文章编号:1671 - 1653(2025)01 - 0079 - 08

隐私信息保护场景中平台人工智能算法的治理路径

——以场景一致性为视角

寿晓明

(安徽大学 法学院,安徽 合肥 230601)

摘 要:随着《中华人民共和国个人信息保护法》《生成式人工智能服务暂行办法》的相继出台,人工智能算法的专项立法亟待推进。数字经济时代,算法、隐私、数据和信息的深度融合使得人工智能算法规制需要同步推进隐私保护和数据治理。当下数据治理刚刚起步,隐私信息保护由于人工智能算法规制的缺失难以有效推进。场景一致性框架下建构个人隐私信息合理流动作为隐私保护的社会本位立场,而人工智能算法规制同样需要回归场景。基于人工智能算法收集、分析和应用的流程本质,隐私保护和算法规制以实时动态的流程作为逻辑契合点,在人工智能算法流程中实现实时动态的隐私保护。人工智能算法技术作为法律规则的底层逻辑,构建了技术规制与法律规制并行的双重机制。建构平台问责制的立法体系,并建立事前算法评估、事后算法解释的算法责任全流程规制体系,形成人工智能算法专项立法的基础。

关键词:人工智能算法;算法流程;场景一致性;隐私信息保护;算法规制

中图分类号:D922.17 **文献标识码**:A **DOI**:10.7535/j.issn.1671-1653.2025.01.010

Governance Path of Platform Artificial Intelligence Algorithm in Privacy Information Protection Scenario: From the Perspective of Scene Consistency

SHOU Xiaoming

(School of Law, Anhui University, Hefei 230601, China)

Abstract: With the promulgation of *The Law* of *The People's Republic of China on the Proteetion of Personal Information* and the *Interim Measures for Generative Artificial Intelligence Services*, the special legislation of artificial intelligence algorithms needs to be promoted urgently. In the digital economy era, the deep integration of algorithms, privacy, data and information requires the artificial intelligence algorithm regulation to simultaneously promote privacy protection and data governance. At present, data governance has just started, and privacy information protection is difficult to effectively promote due to the lack of artificial intelligence algorithm regulation. The reasonable flow of personal privacy information under the

收稿日期:2024 - 06 - 09

作者简介:寿晓明(1996—),男,浙江绍兴人,安徽大学法学院 2024 级经济法学专业博士研究生,主要从事竞争法、数字法研究。

framework of scene consistency is constructed as the social standard position of privacy protection, and artificial intelligence algorithm regulation also needs to return to the scene. Based on the essence of artificial intelligence algorithm collection, analysis and application process, privacy protection and algorithm regulation are logically aligned with real-time dynamic processes, achieving real-time dynamic privacy protection in the artificial intelligence algorithm process. Artificial intelligence algorithm technology, as the bottom logic of legal rules, constructs a dual mechanism of parallel technical regulation and legal regulation. It is necessary to construct the legislative system of platform accountability, and establish the whole process regulation system of algorithm responsibility of pre-algorithm evaluation and post-algorithm explanation, so as to form the basis of special legislation of artificial intelligence algorithm.

Keywords: artificial intelligence algorithms; algorithmic flow; scene consistency; privacy information protection; algorithm regulation

一、引言

当下以“人工智能+”行动推动新质生产力发展成为了时代新命题。数字时代数据在ChatGPT、算法等人工智能技术的加持下成为了一种新兴的生产资料,在数字经济的发展中起到了愈发重要的作用。数据的来源往往涉及个人信息和隐私,实践中,保护这些信息显得尤为困难,这很大程度上是因为在攫取数据背后的经济利益时,没有对个人权利的保护引起足够重视,尤其是在人工智能算法的加持下侵犯个人信息和隐私变得更为隐匿。个人信息很大程度上依赖于知情同意等一系列授权进行规制,但对于更为隐蔽的隐私信息保护更为困难。人工智能时代,算法迅速内嵌社会生活各领域的同时,脱离纯粹的技术属性,形成一种新型的权力形态——算法权力[1]。大数据、算法、平台成为经济发展的关键要素、动能及场域。算法因其技术性特点给生活带来了巨大便利,但算法超出技术工具范畴影响或替代公权力决策,形成新的权力形态而引发的诸多风险也引起了学界的关注[2]。强大的资源调配能力使得算法拥有权力基础而形成算法权力。算法权力嵌入社会之中虽为社会的运行带来一定便利,但亦造成算法歧视、算法侵权、算法与公权力合谋等一系列社会风险[3]。算法“霸权”的形成产生了区别性歧视风险、加剧隐私信息受侵害风险和社会不公平性风险[4]。从规制理论来看,作为规制对象的数字平台私权力兼具人源性与物源性特征,其演变具有动态性与系统性特征,这对作为规制手段的法律系统提出统一价值理性与工具理性、注重个体主体地位、优化规制工具的要求[5]。同时在司法实践中此类案件也日益增多,对于人工智能算法的专项立法过于滞后,使得司法机关对此类案件的重视度不够,无法充分保障数字经济的发展。算法是内在的运行指令,数据是外化的表现形式,信息是数据承载的内容[6]。人工智能算法和个人隐私信息等数据的深度融合导致隐私信息保护和算法规制难以分离。随着《中华人民共和国网络安全法》(以下简称《网络安全法》)、《中华人民共和国数据安全法》(以下简称《数据安全法》)、《中华人民共和国个人信息保护法》(以下简称《个保法》)、《生成式人工智能服务暂行办法》的出台,隐私信息保护缺失的一角,即人工智能算法规制立法被提上日程,如何实现隐私信息保护场景中的人工智能算法治理成为了当下紧迫的时代命题。

二、隐私信息保护场景中人工智能算法规制的基础与风险

人工智能算法依靠自身的机器优势、结构优势和嵌入优势形成了突破数学意义上的社会性权力。“算法技术利用其掌握的数据信息与分析能力,在社会各个领域干预个人、政府等决策,由此衍生出‘算法权力’这一概念。”[7]基于场景一致性[8]理论下的隐私信息保护需要回到平台人工智能算法规制的特定社会环境中,实现隐私信息在人工智能算法规制中的合理流动,做到隐私保护和数据流通利用的利益平衡。

(一)隐私信息保护场景中人工智能算法风险的规制基础

场景一致性框架中的隐私权利究其根本是实现隐私信息的合理流动。《中华人民共和国民法典》(以下简称《民法典》)将“隐私”界定为“自然人的私人生活安宁和不愿为他人知晓的私密空间、

私密活动、私密信息”。其中：私密空间作为一种空间状态存在；私密活动作为一种行为状态存在；私密信息作为和个人信息交叉保护的中间状态存在。争论最大的是私密信息以及数字化的私密空间和私密活动，隐私被泄露往往是由传递数据化信息进而侵犯了私密空间和私密活动所导致。数字化的私密空间和私密活动产生了数字化私密信息，与之相对的是原始的私密信息。很大程度上由于私密信息隐藏的经济利益被平台通过算法进行攫取利用，而隐私信息由于涉及个人自由和人格尊严应当禁止其作为数据进行流通利用。立法者基于保护和平衡社会利益的本位，应积极治理以保护隐私权。隐私权是保障个人信息合理流通的平衡性权利。

场景一致性理论构成了数字时代隐私信息保护的基本框架，在平台人工智能算法规制这个特定社会环境中构建角色、活动、规范和价值理念。角色是指在人工智能算法平台规制中的平台、平台用户、算法设计者和算法使用者等；活动是指平台通过人工智能算法获得用户的隐私信息以及算法的设计者设计算法等行为；规范是指平台使用人工智能算法收集隐私信息时的禁止性行为和可接受行为，以及算法设计者设计算法时的设计义务和权利；价值理念是指在平台人工智能算法规制中秉持的目的和价值。平台人工智能算法规制的价值理念是维持算法的技术中立性，实现个人信息合理流通，既保障信息隐私的安全又促进数据的流通利用。对于隐私信息的保护需要回归平台人工智能算法规制这个特定场景中，人工智能算法规范作为场景一致性理论的核心所在，是平台通过人工智能算法收集隐私信息和算法涉及的强制性规定，规范涉及场景中的权利义务和责任。而如何权衡个人信息商业化利用和信息隐私保护之间的冲突，是算法全面嵌入社会语境下技术风险需要应对的时代命题[9]。

（二）隐私信息保护场景中人工智能算法规制的风险来源

“数字时代生成式人工智能强大的黑箱算法由于获得了海量的个人信息使得其具备了识别个人信息的数据基础和技术基石，黑箱算法的自动化决策可以在人机交互过程中轻松调取相关的个人信息进行输出。”[10]人工智能算法摆脱了数学工具的枷锁，在和数据的深度融合下异化成为新兴的社会型权力。人工智能算法形成的自动化决策使得算法权力的辐射范围日益扩大，而自动化决策导致的算法黑箱使得普通大众对于人工智能算法的运行机理难以有效认知，算法权力的透明性和自主性日益架空个人的法律主体地位，算法权力甚至在一定程度上开始异化个人成为行为的客体，从而颠覆整个法体系的基石。

人工智能算法通过自动化决策掌控数据的输入和输出，演变为资源、财产等社会权利义务的重新配置。民事领域私权力的出现最早可以追溯到市场垄断，然而平台作为商业架构和组织模式的出现意外且快速地拓展了私权力的衍生空间[11]。平台出于自身经济利益的实现和匿名成本的考虑不可能对隐私信息进行匿名化处理。平台在收集隐私信息等信息后通过算法自动化决策分析出个人的行为规律，从而引导个人的行为，使得个人的主观意志被算法潜移默化地深度影响，个人行为甚至受到人工智能算法的远程虚拟操控。同时，公权力为了监管算法和技术辅助不得不依赖算法的自动化决策，平台在运用人工智能算法处理隐私信息时，公权力需要借助算法以技术制约技术，人工智能算法产生的知识鸿沟使得规制算法的底层设计必然是运用技术进行规制，自然规制人工智能算法不单纯是技术层面的后果，也是在技术底层逻辑基础上建构法律的权利义务和责任配置。

（三）隐私信息保护场景中人工智能算法规制的法律风险

在隐私信息保护场景中，平台通过算法自动化决策，收集用户的隐私信息并进行分析处理，然后转化成为数据，从而攫取高额利润，并借助算法黑箱逃避现有法律体系的规制。而现有法律体系出台了《民法典》《个保法》等保护个人信息和隐私信息，但没有出台相应的人工智能算法规制的部门法，从而使保护隐私信息在人工智能算法一环中出现纰漏。人工智能算法通过收集隐私信息掌控用户日常生活的行为规律，并以此作为训练样本形成个性化算法推荐，从而在潜移默化中对个人行为进行预测、引导和操控，以此掌握市场主动权和获取高额利润。若未获得用户的允许，个性化推荐算法，包括基于内容的推荐、协同过滤推荐和基于知识的推荐等，计算判别用户喜好的行为，既干扰了用户的私有领域，又窥探了用户的人格图像[12]。平台借助用户画像实现歧视性行为，导

致平台和用户之间在人工智能算法权力的加持下信息的不对称日益严重。用户相对平台处于劣势地位,信息的不对称性极大程度上影响了用户的主观意思,使得用户在人工智能算法权力面前难以作出正确的判断,从而导致权利义务的不对等。

运用算法对用户数据进行提取、分析和处理,对用户实施长时间不间断的单向监控,训练用户数据,以此预测用户行为并基于用户行为构建相应的调整机制,使用户按照平台希望的行为轨迹发展,人工智能算法再基于用户行为逻辑生成个性化决策,从而作出进一步的预测和引导。平台运用人工智能算法监视用户的目的是获取人工智能算法所需的质料以及对用户未来作出精准预测的机会,用户的隐私信息在实时监视下显露无疑,甚至个人的隐私信息会成为规制个人的数据源,而个人在人工智能算法和数据深度融合的信息茧房中从主体逐渐退化成为算法权力的客体,用户的隐私信息则在人工智能算法计算下源源不断地生成数据剩余价值。

三、隐私信息保护场景中人工智能算法流程规制的逻辑理路

"人工智能的决策过程应该被理解为计算机科学家所制定的问题和目标的实现过程。"[13]技术性正当程序权利是对算法权力的程序性与实质性的双重约束,是算法权利束中统辖其他权利的概念[14]。算法监管涉及政治、经济和法律三个社会子系统,应当通过数据活动顾问这一"接口岗位"实现系统间的结构耦合,从而借助系统间的协力有效监管算法活动[15]。应及时树立以风险防范为目的的监管思路,实行内容与算法并重的双轨审查机制,以及设立平台责任与技术责任双轨并行的责任体系,并对算法的生产性资源数据的收集和使用进行合理限制[16]。在算法场景化规制原则的指引下,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[17]。

(一)人工智能算法事前收集的隐私信息控制

收集阶段是人工智能算法流程中平台运用算法收集用户隐私信息的阶段。人工智能算法流程的开始需要隐私信息等元数据作为质料。人工智能算法规制流程通常具备一个响应新数据输入的自动化决策算法,数据可以从用户手中收集,通过匿名化算法和技术去除个人信息的人格和身份属性,转化成为匿名化信息。自然数据收集可以通过传感器或数据交易获取。此外,可以根据现有数据进行推断和预测,这种推断和预测又会构成新的数据。平台运用人工智能算法收集了海量的隐私信息等元数据,如果不在人工智能算法规制流程的开始便对其进行管控,则难以保证人工智能算法所需的训练样本不会掺杂污染数据。算法的程序正义理论在确立"以人为本"的智能伦理观的基础上,通过算法公开、算法影响评估和问责等新的制度性设计,化解传统程序所面临的正义风险,促进正义价值的实现[18]。

在人工智能算法事前收集阶段,对于元数据的控制需要先界定个人信息和隐私信息,以此从人工智能算法流程的源头保护隐私信息的安全。根据《民法典》《个保法》,隐私信息只涵盖足以识别特定自然人的信息。由于个人信息兼顾人格属性和财产属性,而隐私信息是一种精神性权利,因此两者的保护路径也有所差别。根据《民法典》中规定的知情同意原则,平台在收集用户的个人信息中的非隐私信息可以采取默示知情收集,从而最大程度地减少平台沟通成本和提高收集效率,自然应当赋予用户拒绝的权利作为格式条款的一项进行规制。而对于收集用户的隐私信息,平台则需要用户明示同意,并需要在用户授权后进行匿名化处理,从而保护隐私信息的安全。人工智能算法事前收集阶段对于隐私信息的管控采取合理方式进行流通,建构个人信息和隐私信息的默示知情—明示同意的二元规制路径。界定算法场景中的数据信息是否构成信息隐私时,应采"信息自主决定 + 控制人格图像"的判断标准[19]。由于数据和算法的深度融合,隐私信息保护和人工智能算法流程规制在逻辑上实现契合,隐私信息保护需要在人工智能算法流程这个特定场景中实现,基于场景一致性理论的规范建构,人工智能算法流程规制是以保护个人的隐私信息安全为基础的。

(二)人工智能算法事中分析的隐私信息监管

人工智能算法流程规制的第二个阶段是对获取的数据进行算法分析,并全程实时监管数据中是否遗留隐私信息,同时对出现隐私信息的训练样本进行追查。数据分析需要依赖数据库管理和数据处理软件,如数据分析的前端预处理技术、数据挖掘和支持技术。数据被用来构建逻辑中的假设部分,而假设是数据收集或分析的起点,决定着结论是否相对中立、客观和真实。假阳性和假阴

性的存在风险其实和传统统计数据类似，高度拟合的风险同样较为突出。数据永远不可能完全决定性的重现真实世界。“一个或多个人类决策或评估被记录为结构化数据，并且该数据被自动汇总或合并以生成用于做出决策的总体评分或评估。”[20]数据的选择取决于其对数据组织的时效性和目的性，其效果和代表性并不能完全预测或者进行事后评价。在人工智能算法事中分析阶段的隐私信息监管则依赖上述技术进行人工智能算法规制，同时对于数据分析形成的相关关系进行法律上的因果关系认定，以此为事后责任的分配打下基础。

在人工智能算法流程规制的分析阶段，数据分析所形成的是相关关系的信息，并非因果关系的信息，两个变量或者事件之间的联系可能纯粹出于巧合而并不具有必然性。事实上的相关关系并不等于法律上的因果关系，“依据特征组合作出的预测在某种程度上与目标信号存在因果关系。”[21]将人工智能算法流程分析阶段的相关关系转化成法律上的因果关系，为算法规制打开了大门。根据人工智能算法的作为和不作为进行划分，判读人工智能算法的因果关系标准是行为风险的关联性。从作为层面看，对于平台人为操控人工智能算法产生的作为行为应当认定为因果关系。平台积极地作为行为操控人工智能算法侵犯用户的隐私信息，所实施的侵权行为危害到的不仅仅是用户个体，甚至是群体和社会的利益。对于人工智能算法的不作为行为则可以分为积极的不作为和消极的不作为。积极的不作为由于出于故意的意思表示，对于不作为引发的风险存在预见的可能性，应当对人工智能算法的因果关系进行认定；消极的不作为，由于主观上出于过失的意思表示，不作为的义务来源要求平台对风险具备注意义务和作为义务，且产生的风险具有可避免性。此时的风险应当限制在可认知的范围之内，对于超出认知范围的风险不应当认定为法律上的因果关系。

（三）人工智能算法事后应用的隐私信息责任

在人工智能算法流程规制的事后应用阶段，应用则是从数据分析中获得信息，获得的信息可能是知识、模型或者预测，通过分析得出的信息可能会形成通用的决定或者针对个人的决定。依据事后算法自动化决策是否直接影响用户的权利义务，进行相应的责任追究。对于侵犯个人合法权利的决策，可以要求平台作出相应的合理解释，从而做到隐私信息保护的事后救济。使用数据作出针对个人的决策可能是通过自动化方法得以实现，或者由个人或机构通过基于分析产生的知识作出决策。知识是通过汇集来源广泛的数据集而产生，这些数据集则是影响用户特定个体决策背后的依据。在绝大多数情况下，算法所作出的一般性决策会对用户特定个体的行为策略产生影响，甚至在不需要处理用户特定个体的数据时影响其内心主观想法。但是当算法作出的决策是针对用户特定个体本身的时候，数据仍然主要来自分析阶段的其他来源，以及在应用阶段的有限信息决定了结果或者决策。人工智能算法事后应用所作出的自动化决策一旦侵犯到用户的合法权益，用户则有权要求平台作出合理解释，以此作为隐私信息事后救济和人工智能算法事后规制的契合点。

综上，分析阶段的大量数据加上应用阶段的少量信息，是人工智能算法流程规制中影响决策、预测和新信息的因素。人工智能算法流程是一个极其复杂的程序，从实践和法律视角出发可将其分为获取、分析和应用三大阶段。获取阶段涉及大量特定用户个体的数据；在分析阶段，利用各式各样的人工智能算法分析数据提炼知识；在应用阶段，收集的用户特定个体的数据会被用于平台这个特定场景中，但知识和信息的主要来源并非一定来自用户特定个体的自主数据，也可能来自其他数据。显然获取阶段、分析阶段和应用阶段是分离的。这已暗示了人工智能算法流程中不同法律规制、不同人工智能算法流程的不同部分、不同阶段需要运用不同的法律手段和解决方案。人工智能算法流程规制在场景一致性框架下建构起来，在隐私信息保护这个特定场景中，显示出人工智能算法流程是由不同阶段组成，不同的法律规制适用不同的行为，不同的行为则伴随着不同的风险和责任。这是人工智能算法流程划分为不同阶段进行规范分析和法律分析的核心和优势。通过规范的计算化、规制的技术化、技术的反馈化三个界面的勾连，使规范穿透算法的外衣，直达算法的后台[22]。

四、面向场景一致性的人工智能算法问责规制路径

基于场景一致性框架下的隐私信息保护，对

于人工智能算法的监管应当贯穿隐私信息保护全过程,从事前、事中和事后对平台使用的人工智能算法进行全流程法律规制。“厘清法律责任是对算法进行规制的前提,而算法黑箱对因果关系的重构会直接影响到法律责任的分担。”[23]平台人工智能算法规制的立法路径应当建构的是平台问责机制,对于产生具体侵害结果的人工智能算法损害归责而言,分别设置因果支配型的责任与非因果支配型的义务型责任[24]。

(一)隐私信息事前保护的人工智能算法评估机制

基于场景一致性理论,在隐私信息侵权事前建构人工智能算法评估机制,从而在事前对隐私信息予以保护。为了管控因赋能导致的衍生性应用风险,一方面立法应当强调人的主体性,另一方面立法应根据各类应用场景中的利益顺位,灵活配置权利[25]。平台应当在事前建立人工智能算法评估备案机制。平台因人工智能算法被问责的原因是平台在保护个人隐私信息时没有履行相应的法律责任和道德义务,无法向用户解释和证明平台在侵犯隐私信息时主观并不存在过错。合理的平台问责机制有赖于清晰的主观过错认定,这需要通过人工智能算法评估和算法备案对平台事前和运行中的问责点予以固定,从而评估平台是否可以评估、控制和纠正人工智能算法在侵犯隐私信息时带来的责任和危害。设置平台问责点需要依靠人工智能算法评估机制的运行,人工智能算法评估机制可以有效预防平台事后隐瞒和错误披露的问题,也可以预防错误计算反复适用于海量主体带来的损害扩散化。以风险预防为目的的事前人工智能算法监管成为各国立法实际的选择。我国《个保法》第五十四条、第五十五条就规定了个人信息处理者应当对个人信息遵守法律、行政法规的情况进行合规审计,个人信息处理者对利用个人信息进行自动化决策进行评估。应当按照人工智能算法处理的数据所涉及的利益主体、对各方行为的干预程度作为基本准则,确立不同等级的人工智能算法风险机制并设立不同的监管标准。而对于涉及的隐私信息的风险等级和监管标准则应当最为严格。为解决现行隐私保护的法律标准不够灵活、不够中性的理论困境,我国应重构自主算法隐私保护的评价标准[26]。平台需要着重设计部署算法的目的、风险和控制能力,以便日后进行人工智能算法问责时可以及时追溯和审查。

“算法备案有效实施的关键在于其法律属性的廓清,算法备案与行政许可、行政确认等行政行为存在内容、结构及性质层面的显著区别。”[27]为了以备日后审查,人工智能算法评估的内容需要进行相应的备案,平台需要将运用的人工智能算法提交给国家监管部门作为存档审查的内容,国家监管部门获得平台设计部署具有潜在风险和危害的算法系统的相关信息,从而固定隐私信息侵权的算法问责点便于日后追责。根据人工智能算法风险等级的差异,对人工智能算法进行平台的元规制或者向国家监管部门进行独立规制。平台进行人工智能算法备案是为了事后追责确定人工智能算法的问责点,而并非作为行政许可的前置程序或者审批程序,法无禁止,即在自由的基本准则下人工智能算法备案只是起到监管平台的作用,不能作为阻碍平台发展的行政障碍。平台应当根据监管第三方提供的格式要求提前制作人工智能算法备案内容,内容应当包括算法目的、流程、风险等,备案的内容可作为隐私信息侵权事后追责平台责任豁免的相关证据,也可作为监管平台持续性的定期审核。

(二)隐私信息事后保护的人工智能算法解释机制

基于场景一致性理论,在隐私信息侵权事后建构人工智能算法解释机制,从而在事后对隐私信息予以保护。“算法解释承载着权益保障、社会交往和风险治理三重意义,其在技术层面上的障碍正在逐渐被突破,可以通过多种技术机制实现。”[28]人工智能算法评估备案是判断平台在算法运行侵犯隐私信息时主观上是否故意或者过失的依据,同时根据客观上人工智能算法侵犯隐私信息造成的损害结果进行行政处罚和问责,被侵权人则可以提起民事诉讼要求平台进行侵权责任赔偿。人工智能算法解释是平台侵犯隐私信息时认定主观过错的重要内容。在平台被处以行政处罚前,平台对于行政处罚可以作出相应陈述和申辩,这是程序正当作为行政法基本原则赋予行政相对人的程序权利。平台因为人工智能算法侵权被处以行政处罚,作为法律上的不利后果,如果缺少平台对此作出的陈述和申辩,则行政处罚存在严重的程序瑕疵,而人工智能算法解释则是平台遭受行政处罚时所作的陈述与申辩。人工智能算法解

释应当成为一个独立的体系性解释,实践中人工智能算法解释只是作为平台排除或者认定其法律责任的限度,只是在法庭调查时发起并没有成为一个独立的程序,这显然不利于保护平台的合法权利。应当适用平衡工具与人的主体性之间的关系、贯彻正当程序原则、严格审慎而不失灵活性的立法态度等法治要求,进而构建算法规制工具的内部优化和决策矫正的法治框架[29]。

人工智能算法解释应当作为平台人工智能算法问责机制的一个独立环节。平台的人工智能算法解释作为监管第三方审查的必要内容,应在事后,即个人隐私信息客观损害发生后作出,平台人工智能算法解释的内容同样会成为监管第三方法律评价的内容之一。平台如果在人工智能算法解释中隐瞒、虚构解释同样可能会被要求承担相应的法律责任。同时,国家监管部门在启动人工智能算法审查和相应的行政处罚时,平台应当承担举证自身合规的举证责任,即事后作出人工智能算法解释。“为了便于有效审查算法系统的运行情况,需考虑所讨论的系统、其部署的领域以及目的。”[30]人工智能算法解释的内容本身也是对人工智能算法合法性和合理性的审查,可参照对抽象行政行为的附带性审查。由于平台处于强势地位,因此证明人工智能算法的合法性、合理性及无歧视性等责任应当由交易平台承担,从而平衡各方的权利义务。由于知识鸿沟的存在,监管第三方需要将平台的人工智能算法解释作为一个独立环节对待,独立清晰的人工智能算法解释环节可以提供相类似的人工智能算法监管经验和数据,以便监管第三方日后的人工智能算法监管。有必要引入“算法发展”平衡算法安全、“权利公平”补充算法公平、“私人自主”调和算法向善,进而推动规制主体分工协作、拓展规制对象范围、优化规制工具,以铸就彰显中国风格、体现中国智慧的算法规制体系[31]。

(三)隐私信息全流程保护的人工智能算法责任机制

基于场景一致性理论,在隐私信息侵权全流程建构人工智能算法责任机制,从而在全过程对隐私信息予以保护。“有必要对原有责任制度进行扩充,打破平台公司的过错和算法的不可解释性这两大归责障碍。”[32]传统的追责逻辑遵循的是“主体—行为—责任”链条,需要厘清隐私信息侵权责任链条的三要素。结合场景一致性理论的隐私信息规范,平台是隐私信息侵权的责任主体。虽然人工智能算法侵犯隐私信息是多方因素造成的,但平台作为人工智能算法使用者和算法利益既得者,基于权责一致的原则应当对人工智能算法造成的隐私信息侵权承担一定的责任和义务。根据事前固定的人工智能算法评估备案和事后认定算法解释,平台侵犯隐私信息的法律责任应当根据主客观过错的程度轻重分层设置。从平台客观角度出发,平台作为算法运营者对于算法决策结果存在着根本性影响。平台或多或少会利用编程代码内嵌规则在人工智能算法运行中嵌入自身的主观意图。即使人工智能算法运行规则方式输出的结果不为平台所控制,但算法深度学习输出结果的技术路线依然无限接近平台设定的输出目标。同时平台对于算法决策结果负有注意义务,基于《网络安全法》《数据安全法》等法律法规,平台可采取备案、审查、复核、验证等必要手段,确保算法决策结果合法合规。

此外,平台算法问责需要依赖清晰的主观过错认定。引入“算法向善”理念,回归侵权法过错归责框架,个案中算法平台注意义务的界定需要综合考虑算法平台性质、算法应用场景以及信息管理能力等三类因素[33]。平台的隐私信息侵权过错应当包括不得操控用户行为、保障用户的自治性等技术伦理内容,同时应当鼓励平台主动承诺扩大自身主体义务的涵盖范围。对于人工智能算法中的关键性算法,国家应当制定统一的技术标准,平台倘若违反统一的技术标准,则可以直接推定平台的主观故意。对于平台的人工智能算法责任,监管三方治理平台应当基于人工智能算法流程的特征履行其治理责任,需要在整个人工智能算法流程的事前、事中和事后进行全方位的归责。当平台使用人工智能算法侵犯用户隐私信息时,监管第三方应当首先对平台是否对人工智能算法进行了有效控制进行检查,从而确保人工智能算法的正常运行。当发生了重大损害并且平台对此不存在控制措施,或者平台在满足人工智能算法运行标准出现了疏忽过失时,那么平台则有可能承担相应的责任和处罚。但如果平台采取的控制措施是适当的,监管第三方则可以认定平台主观上不存在故意或者过失。“在重构算法信任和算法问责的基础上规制算法,重塑法

律层面的主体—行为—责任链条。”[34]

五、结语

随着欧盟《人工智能法案》的出台和“中华人民共和国人工智能法(学者建议稿)”的提出,人工智能算法治理被提上日程。基于场景一致性理论框架下的人工智能算法全流程规制思路,隐私信息保护在此找到了逻辑上的契合点,对于人工智能算法的规制同样需要进行场景化规制。基于场景一致性理论建构人工智能算法流程的平台责任制,以事前人工智能算法评估、事后算法解释的全流程算法责任规制体系为基础,建构人工智能算法专项立法的基石。我国已实现了人工智能算法治理与平台治理的深度融合,数据保护立法刚刚起步,唯有将人工智能算法规制、隐私信息保护和数据治理同步推进,才能促进人工智能算法治理体系的建构。

参考文献:

[1]赵一丁,陈亮.算法权力异化及法律规制[J].云南社会科学,2021(5):123.

[2]于淼.数字经济视域下算法权力的风险及法律规制[J].社会科学战线,2022(2):126.

[3]戴曾盛.算法权力规制的法律问题研究[J].福建金融管理干部学院学报,2022(1):53.

[4]赖丽华,邱琳.数字经济发展视域下算法“霸权”的法律规制研究[J].企业经济,2023(5):151.

[5]贺螾,张旖华,邓沛东.风险视角下数字平台私权力的法律规制[J].西安财经大学学报,2023(5):105.

[6]任颖.算法规制的立法论研究[J].政治与法律,2022(9):98.

[7]路于婵.算法权力的证成与法律规制[J].武汉交通职业学院学报,2023(4):41.

[8]海伦·尼森保姆.场景中的隐私:技术、政治和社会生活中的和谐[M].北京:法律出版社,2022:122.

[9]杜永欣,周茂君.透明化生存中的自主性构建:算法推荐的隐私问题与规制路径[J].未来传播,2022(6):10.

[10]寿晓明,曹贤信.生成式人工智能场景中个人信息保护的风险、逻辑与规范[J].昆明理工大学学报(社会科学版),2024(1):21.

[11]何昊洋.大数据杀熟背后的平台私权力及其法律矫正[J].重庆大学学报(社会科学版),2023(6):220.

[12]张慧.算法时代传统隐私理论之困境与出路:以个性化推荐为场[J].重庆社会科学,2021(2):125.

[13]KRUPIY T. A vulnerability analysis: theorising the impact of artificial intelligence decision-making processes on individuals, society and human diversity from a social justice perspective[J]. Computer Law & Security Review, 2020,38(9):7-8.

[14]温昱.算法权利的本质与出路:基于算法权利与个人信息权的理论分疏与功能暗合[J].华中科技大学学报(社会科学版),2022(1):59.

[15]林洹民.自动决策算法的法律规制:以数据活动顾问为核心的二元监管路径[J].法律科学,2019(3):43.

[16]张凌寒.风险防范下算法的监管路径研究[J].交大法学,2018(4):49.

[17]丁晓东.论算法的法律规制[J].中国社会科学,2020(12):138.

[18]郭春镇,勇琪.算法的程序正义[J].中国政法大学学报,2023(1):164.

[19]张慧.信息隐私和个人信息功能定位的再区分[J].北京社会科学,2022(1):107.

[20]BINNS R, VEALE M. Is that your final decision? Multi-stage profiling, selective effects, and Article 22 of the GDPR [J]. International Data Privacy Law, 2021,11(11):323.

[21]凯伦·杨,马丁·洛奇.驯服算法:数字歧视与算法规制[M].上海:上海人民出版社,2020:57.

[22]蔡星月.以算法规制算法[J].华中科技大学学报(社会科学版),2023(4):111.

[23]高艳东,王莹.数字法治:数字经济时代的法律思维[M].北京:人民法院出版社,2021:46.

[24]王莹.算法侵害责任框架刍议[J].中国法学,2022(3):66.

[25]林洹民.自动决策算法的风险识别与区分规制[J].比较法研究,2022(2):188.

[26]张慧.自主算法隐私保护的规范与技术分析[J].兰州学刊,2021(3):120.

[27]韩世鹏.生成式人工智能算法备案的法律属性与控制路径[J].河南财经政法大学学报,2024(2):119.

[28]苏宇.算法解释制度的体系化构建[J].东方法学,2024(1):81.

[29]闫海,王洋.算法规制工具的功能悖论及其法治实现[J].法治研究,2022(2):105.

[30]COBBE J, SINGH J. Reviewable automated decision-making[J]. Computer Law & Security Review, 2020,39(11):3.

[31]许可.算法规制体系的中国建构与理论反思[J].法律科学,2022(1):124.

[32]张凌寒.权力之治:人工智能时代的算法规制[M].上海:上海人民出版社,2021:69.

[33]邵红红.破解算法侵权责任界定的中立性难题:以“算法推荐第一案”为切入点[J].新闻界,2023(9):71.

[34]丁国峰,寿晓明.生成式人工智能算法的法律风险及其规范化防控[J].云南大学学报(社会科学版),2024(3):122.