规制模式》,《现代法学》2024年第3期,第187—188页。

^②肖梦黎《算法行政责任的分布式重建》,《国家检察官学院学报》2023年第2期,第52页。