

算法风险国家预防义务的证成及展开*

洪丹娜 黄凯琪

(华南理工大学法学院, 广东 广州 510006)

[摘要]我国算法风险预防体系呈现出规范依据分布“碎片化”、行政规制主体多头分散、规制对象无差别化等特点,并由此衍生出制度衔接困难、法律依据空白、监管真空或重复、规制过度或不足等困境。究其根源,算法规制非预防手段在责任认定和风险预防效果上具有明显的局限性,现有个别化的制度因应对算法风险的系统性、潜在的危害对象、影响范围、危害可挽回程度等因素考量不足,加之“权力—权利”结构进一步失衡可能使算法应用逃逸现有的规制轨道。因此,亟待确立算法风险的国家预防义务,从宪法层面回应算法技术应用的非理性结果,为算法风险预防提供统一秩序引导和宪法依据,进而搭建系统性的算法规制框架,优化既有法律体系。基于保障人权和维护社会秩序的目的,算法风险国家预防义务的宪法基础根植于基本权利条款以及宪法序言、宪法总纲中的国家任务或者国家目标条款。算法风险国家预防义务是一种工具性宪法机制,兼具主观法和客观法的双重属性,其具体内容因所依据的宪法规范不同而存在差异,需要基于禁止不足原则和禁止过度原则确定义务的动态边界,通过立法完善、行政监管、司法审查和多元监督等路径协同落实。

[关键词]算法风险预防 国家义务 基本权利保障

[中图分类号] D921

[文献标识码] A

[文章编号] 2096-983X(2025)04-0094-12

一、引言

伴随着算法技术应用日益普遍化,算法歧视、算法黑箱、算法共谋、信息茧房等问题日益凸显,暴露出算法应用存在的非理性结果和对社会秩序构成的潜在威胁。算法风险已成为现代复杂社会系统不可避免的重要挑战,法律体系应适时回应。《关于加强互联网信息服务算法综合治理的指导意见》(下称《意见》)确立算法治理的风险预防原则,《个人信息保护法》《互联网信息服务算法推荐管理规定》等法律

文件构建起算法风险预防的制度框架。目前,风险预防已成为我国算法规制实践的原则指引,也有部分学者针对特定算法风险预防手段或者具体领域的算法事前规制路径进行研究。^[1-2]然而,我国初步建立的算法风险预防体系仍存在诸多不足,当前学界相关研究也具有一定局限性。第一,算法风险预防相关的现有研究多聚焦于论证必要性或者基于特定算法风险展开,尚未形成对算法风险预防整体正当性的系统论证。第二,对算法风险预防的质疑虽未全盘否定事前规制的价值,但其指出的算

收稿日期: 2025-04-10; 修回日期: 2025-05-13

*基金项目: 国家社会科学基金项目“算法时代数字弱势群体的人权保障研究”(22BFX199)

作者简介: 洪丹娜, 法学博士, 副教授, 主要从事宪法学与行政法学研究; 黄凯琪, 硕士研究生, 主要从事宪法学与行政法学研究。

法风险预防侵蚀私法自治的可能性和不具有可行性等局限性,客观上对我国现有算法事前规制制度及其实践的正当性构成了挑战。^[3]第三,我国宪法学界在整体上关注的风险预防议题,主要集中在风险社会与科技风险的宪法应对机制以及国家在风险社会中的整体性预防义务等方面,未针对算法风险展开探讨。^[4-7]此外,也有部分研究关注算法应用中的人的尊严与人权保障问题。^[8-10]这些研究为国家规制算法应用行为和风险社会下的风险预防提供了正当性基础,但整体风险预防的宪法正当性无法推导出算法领域风险预防的宪法正当性,^[11]通过风险预防手段规制算法的行为仍有待证成。

现有研究尚未系统探讨国家在整体层面预防算法风险的宪法基础,导致实践中算法风险预防措施缺乏充分的宪法依据,相关制度困境也亟待宪法层面的回应。鉴于此,本文首先检视我国算法风险预防体系,指出确立算法风险国家预防义务的现实基础,通过规范分析阐明该义务的宪法基础并实现该义务的证成,进而围绕预防义务定位、基本内容和义务实现进一步展开,为我国算法风险预防体系的完善提供理论支撑与构建方案。

二、我国算法风险预防体系的检视

随着算法问题日益凸显,我国目前已将算法风险纳入法治范畴,逐步建立起算法评估制度、算法备案制度以及算法风险监测和监管制度等一系列算法风险预防制度。这些制度以行政性风险规制为主,形成了我国算法风险预防的体系框架,并在实践中有序推行,对规范我国算法服务市场秩序、预防算法风险产生积极作用。但也应当认识到,当前的算法风险预防体系仍处于发展初期,其规范依据、规制主体和规制对象等三个维度呈现出鲜明的制度特征,并由此衍生出若干亟待解决的预防困境。

(一)规范依据分布“碎片化”及其预防困境

我国算法风险预防的规范依据整体分布情况呈现“碎片化”特征,主要体现在横向场景规制和纵向立法层级两个维度。从横向上看,相关规范分散于不同算法应用场景对应的法律中,即算法风险预防规范通过个别条款分别嵌入《个人信息保护法》《电子商务法》《反垄断法》等法律中。场景规制的思路本质上是回应型立法的体现,能够适应算法与不同应用场景结合所产生的不同风险规制需求,有效回应现实中已经凸显的算法问题。^[12]然而,场景化规制也加剧了制度间的衔接难度,难以形成制度合力进而对算法风险形成体系化的约束,也可能导致部分场景中的算法风险预防缺乏上位法的支撑。^[13]从纵向上看,算法风险预防相关规范分布于法律、法规和部门规章等不同层级的法律文件中,但以位阶较低的规章和其他规范性文件为主。一方面,不可否认,考虑到我国现行算法技术发展不确定和算法治理领域立法经验不充分等,我国以较低效力等级的法律文件作出算法专项规定能够将立法风险灵活限制在可控范围内,是当前应对算法风险不确定性的合理选择。^[14]然而,相关规定以算法技术类型为标准,将不同技术类型的算法风险规制规范分别规定在这些部门规章中,算法风险预防规范在整体上仍呈现“碎片化”分布,且无法覆盖其他类型的算法技术。另一方面,在法律层面基于算法可能侵犯的具体利益将算法风险预防融入数据安全保护和网络安全保护等领域,有利于快速响应算法风险规制的重点需求,但却不利于构建统一的算法风险预防框架。总体而言,不同层级的算法风险预防规范都在不同程度呈现“碎片化”分布的特征,增加不同层级规范间的衔接难度,可能在现实中模糊算法合规的标准。

此外,集中于部门规章的算法风险预防规定由于其载体效力等级较低,也存在一定的局限性。第一,部门规章的制定过程缺乏公众参与,容易以技术为单一导向而忽视多方利益平衡,导致规则与实际需求的脱节。^[15]第二,受法

律保留原则限制,部门规章无法创设新的法律责任形态,手段上局限于传统行政监管工具,缺乏应对新型风险的技术治理手段。^[16]第三,有关算法专项治理的部门规章主要调整行政法律关系,忽略其他法律关系中的算法风险调控。由此可见,长久的算法风险预防仍需要在效力等级更高的立法层面进行统一立法。

(二) 行政规制主体多头分散及其预防困境

面对算法风险的威胁,政府需要基于行政管理职能坚守执行机关和行政机关的机构定位,成为算法风险预防制度的主要执行者和算法风险的主要规制主体。事实上,根据各部门的职权范围和相关法律法规规定,网信部门负责统筹协调我国整体的算法风险预防工作,其他有关部门则在各自职责范围内防控算法风险,我国形成“网信部门+有关部门”的算法风险规制主体格局。^①算法嵌入式分布的特点使得算法风险分布于社会各个领域,而这样的算法风险规制主体格局能够通过网信部门主导、有关部门配合的形式在统筹不同领域资源的情况下实现对不同领域算法风险的防控,提高算法风险预防的效率。但是,算法风险由算法系统与社会子系统交互形成,其不仅涉及不同领域的风险,还包含技术风险、伦理风险、法律风险和异化风险等不同性质的风险。此时,即便有网信部门的统筹协调,以职权为分工标准的多头治理也可能在交叉领域的算法风险预防工作中出现职权重叠、职责不清、监管责任分散或者监管责任推诿的情况,进而产生监管真空或重复等问题。

(三) 规制对象无差别化及其预防困境

根据相关算法风险预防规定,我国算法风险预防所规制的对象主要以是否涉及算法的应用这一标准来确认范围。此标准没有将公权力机关排除在算法风险规制对象之外,算法风险预防的相关规则基本能够无差别化适用于公私

法主体。然而,普遍认为公法主体和私法主体之间存在本质上的异质性,无差别适用算法风险预防相关规则很有可能导致公法领域的算法风险预防不足或者私领域算法风险预防过度。^[17]事实上,从网信部门公开的算法备案清单可见,以算法备案为代表的算法风险预防制度主要适用于私主体,公权力机关的算法应用在实践中受约束的范围有限。我国算法风险预防制度在应然和实然层面的规制对象存在偏差,实践中公权力机关算法应用所产生的风险并未得到充分重视和有效防控。

虽然算法评估、算法备案等算法风险预防制度以“具有舆论属性或者社会动员能力”为标准要求相关算法主体履行强制性算法风险预防义务,能够差异化规制此标准内和标准外的算法主体,但是这一落实分级管理理念的做法,实际上并未进一步区分规制对象“舆论属性或者社会动员能力”的强弱,而是无差别地适用于所有具有此属性或能力的算法服务提供者。^②这同样可能导致舆论属性或者社会动员能力较强主体受到的约束不足或者较弱主体受到过多约束。不仅如此,这种无差别的规制也难以有效衔接算法、个人信息和数据分级分类制度。

三、算法风险国家预防义务确立的现实基础

算法风险预防困境的根源在于缺乏统一的价值指引和规则体系,亟待宪法层面确立算法风险国家预防义务予以纾解。算法自身的技术特征,使得以个人责任为主的事后救济手段难以应对算法技术带来的社会问题。^[12]算法风险作为现代风险的新表现形式可能带来极端危害以及公民无力抵抗算法权力的现实,产生了在宪法层面确立相应国家义务对其予以回应的现实需要。

①参见《互联网信息服务算法推荐管理规定》第3条、《互联网信息服务深度合成管理规定》第3条和《生成式人工智能服务管理暂行办法》第16条。

②参见《具有舆论属性或社会动员能力的互联网信息服务安全评估规定》第2条。

（一）算法风险预防困境需寻求根本法层面的回应

算法风险预防制度的构建和实施本质上是国家权力的行使过程和结果。当前我国由算法风险预防制度规范碎片化、主体分散化和对象无差别化等特征带来的预防困境，事实上是算法风险预防系统化、精细化缺失的体现，根源在于国家算法风险预防缺乏宪法层面的阐释，无法在整体层面对国家权力预防算法风险行为形成统一的价值指引和规则体系。而在宪法层面确立算法风险国家预防义务的首要意义在于，通过宪法作为根本法的共识性和权威性，构建统一的算法风险预防制度框架，增强算法风险预防体系的整体性和协调性。第一，宪法所蕴含的价值共识和共性要求能够为算法风险预防提供明确导向，形成稳定的价值预期，进而规范具体预防行为、增强预防体系的稳定性。^[18]第二，算法风险国家预防义务的确立代表着算法风险预防不仅是国家的政治决定，还是国家需要承担相应法律责任的法律义务，涵盖立法、行政、司法和监督等方面。尤其是立法义务的推进，有利于填补当前法律体系中算法风险规制的空白，加快算法领域专项立法进程。第三，算法风险国家预防义务为国家所有具体算法风险预防提供正当性来源和宪法依据，对整个预防体系具有统筹作用。宪法层面的讨论，不仅能够在价值和底线层面统一算法风险预防，还可以实现国家资源和国家权力的统合和整体调度。^[19]

（二）算法规制非预防手段之局限呼唤风险预防

风险社会中，具有相对稳定性、面向确定性的法律规范仍主要以事后救济为主，但立法者基于特定领域和场景的风险预防需求在立法技术上作了相应调整，例如在环境法中立法者将

某些环境行为或状态拟制为违法行为或危险状态，以预防环境污染风险。在现有法律体系中，仍可以对算法技术应用所带来的潜在危害采用事后救济的方式，在损害发生后再对相应个人或者社会秩序进行救济和修复。然而，抛开所有规制手段都面临的成本问题，事后救济在算法问题上存在责任认定的局限。首先，算法黑箱构成责任认定的重大阻碍。虽然专家辅助人等制度可以缓解公众基于认知局限的黑箱现象，但是技术复杂化趋势以及算法主体利益保护制度的存在，使得认知意义上和人为故意的算法黑箱对取证、举证和因果关系认定造成严重阻碍。^①算法损害行为事后维权和事后惩戒之困难，将在很大程度上减损算法领域事后救济手段的威慑性。其次，在法律未规定特殊归责原则的情况下，机器学习算法的动态演进特性使传统归责原则面临挑战，特别是深度学习系统通过数据训练产生的非预设性输出，增加了行为与结果之间因果关系的认定难度。^[20]最后，在算法普遍分布且日渐融入人们日常的情况下，事后救济的局限性还体现在，其只能通过个人责任的追究对已经发生的损害及相关个人进行救济，而不能阻止同类算法应用行为对他人造成同类损害的发生，预防作用有限。^[21]这些局限在算法技术快速迭代、深度融入社会生活的背景下尤为凸显，单纯依靠事后救济难以有效应对算法问题。

确定性要求高于风险预防的危险消除手段虽然更符合法律的规范性要求，但是并不足以涵盖风险预防的需求，同样无法取代预防手段对算法风险的防控作用。第一，算法风险具有突发性和系统性，算法风险预防多基于非线性因果关系，与危险消除遵循的“传统（机械）物理学的因果和线性时间观”冲突。^{[22](P148)}第二，

^① 认知意义上的算法黑箱指的是，基于算法技术性、复杂性、动态性等自身技术特性所造成人们难以理解算法或者算法难以透明的问题；人为主观造成的算法黑箱则是指由于算法设计者、算法使用者等主体基于商业秘密保护、算法安全等原因不公开算法，造成人们难以知悉算法内部运行的状况。参见赵泽睿：《论证程序的意义——对法律规制算法的另一种思考》，载《中国政法大学学报》2023年第1期；谭九生、范晓韵：《算法“黑箱”的成因、风险及其治理》，载《湖南科技大学学报(社会科学版)》2020年第6期。

算法的技术性、复杂性和动态性在困扰事后救济的同时,也阻碍着人们发现和证明危险的存在,以个体提起诉讼的方式主张危险消除存在一定难度。第三,算法的高速运算特性使得损害常在危险被发现前就已发生,因而算法运行过程的消除危险无从谈起。事实上,多数的算法潜在危害恰恰是通过风险预防制度的识别才得以避免发展成具有紧迫性的危险和现实的损害,例如在算法歧视风险的预防中,算法在投入使用前的影响评估是发现歧视性因素的主要手段。

(三) 算法风险的危害性亟需确立相应的国家义务

算法风险作为现代风险的表现形式,可能减损公民权利、侵扰社会子系统的正常秩序和损害国家利益。其一,算法应用存在减损人的尊严之虞。嵌入式部署特点和隐蔽的运行方式使得算法能够在庞大数据、强大算力和极具传播力的互联网加持下,“以前所未有的速度、范围和规模”悄无声息地实现对人们行为的规训,进而侵蚀人的自主性。^[23]当算法基于用户画像实施各类算法歧视等非理性行为时,人即成为算法权力的客体,其尊严也随之减损。^[10]其二,公共领域的算法应用也可能产生溢出效应。例如,2023年巴以冲突中“福音”人工智能系统造成平民伤亡激增以及剑桥分析公司算法操纵选举等案例,表明算法已经威胁到公民生命权、健康权和国家安全。2010年美国股市由于高频交易算法导致的“闪崩”事件,也早已表明算法风险对社会子系统正常秩序的威胁。^[24]此外,算法风险的危害性还体现为影响范围的广泛性和危害的难以挽回。一方面,算法与互联网技术相结合,能够轻松突破空间的限制,在更大范围内、以更快的速度造成危害。另一方面,以隐蔽的方式构成其应用领域行为模式且发挥规训作用的算法,若具备某一缺陷很大程度将引发结构性困境(如算法歧视),并且其所侵害的内容往往以非物质性危害为主,难以修复和弥补。

在算法普遍化应用背景下,算法风险危害

的极端性成为国家在宪法层面承担该风险预防义务的理由之一。然而,增设国家义务会增加财政压力,为国家权力介入私域提供更多空间,国家承担算法风险预防义务需要更为根本的正当理由。究其根本,国家通过履行算法风险预防义务调整和介入应用算法的自由,本质上源于公民无力抵抗算法赋权下权力主体的侵害行为。第一,算法行政能够延伸国家权力的管理能力,扩大国家监控的范围和深度,进而加剧国家作为庞大的权力机器与公民之间存在的结构性、对抗性的“权力—权利”张力,分散的公民个体难以对抗集中行使的国家权力。第二,私主体在应用算法的过程中通过算法赋权、赋能获得了技术权力和资本权力交织形成的私权力,^[25]而互联网平台则基于其监管义务拥有了“监督、定义和影响”相关主体的权力。^[26]原本“公权力—私权利”的二元格局逐渐向“公权力—私权力—私权利”的三元格局转变。在此格局下,私主体间的法律关系可能突破“平等主体”间意思自治、契约自由的假设,形成私权力主体和私权利主体的不平等关系,国家进而具有了介入私主体间法律关系的空间,公法性质的宪法规范也有了及于私主体的余地。^[27]综上,算法风险的危害性为宪法层面讨论算法风险预防议题提供了现实必要。

四、算法风险国家预防义务的宪法基础

在宪法规范层面,有关国家风险预防的直接体现是第26条环境保护条款,而其他领域国家风险预防义务的宪法基础则需要对相关条款进一步解释。算法技术的应用威胁着人权、社会秩序和国家利益,确立算法风险国家预防义务实际上是为了保障人权、维护社会秩序和国家利益,因而可以从相关规定中寻找算法风险国家预防义务的宪法基础。

(一) 基本权利条款与算法风险国家预防义务

在“基本权利—国家义务”框架下,算法技

术的应用对公民基本权利具有损害可能性,国家出于基本权利保障的目的,以基本权利条款为规范基础,承担算法风险预防义务。^[28]

1. 概括性人权条款与算法风险国家预防义务

《宪法》第33条第3款规定了“国家尊重和保障人权”,即概括性人权条款。此条款本质上是基于人民主权原则的义务性条款,为国家设定了宪法义务,包括国家不侵犯人权的尊重义务和国家积极履行或者提供人权实现条件的保障义务。^[29]基于此,国家在面对算法风险时,需要承担两大类型的义务:一是国家作为算法应用者,应当在算法应用过程中避免或减少对公民基本权利的侵害;二是国家作为算法规制主体,应当为算法应用已造成或者可能发生的损害提供相应的保障,为算法权利的实现提供应有条件。

首先,国家的尊重义务不宜作为算法风险预防义务的规范基础。智慧政府、智慧司法、智慧检务等进程的推进,意味着国家已然成为算法的应用者,并且和其他算法应用主体一样存在侵害公民权利的可能。尊重义务要求国家在算法应用过程中避免或减少对公民基本权利的侵害,而防止国家不履行尊重义务、利用算法侵害公民基本权利的制度设计,在一定程度上会发挥预防作用。但是,这种预防作用并非由国家履行尊重义务直接产生,本质上仍是国家积极义务的履行结果。更何况,“预防”一词本身具有事前做好应对措施的意涵,这一积极面向使得要求国家不作为的尊重义务与要求国家积极作为的风险预防义务相对立,故不宜将尊重义务认定为算法风险预防义务的规范基础。其次,算法风险国家预防义务应当立足于具有积极面向的人权保障义务。^①一方面,防止国家不履行尊重义务而利用算法侵害公民基本权利的制度设计属于人权保障义务的内容。另一方面,“风险预防”在文义上具有积极作为的内涵,与属于积极义务的人权保障义务相

契合。在基本权利功能体系中,国家保障义务的内容为通过积极作为保护人权,促进人权的实现和更好的发展,具体包括提供落实、保障和救济制度,履行国家给付等等。^{[30](P60-71)}作为一种人为风险且威胁着人权的存在,科技风险预防应当在风险社会中被纳入国家保障义务的范围,国家有义务保障人权免受科技风险的侵害。^[31]由现代算法产生的科技风险应当被纳入国家风险预防义务的规制范围,进而形成算法风险的国家预防义务。

2. 具体基本权利条款与算法风险国家预防义务

“国家尊重和保障人权”条款从抽象的层面为算法风险国家预防义务提供宪法基础,但是仅仅基于概括性人权条款的算法风险国家预防义务过于抽象和空洞,还需要具体基本权利条款的支撑。

算法风险是一种风险类型而非风险概念。

“概念是对特定事实内在典型因素进行准确、全面的概括,在适用时,被适用对象必须与概念的外延和内涵完全对应;而类型则是一种比较松散的认识模型,在使用时只要求所指对象满足其基本的要素或部分要素,而不一定是全称对应。”^{[32](P99)}具体而言,算法风险包含跨越多个领域的风险,其所可能侵犯的基本权利也各有不同。部分算法应用可能侵犯公民的平等权、言论自由权、劳动权、隐私权等,而部分算法应用则可能侵犯公民的一般人格权。例如,自动化行政对行政相对人程序性权利的挤压和忽视,就是算法损害行政相对人人格尊严的体现。^[33]国家义务的履行主要是为了满足基本权利实现的要求,不同的基本权利有不同的权利内容,故国家在履行具体算法风险预防义务时应当考虑算法应用可能直接侵犯的具体基本权利条款之意涵,以确定具体履行的目标。

(二) 基于国家任务或国家目标的算法风险国家预防义务

虽然近现代国家义务理论的发展得益于

^①此处的人权保障义务意指基于宪法第33条第3款的“国家……保障人权”衍生出的狭义层面的国家保障义务,而非泛化的国家人权保障义务。参见郭道晖:《人权的国家保障义务》,载《河北法学》2009年第8期。

人权理论和实践的推动,宪法基本权利条款在“权利—义务”框架下可以证成相应的算法风险国家预防义务,但是该条款并非算法风险国家预防义务的唯一宪法基础。宪法序言和总纲规定的国家任务或者国家目标条款并非仅是一种政治宣示,还会对国家机关施以相应的国家义务。^[34]基于此,国家任务和国家目标条款存在证成算法风险国家预防义务的可能。

除保障基本权利外,宪法与算法风险预防的勾连还在于对秩序的维护。宪法的存在是为了维持共同体的秩序,其本身也意味着一种秩序,秩序价值是宪法的基本价值。^{[35](P69-71)}宪法所保护的社会秩序是包括各社会领域秩序安全和国家安全在内的大社会秩序,主要体现在宪法序言第7至12自然段,正文第4条、第15条、第28条、第36条、第40条、第53条和第54条等规定中。算法风险国家预防义务可以从其中部分规范中衍生出来。首先,宪法序言中有关国家根本任务的规定是算法风险国家预防义务的宪法基础之一。宪法序言第7自然段规定国家建设社会主义现代化强国的根本任务。社会主义现代化建设的过程必然会遇到各种风险的冲击,若不对其中的重大风险进行预防,该进程必然受阻,故而风险预防对我国现代化建设有着重要意义。^[7]算法技术的应用可能会引发政治危机、损害国家利益,“剑桥分析”对各国选民实施操纵的案例是这一算法风险存在的证明。^[36]作为建设社会主义现代化强国的条件之一,我国多民族国家的现实基础也可能会因为由算法推荐形成的有关民族分裂的信息环境而遭到破坏,进而影响宪法根本任务的推进。因此,国家要实现宪法序言规定的根本任务应当防范算法风险。其次,算法风险预防义务的宪法基础还包括宪法第15条第3款和第28条所设定的国家维护社会秩序的目标条款。算法风险对社会秩序具有潜在危害,国家通过相关制度安排预防此类算法风险,可视为落实宪法“维护社会秩序”目标的具体举措,算法风险国家预防义务由此获得宪法规范的正当性支持。值得一提的是,

虽然上述规定中部分使用“禁止”等似乎与“预防”积极面向相对立的词语,但实际不然。尊重义务的履行主体为国家,而上述规定包含为国家以外的组织或个人设定的消极义务。确保这些组织或个人履行相关消极义务,需要国家积极制定和落实相关法律,因而上述规定使用具有消极面向的词语并不影响国家承担具有积极面向的风险预防义务。

五、算法风险国家预防义务的展开

(一) 算法风险国家预防义务的定位

首先应当明确的是,算法风险国家预防义务不以风险预防本身为终极目的,而是一种工具性宪法机制,服务于更高层级的价值目标。该义务的生成具有双重逻辑:一方面源于数字时代国家对算法技术引发的系统性风险规制需求,是对算法技术可能导致的权益侵害、市场失序和公共利益损害等风险的制度性回应;另一方面则体现了宪法对数字时代的适应性调适,是基本权利保障体系向算法领域延伸的必然要求。其核心功能在于为国家规制算法风险提供宪法依据和价值指引,在本质上属于现代风险社会下法律范式转型的体现。其次,算法风险国家预防义务既可以从基本权利条款中推演出来,也可能源于宪法国家任务或目标条款,兼具主观法和客观法性质。基于基本权利条款的算法风险国家预防义务具有权利驱动性,是基本权利国家保护义务在算法领域的具体化。此时,算法风险国家预防义务包含国家采取积极作为最低限度确保公民基本权利不受算法风险侵害的主观法义务,以及超最低限度履行算法风险国家预防义务的客观法义务。^[27]而基于国家任务或目标条款的算法风险国家预防义务则不以特定主观权利存在为前提,其核心在于维护公共利益。由于规范基础采用结果性结构,仅规定国家需要实现的目标而未明确具体实现路径,该类义务具有高度不确定,因而尊

重国家裁量自由空间成为必要,司法救济作用有限。^[37]此时,算法风险国家预防义务更倾向于一种客观法义务。

(二) 算法风险国家预防义务的基本内容

算法风险国家预防义务是由国家机关实施的积极义务,基本内容包括算法风险预防体系建构和完善义务、算法系统安全保障义务、算法风险监管义务、权益救济义务和监督国家落实算法风险预防的义务等,要求国家通过立法、执法、司法、监督等手段予以落实。

第一,基于国家任务或目的条款的算法风险国家预防义务作为客观法义务,以维护社会公共秩序和国家安全为核心目标,国家在实现手段和具体内容上受到相应条款的限制较少,享有较大的裁量权,司法机关可救济的空间较少。此类义务的权益救济义务仅针对社会权益救济,倾向于通过公益诉讼实现对社会利益的救济,而不包括该类义务履行不适当的司法救济情形。因而,此类义务应当重视具有监督权的国家机关保障算法风险预防义务具体落实的重要作用,通过工作报告、执法督察、算法审计等制度落实监督职能。第二,基于基本权利条款的算法风险预防义务以保障人的尊严和具体基本权利为根本目的,其内容与基本条款的规范内涵紧密相连,在实现手段和具体内容上受到的限制较多。其权益救济义务的具体内容包括为公民提供针对此类义务履行不适当的司法救济与公民具体权益遭受算法风险侵害的司法救济。前者意味着,监督国家落实算法风险预防的义务主体包括了司法机关,即司法机关可以通过行政诉讼及附带性审查监督行政机关履行算法风险预防义务行为。后者则需进一步明确该类义务规制对象的双重面向:基于概括性人权条款以及具体基本权利条款与该条款的衔接,算法风险国家预防义务既要防范私主体算法应用行为对权利的侵害,也要重视国家机关算法应用行为可能造成的权利损害。^[38]对于前一种情况,此类义务所包含的算法风险预防体系建构和完善义务应当注重构建和完善权利本

位的算法风险规制体系,通过权利义务的分配尽可能平衡算法主体和算法用户之间的不平等关系;对于后一种情况,此类义务的履行应当侧重于通过落实制度性保障义务规制公权力主体算法应用行为,包括为公权力机关增设程序性义务、为公民增加程序性权利和设置监督制度等。此外,当基于基本权利条款的算法风险国家预防义务落入客观法义务范围时,监督国家落实算法风险预防的义务则与基于国家任务或目的条款的算法风险国家预防义务一致。

(三) 算法风险国家预防义务的实现

算法风险国家预防义务的实现,需要建立在系统化的制度设计和精细化的治理框架之上。当前我国算法风险预防面临的一系列结构性困境不仅制约了治理效能,而且可能导致“规制不足”与“规制过度”的双重风险,进而致使算法安全与创新发展价值之间的失衡。基于此,算法风险国家预防义务的落实应当以禁止不足原则和禁止过度原则为双重标准,通过相应的实现路径,努力实现算法安全与创新发展之间的平衡。

1. 国家履行算法风险预防义务的原则

禁止不足原则和禁止过度原则既是评价国家履行算法风险预防义务的标准,也是宪法对具有较大不确定性的算法风险预防义务之约束,是义务边界所在。禁止不足原则是宪法对基于基本权利条款的算法风险国家预防义务提出的要求。国家履行和落实算法风险预防义务应当满足宪法对基本权利保护的最低要求,即国家通过算法风险预防能够使所欲保护的基本权利之本质核心受到最低限度保护。具体应当以受保护主体的“期待可能性”为基准,从理性人的视角考量算法风险可能造成的损害大小、权利威胁的紧迫性和发生可能性、风险行为的可回避性和受影响主体的自救可能性,以及落实该义务能否确保最低限度保护等。^[39]值得一提的是,禁止不足原则一般不适用于基于国家任务或目标条款的算法风险国家预防义务。原因在于,我国宪法中的国家任务条款和国家目标

条款多适用目标程式,赋予了国家机关(尤其是立法机关)在具体目标、实现手段和实现内容等方面较大的裁量空间。以立法机关为例,就社会主义经济秩序安全的目标(宪法第15条第3款)而言,其裁量权包括了立法时机、规制对象(如金融秩序或交易秩序)、规制手段(民事、行政或刑事)的确定,以及是否就算法技术引发的经济秩序风险进行立法的判断。将禁止不足原则适用于基于国家任务或目标条款的算法风险国家预防义务往往受到国家机关裁量权的限制。

如果说禁止不足原则确立了算法风险国家预防义务的最低履行标准,那么禁止过度原则则指向该义务履行的动态约束边界。限制算法风险国家预防义务的原因在于权力控制、次生风险防控和基本权利保障。第一,算法风险的不确定性要求扩大国家机关履行义务的裁量空间,可能引发权力滥用风险,进而可能侵蚀私主体自治空间,破坏社会子系统自主运行机制,或者加剧公权力主体和公民间的不对等,最终减损公民的合法权利。^[40]第二,算法风险预防涉及事实判断和价值判断,会受到决策者认知局限性的影响,进而可能产生次生风险。^{[41](P17-22)}第三,算法风险国家预防义务确立和履行基于基本权利保障,反之受到基本权利保障的限制。一方面,基本权利保障存在限制,由基本权利保障衍生的算法风险国家预防义务需要受到基本权利的内在限制和外在限制的约束。^{[42](P334-338)}另一方面,为保护特定基本权利或公共利益而对他人基本权利的限制应当受到限制,这是对他人基本权利保障的要求,亦是基本权利限制的限制之意涵。^{[42](P340-350)}具体而言,应当以比例原则作为其动态评价标准,通过正当性、适当性、必要性和均衡性等四个方面的论证划定义务的具体限度。根据《意见》,我国算法风险预防强调技术创新发展和算法安全之间的平衡。遵循比例原则意味着,国家在追求超最低限度的算法风险预防效果时,应当警惕过度强调技术创新发展价值或者算法安全价值,致使个体

或群体权益的核心内涵遭受侵害。

2. 算法风险国家预防义务的实现路径

算法风险国家预防义务可以在禁止不足原则和禁止过度原则的指导下确定其实现的主要路径,对我国算法风险预防困境作出回应。

算法风险预防立法完善的实现路径。通过立法手段落实算法风险预防义务应当立足专项立法和分散立法相结合的立法模式,构建框架统一、层次分明、重点突出的算法风险预防规范体系。面对算法技术特有的系统性风险,专项立法能够弥补场景化规制的局限性。虽然不同领域的算法应用和不同类型的算法技术所可能带来的危害不同,但是共性的风险形态和规制方案仍然存在,这使专项立法具有可能。例如,在技术层面涉及数据偏差、模型黑箱等问题,在法律层面关涉透明、公平等共性原则,在治理层面多个领域都涉及到算法评估、算法备案等共通制度。专项立法的作用在于:第一,确立基本原则和统一框架,统筹协调我国算法风险预防实践;第二,构建基础性制度,提升算法风险预防体系的整体性;第三,明确网信部门和有关部门的监管职责划分,统一网信部门不同领域、不同类型的算法风险监管职责;第四,共性规范为存在法律空白的算法风险预防提供基础性法律依据,为新兴算法风险预防提供前瞻性法律框架。在当前算法立法经验有限、算法风险规制实践初具雏形的背景下,通过总则和分则的方式,在单一法律中塑造预防体系框架并应对不同领域、不同类型的算法风险,尚不具备条件。^[12]专项立法和分散立法相结合的立法模式应当以嵌入式条款回应不同领域的算法风险预防热点需求,以单行法防控重点领域的算法风险,以一般法就算法风险预防原则、规制主体、共性风险以及共同防控规则等建构算法风险预防的体系框架。^[43]

算法风险差异化规制是禁止不足原则和禁止过度原则的共同要求,在立法上体现为立法机关针对不同规制对象差异化配置权利义务。第一,我国以私人场景为基础的算法风险预防

规则标准将使得对公权力主体应用算法行为的规制力度不足与重点偏离,^[44]需区分公权力机关和私主体,并对前者采取更加严格和周延的风险防控手段。第二,衔接个人信息保护、数据安全保护和算法管理分级分类体系,细化“舆论属性或者社会动员能力”标准,根据不同强弱程度为算法主体设置不同程度的义务。

在算法风险预防义务的履行过程中,禁止过度原则要求立法机关必须要在算法安全与技术发展之间寻求平衡。基于我国数字经济建设的战略需求和算法技术发展的现实阶段,程序规制路径应当成为优先考虑的立法选择。第一,相较于实体规则,以程序进路规制算法风险对算法主体的实体权益影响较小,能够在坚持安全底线的同时为技术创新发展保留更多的空间。第二,相较于实体规则的静态性,算法备案制度、算法风险监测制度等程序规制路径更具弹性,能够灵活追踪风险动态,适应算法技术的发展。强调程序规制优先性,绝不意味着否定实体规制的价值。一方面,程序规制进路旨在防控算法运作过程的算法风险而非确定地阻止损害发生,需要实体规则弥补其局限性。^[20]另一方面,实体规则在底线安全领域、基本权利保护和关键基础设施等领域仍具有不可替代的作用。

算法风险预防行政监管的实现路径。在行政领域落实算法风险国家预防义务,需要在比例原则的基础上提高行政机关的监管效能。首先,行政机关规制算法风险的逻辑起点是明确的规制目的,核心在于协调国家(监管者)、算法主体和算法用户三方权益关系,化解个体权利之间、个体和社会利益之间的冲突。行政机关在启动算法风险预防行动前原则上应当深入调研,准确把握算法技术的发展动态及其对不同利益的影响,并通过蕴含公众参与的透明决策机制,确定既能平衡技术发展和算法安全,又具有实践性和合法性的算法风险规制目标。其次,加强行政机关的专业化建设是提升算法风险监管能力的关键。鉴于算法技术的复杂性

和专业性,行政机关应加大对监管人员的培训力度,以提升其识别、评估及应对算法风险的能力。算法风险监管决策过程增设专家咨询机制也是提升监管科学性的方法之一。再次,通过融合“算法风险可接受”与“算法可信”理念将规制目标具体化,提升行政机关算法风险预防的确定性。后者要求算法主体建立健全算法机制机理审核、科技伦理审查、安全评估监测、安全事件应急处置等管理制度和技术措施,侧重过程规制但无法提供稳定的算法安全预期。而前者则通过“概率风险评估法”等量化方法将预防目标定性和定量化,为算法风险预防提供具体的实体要求,进而提高行政算法风险规制的明确性和可预期性。^[45]最后,在监管组织体系上,应强化网信部门统筹协调的作用,构建跨部门协调监管机制。可以尝试制定统一的监管标准、搭建信息共享平台,促进部门间的信息共享和协作,减少监管空白,形成监管合力。

算法风险预防司法审查和多元监督的实现路径。在司法领域落实算法风险预防义务,应建立算法纠纷类型化裁判体系,通过设置专属案由和场景化裁判规则,化解技术复杂性导致的法律适用困境。^[13]面对算法技术快速迭代与法律固有滞后性之间的张力,司法机关应通过法律解释技术填补规则漏洞,实现对算法风险预防立法滞后性缺陷的补强。针对算法系统性风险探索建立公益诉讼制度,为公共利益提供司法救济。最后,通过行政复议和诉讼加强对行政机关部署和应用算法行为的规制,并通过抽象行政行为的附带性审查监督行政机关算法风险预防义务的落实。

保障算法风险国家预防义务的落实,应当构建多元监督体系。权力监督方面,各级人大将算法治理纳入年度监督重点,可以建立立法后评估机制和通过审议其他国家机关有关算法风险预防的工作报告,实现对其他国家机关落实算法风险预防义务情况的监督;行政监督方面,上级监管部门应定期对下级部门的算法治理工作进行督导和完善行政执法考核机制;社

会监督方面,建立重要领域算法应用的投诉举报平台,为公众提供便捷的参与渠道。

六、结语

以机器学习算法为代表的算法技术在21世纪高速发展,为我国带来了前所未有的发展机遇,推动着产业升级、创新驱动和治理效能提升,也催生了复杂多元的算法风险,算法风险预防机制理应迭代升级。确立算法风险国家预防义务,是宪法对算法社会的回应,也是国家具体算法风险预防行为的正当性基础,有助于挣脱部门法视角的桎梏,在整体视角下统筹、规范算法风险预防。然而,国家如何切实有效地落实算法风险预防义务,至今仍是一个开放性问题,各国都需要基于自身国情写下“答卷”。对于我国而言,在算法技术发展与安全保障需求日益凸显的双重背景下,如何把握算法技术发展与安全之间的平衡点,构建既符合我国国情又保有国际竞争力的算法风险预防体系,仍有待在实践中继续摸索和探讨。

参考文献:

- [1]闫静.算法垄断风险的事前预防:理论证成、实践反思与路径优化[J].华中科技大学学报(社会科学版),2023,37(1):108-118.
- [2]李文超,武一帆.算法侵害行为的事前规制与侵权救济研究[J].法律适用,2023(3):119-128.
- [3]沈伟伟.算法透明原则的迷思——算法规制理论的批判[J].环球法律评论,2019,41(6):20-39.
- [4]石晶.生命科技风险的宪法回应[J].法学,2023(2):47-61.
- [5]张海涛.“风险社会”的宪法结构分析[J].湖北社会科学,2021(4):128-138.
- [6]李忠夏.风险社会治理中的宪法功能转型[J].国家检察官学院学报,2020,28(6):3-15.
- [7]王旭.论国家在宪法上的风险预防义务[J].法商研究,2019,36(5):112-125.
- [8]张凌寒.自动化决策与人的主体性[J].人大法律评论,2020(2):20-48.
- [9]陈姿含.人工智能算法中的法律主体性危机[J].法律科学(西北政法大学学报),2019,37(4):40-47.
- [10]洪丹娜.算法歧视的宪法价值调适:基于人的尊严[J].政治与法律,2020(8):27-37.
- [11]段沁.“环境国”下的国家扩张与宪法应对——以环境风险预防为例[J].中德法学论坛,2021(2):216-236.
- [12]丁晓东.全球比较下的我国人工智能立法[J].比较法研究,2024(4):51-66.
- [13]陈增宝,张凌寒.算法技术的法律规制:治理困境、发展逻辑与优化路径[J].中国应用法学,2024(4):155-166.
- [14]张凌寒.中国需要一部怎样的《人工智能法》?——中国人工智能立法的基本逻辑与制度架构[J].法律科学(西北政法大学学报),2024,42(3):3-17.
- [15]蔡琳,卢青霞.从技术导向型到以需求回应型为主的转变——论算法规制的路向问题与破解[J/OL].科技进步与对策:1-12(2025-01-23)[2025-03-26].<http://kns.cnki.net/kcms/detail/42.1224.G3.20250122.1738.038.html>.
- [16]任颖.算法规制的立法论研究[J].政治与法律,2022(9):98-111.
- [17]侯卓.“领域法学”范式:理论拓补与路径探明[J].政法论丛,2017(1):87-96.
- [18]刘茹洁.论宪法的稳预期功能[J].北方法学,2023,17(4):151-160.
- [19]王锡锌.个人信息国家保护义务及展开[J].中国法学,2021(1):145-166.
- [20]王莹.算法侵害责任框架刍议[J].中国法学,2022(3):165-184.
- [21]赵鹏.风险社会的自由与安全[M]//沈岍.风险规制与行政法新发展.北京:法律出版社,2013:11-14.
- [22]乌尔里希·K·普罗伊斯.风险预防作为国家任务——安全的认知前提[M]//刘刚.风险规制:德国的理论与实践.北京:法律出版社,2012.
- [23]约翰·丹纳赫.日常生活算法外包的伦理[M]//彭诚信.驯服算法:数字歧视和算法规制.林少伟,唐林垚,译.上海:上海人民出版社,2020:111-126.
- [24]U. S. CFTC, U. S. SEC. Findings regarding the market events of May 6, 2010: Report of the Staffs of the CFTC and SEC to the joint advisory committee on emerging regulatory issues[EB/OL]. (2010-09-30)[2024-11-20]. <https://www.sec.gov/news/studies/2010/marketevents-report.pdf>.
- [25]张凌寒.算法权力的兴起、异化及法律规制[J].法商研究,2019,36(4):63-75.
- [26]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[J].法学研究,2020,42(2):42-56.
- [27]李海平.基本权利的国家保护:从客观价值到主观权利[J].法学研究,2021,43(4):39-54.

- [28]陈醇. 论国家的义务[J]. 法学, 2002(8): 15-19.
- [29]焦洪昌. “国家尊重和保障人权”的宪法分析[J]. 中国法学, 2004(3): 44-51.
- [30]张翔. 基本权利的规范建构[M]. 北京: 法律出版社, 2017.
- [31]石晶. 现代科技风险国家预防义务的限定[J]. 中国特色社会主义研究, 2024(4): 95-109.
- [32]姚建宗, 等. 新兴权利研究[M]. 北京: 中国人民大学出版社, 2011.
- [33]张涛. 自动化行政对行政程序的挑战及其制度因应[J]. 华中科技大学学报(社会科学版), 2022, 36(5): 60-69, 89.
- [34]王锴, 刘彝昊. 宪法总纲条款的性质与效力[J]. 法学论坛, 2018, 33(3): 27-34.
- [35]陈雄. 宪法基本价值研究[M]. 济南: 山东人民出版社, 2007.
- [36]数万文件揭露剑桥分析操纵各国选民[EB/OL]. (2020-01-14)[2024-11-20]. <https://news.qq.com/rain/a/20200114A092NW00>.
- [37]李海平. 区域协调发展的国家保障义务[J]. 中国社会科学, 2022(4): 44-63, 205.
- [38]王建学, 高强. 国家保障人权: 思想溯源与规范重述[J]. 北方法学, 2024, 18(3): 38-53.
- [39]陈征. 宪法中的禁止保护不足原则——兼与比例原则对比论证[J]. 法学研究, 2021, 43(4): 55-72.
- [40]周辉. 算法权力及其规制[J]. 法制与社会发展, 2019, 25(6): 113-126.
- [41]乌尔里希·贝克. 新的现代性之路[M]. 张文杰, 何博闻, 译. 南京: 译林出版社, 2018.
- [42]林来梵. 宪法学讲义[M]. 4版. 北京: 清华大学出版社, 2023: 334-338.
- [43]曾雄, 梁正, 张辉. 欧盟人工智能的规制路径及其对我国的启示——以《人工智能法案》为分析对象[J]. 电子政务, 2022(9): 63-72.
- [44]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.
- [45]胡帮达. 安全和发展之间: 核能法律规制的美国经验及其启示[J]. 中外法学, 2018, 30(1): 208-230.

【责任编辑 史 敏】

The Justification and Operational Construction of the State's Obligation to Prevent Algorithmic Risks

HONG Danna & HUANG Kaiqi

Abstract: China's algorithmic risk prevention system exhibits three structural flaws including fragmented regulations, dispersed enforcement agencies, and undifferentiated regulatory targets, resulting in systemic deficiencies like regulatory gaps, overlapping jurisdictions, and disproportionate interventions. Fundamentally, these deficiencies stem from the limitations of non-preventive measures, systemic gaps in current mechanisms and the worsening power-rights imbalance. This necessitates establishing the state's preventive obligation to counteract algorithmic applications' irrational outcomes, and provide unified constitutional guidance, thereby constructing a systematic regulatory framework and optimizing the existing legal regime. Rooted in fundamental rights provisions and the constitutional mandates in the Preamble and General Principles, this obligation functions as an instrumental constitutional mechanism with dual subjective/objective legal nature. Its implementation must satisfy minimum protection requirements without exceeding proportionality limits, achieved through coordinated legislative, administrative, judicial, and oversight measures.

Keywords: algorithmic risk governance; state's constitutional obligations; fundamental rights protection