

# 算法解释权：赋权争议、适用逻辑与适用机制

本照清

(郑州大学 法学院, 河南 郑州 450001)

**摘要：**自动化决策算法在商业领域的应用日渐广泛，算法治理成为法学研究的热点。在算法治理探索的路径上，算法解释权的引入具有重要意义。我国《个人信息保护法》第 24 条第 3 款规定了具有中国特色的算法解释权，但在权利构造上依旧模糊。因此需要围绕算法解释权的构成要件展开分析，从主体、时间和内容三个方面着重讨论算法解释权的适用机制，有助于为算法解释权适用提供充分的理论依据和具体的操作指引。

**关键词：**自动化决策；算法解释权；算法黑箱

中图分类号：D90 文献标识码：A DOI：10.19358/j.issn.2097-1788.2024.03.010

引用格式：本照清. 算法解释权：赋权争议、适用逻辑与适用机制 [J]. 网络安全与数据治理, 2024, 43(3): 63-68.

## Algorithm explanation rights: controversial empowerment, application logic and application mechanisms

Ben Zhaoqing

(Law School, Zhengzhou University, Zhengzhou 450001, China)

**Abstract:** Business automation decisions are widely used in the commercial field, and algorithm governance becomes a hot topic in theoretical research. In the path of algorithm governance exploration, the introduction of algorithm explanation rights is of great significance. Paragraph 3 of Article 24 of Personal Information Protection Law stipulates the right to explanation of automated decision-making with Chinese characteristics, but the institutional construction of rights is still vague. Therefore, it is necessary to analyze the constituent elements of the right by taking account of the subject of rights, explanation content of rights and application time of rights. It can help to provide sufficient theoretical basis and specific operational guidelines for the application of the algorithm explanation rights.

**Key words:** automated decision-making; algorithm explanation rights; algorithm black box

## 0 引言

2022 年我国数字经济规模达到 50.2 万亿元，作为数字经济的核心技术，算法对人类的生存、生产和生活方式产生了全方位的深远影响。飞速发展的算法技术在便利生活的同时也带来新的社会问题，如算法黑箱、信息茧房等。为解决算法引发的新型规制难题，国内外学者围绕算法治理积极开展研究，提出新型权利算法解释权来规制算法自动化决策的。尽管算法解释权存在争论，但算法解释权在破解算法黑箱，抑制算法权利上有着重要意义。通过深入剖析算法解释权的适用逻辑和适用机制，平衡效率与公平，让算法解释权在算法治理中发挥其应有的功能。

## 1 商业自动化决策的现实隐忧

随着第四次科技革命的开启，受益于机器学习和大

数据技术等数字技术的飞速进步，算法已经深入社会各个领域，如自动化交易系统、医疗诊断工具、个性化预测分析等。算法从辅助决策逐步发展为主导决策。与人工决策相比，算法决策的稳定性、准确性、效率性十分突出，其“绝对客观”，即便是以专业、客观、公平为内核的司法决策也难以企及。然而随着自动化决策算法的广泛应用，其弊端逐步显露，算法歧视、大数据杀熟等现象对其标榜的客观性和公正性造成冲击。

### 1.1 技术维度：固化算法黑箱

自动化决策算法是以数据作为驱动的，它需要依靠庞大的数据集来学习数据模型并根据数据做出预测。无监督的机器学习算法广泛应用于用户细分、推荐服务等场景。机器学习的独特之处在于它使用任意的黑盒函数来作出决策。这些黑盒函数可能极其复杂，并且具有由

数百万个相互依赖的值组成的内部状态。关键的是，操作人员不需要理解算法产生的决策规则的基本原理<sup>[1]</sup>。在算法的输入端和输出端中间，存在外界既无法观测又无法理解的隐层，也就是算法黑箱。算法会存在错误代码，这些错误代码或是技术原因导致，或是科学原因导致，或是法律原因导致。算法无法被观测、理解，算法决策无法被追踪，算法决策合法与否无法被确定，就会导致新的安全风险和责任空白。算法黑箱造成的不透明性成为自动化决策算法治理的关键性挑战之一。

如果算法成为企业经营的核心技术，会具有严格的保密性，可能会限制对其内部的访问。以定价算法的使用为例，企业运用自动化决策算法，在对“高价值”消费群体进行高定价的同时对价格敏感消费群体保持低定价，以实现企业整体利润的提升。对经营者来说，实行定价算法的成本极低，但是带来的收益却极高，定价算法就会成为企业攫取利润的强力工具。虽然算法提高了企业的整体利润，但是消费者的合法权益受到了危害，而且这也对分配公平造成隐患。不仅如此，算法所具有的隐蔽性为经营者提供垄断合谋提供了便利，经营者甚至不需要公开商讨垄断协议，只需要使用算法进行价格共谋就可以实现限制竞争的目的。算法的不透明性成为经营者攫取巨额垄断利润的一堵保护墙，人为增强算法的不透明性，固化算法黑箱。

## 1.2 价值维度：异化法的价值

算法歧视下的差别待遇。行为画像是从广告业发展而来，一开始它只是为了精准定位潜在客户，然而随着机器学习算法的进步，这种对人类行为相类似数据分析并进行定制化的算法开始渗透进社会生活的方方面面。互联网消费中定价算法造成的“大数据杀熟”现象逐渐为人们所关注。2021年北京消费者协会进行了互联网消费“大数据杀熟”问题调查，在对16个平台进行的32个模拟消费体验样本中，有14个样本新老账户的价格不一致<sup>[2]</sup>。经营者收集消费者的特征信息数据，并借助自动化决策算法分析每个消费者的消费习惯。当算法认定该用户“高价值”后，算法控制者把更高价格的商品或服务推荐给用户，用户在不知情的状态下被动承担额外成本。相比传统商业实践中的价格歧视现象，算法决策分析技术极大地加强了价格歧视的隐蔽性。以电子商务领域为例：首先每位用户在购买商品或服务时处于相对独立的状态，用户间的沟通受限，算法导致的价格差别难以被发现；其次算法价格歧视是面对整个平台的所有用户，任何用户只要满足算法要求的行为参数，就会自动触发系统运行实现精准定价。而绝大多数的普通用户难以感知到算法的运行逻辑，并对产品的价格构成和浮动

状况也不了解，只能被动地支付更高的价格。

算法推荐加剧社会极化。《网络共和国》开篇就提到“个人日报”现象，其本质就是一种信息选择行为。自动化决策算法应用到信息推送服务领域后，这种信息选择无时无刻不在发生。在互联网上，用户甚至不需要主动去搜索，网络平台仅仅根据用户过往的习惯就可以推测出用户的兴趣点，然后增加对这些内容的推送力度，以达到增强客户黏性的目的。这种长时间的主动和被动的信息选择，为用户构建出了一个信息茧房<sup>[3]</sup>。以网络社交平台中的两性问题为例，网络上流传一则视频，内容是一名男性发现，在一则男女吵架的抖音视频评论区，其妻子和其本人所看到的评论内容截然相反。该视频清楚明白地展现了当下的信息茧房。自动化决策算法根据用户既往数据推测出用户可能会希望看到的信息，用户可能会支持的立场，或者可以激发用户讨论兴趣的热点话题，再通过算法个性化推荐给用户，以此达到取悦用户、增加用户使用时间的目的。然而长期的信息茧房使得偏见成为“真理”，人们的思想愈发极端，对营造良好的网络空间氛围形成障碍，甚至会影响到现实生活空间的社会安定。

## 2 算法解释权的赋权之争

为应对算法时代人类面对的诸多社会新型风险，学者开始积极研究算法治理问题，主要集中在三方面：第一，以数据安全角度切入，通过保证数据安全降低算法依赖的方式进行算法治理；第二，从算法伦理角度进行算法治理；第三，从技术角度，打开黑箱的外壳，实现有意义的算法透明度。为解决算法黑箱效应带来的难题，“算法透明”成为提高算法的可问责性的重要工具<sup>[4]</sup>。但是透明概念存在局限性，其可能会导致系统进行策略博弈<sup>[5]</sup>。为弥补这种局限性，有学者提出向受到算法决策重大影响的人们提供解释说明，以降低算法黑箱的负面效应，也就是算法解释权的起源。从学理上的“权利”到法定权利，算法解释权的落地历经争论。支持者和反对者都提出了有力的论证。双方争议的焦点主要在于算法是否具有可解释性和算法解释权与商业秘密保护的冲突。

### 2.1 算法是否具有可解释性

算法可以被解释是算法解释权成立的前提条件。算法的可解释性，是算法治理理论中的一个技术概念。反对者认为能够被解释的只有简单初级的算法，而复杂的机器学习算法，尤其是深度学习算法，其模型输出结果产生的原因是很难被解释的<sup>[6]</sup>。自动化决策算法是在进化计算、模糊逻辑、神经网络发展成熟的基础之上，融

合发展而来的方法，经历市场的迭代和技术人员的不断优化后，包含着复杂多层的网络结构<sup>[7]</sup>，即使是算法的开发者也很难清楚地解释。因此反对者坚持认为：因为技术门槛，即使公开算法，也并非所有人都能理解。相反，持肯定态度的学者则认为绝大多数算法具有可解释性<sup>[8]</sup>。算法鸿沟的存在值得重视，但是如果仅用专业性过高作为无法解释的理由，难免有些牵强。

莱斯格曾做出“代码即法律”的夸张表达，甚至有学者认为算法会对法律进行一定程度上的替代。虽然算法和法律属于不同范畴，但二者存在相似之处。对法律进行解释是法学理论研究的重要内容。算法解释也存在演变成一门专业技术的可能。目前计算机领域已经有规则提取技术的研究了<sup>[9]</sup>。如以实例、可视化的图标，简单明了的语言向一般人讲解，或者通过改变输入数据，展示对输出结果的影响程度让大众理解输入数据和输出结果间的关联关系。张志铭教授曾指出，法律解释就是对法律文本意思的理解和说明，其中“理解”是解释者对法律文本意思的内心把握，“说明”是对理解的外在展示，是对法律解释的结果要用某种恰当的方式、正当的方式表现出来<sup>[10]</sup>。对算法规则进行解释，实际上也是对算法决策过程理解和说明。

## 2.2 算法解释权与商业秘密保护是否存在冲突

算法规则的背后是商业利益。算法解释权要求算法使用者对算法的决策逻辑进行可理解的解释说明，这种要求是否会侵犯到企业的商业秘密，是争议的焦点之一。

否定者认为，算法解释权的行使会侵犯使用者的商业秘密<sup>[7]</sup>。算法解释权的行使必然会导致商业秘密的泄露，这将影响企业的核心竞争力。甚至有学者主张全然废除算法解释制度。算法解释权的立法目的在于保证一定程度上的算法透明度，这与有严格保密需求的商业秘密保护制度存在天然的价值冲突。在这样的冲突面前，非此即彼的取舍是不理性的，法官的自由裁量和利益衡量或许能够提供一些帮助，但是核心在于法律能否通过对算法解释规则的解释与修正实现冲突的消解。

赞成者则主张，通过对算法解释权的适用范围及启动条件的界定，可以使得算法解释与商业秘密间的冲突得到协调。算法解释权并不要求算法使用者公布算法的源代码，仅仅是追求合理的透明度，因此可以通过控制算法解释权的适用范围等手段来规避算法解释权对商业秘密造成侵害的可能。例如在“陈某诉杭州某软件服务公司网络服务合同纠纷案”<sup>[9]</sup>中，虽然软件服务公司算法过滤系统已经被法院认定是商业秘密，但法院仍然采用中立专家小组对其审核，并对平台提出解释要求：就算法过滤系统的一般性原理及具体操作，给原告做出合理

的解释。算法确实可能成为商业秘密的保护对象，但商业秘密也会受到法律的限制。

## 3 算法解释权的适用逻辑

在上述焦点争议的不断探讨中，算法解释权在学理上的可行性得以显现。探索规制算法乱象的路径有许多，基于算法本身的复杂性、多样性以及算法应用的广泛性，设置算法解释权是制约算法权力的破局之法，也是我国法律面对此种问题的积极应对。

### 3.1 制约算法权力的必然选择

在数字社会中，算法成为法律，代码成为规则，代码为人的行为提供指引并设定限制<sup>[11]</sup>。然而算法借助技术中立的掩盖，从技术规则异化为社会权力，在既存的法律规制体系中逃逸。算法权力就是某个主体以算法为工具，对他人或社会实施的影响力和控制力<sup>[12]</sup>。刑不可知，则威不可测。算法权力借助算法黑箱特性悄然发展，决策者和被决策者之间存在着巨大的信息差异，双方地位严重失衡。算法权力依靠强大的数据资源和自身技术优势，间接指引或直接代替人做出决策，以达到对社会进行控制的目的。它的出现，某种程度上打破了原有的个人权利与国家权力二元结构基础上的理论范式。

因此，算法解释权的设立不仅仅是对自动化决策算法这一技术的规制，同时也是对算法权力的规制。首先，它作为算法基础权利，以权利制约权力的方式对算法权力进行规制；其次，它以分配给算法主体责任的制度性安排进行算法规制；最后，它要求算法具有可解释性，是从算法设计源头对算法进行规制<sup>[13]</sup>。

### 3.2 《个人信息保护法》算法解释条款的支撑

我国法律在自动化决策算法的治理中不断探索。《电子商务法》以“尊重和平等保护消费者合法权益”作为个性化推荐算法的底线。国家标准《信息安全技术个人信息安全规范》以要求“个人信息控制者应向个人信息主体提供针对自动决策结果的投诉渠道，并支持对自动决策结果人工复核”来保证个人权益。《数据安全管理办办法》（征求意见稿）要求互联网平台运营者建立算法策略披露制度，同时也明确了互联网平台经营者在使用定价算法和个性化推荐算法时需要遵守的法律规定。

我国2021年11月1日生效的《中华人民共和国个人信息保护法》（简称为《个保法》）第24条是关于利用个人信息自动化决策的规制的规定，该条第1款确定了自动化决策应当遵循透明公平公正的原则；第2款要求个人信息处理者“应当提供不针对其个人特征的选项，或者向个人提供便捷的拒绝方式”，赋予个人在自动化决策下的个性化营销的选择权；第3款阐明了如果自动化决

策做出对个人权益有重大影响的决定，个人可以要求个人信息处理者对自动化决策作出说明，可以拒绝个人信息处理者仅通过自动化决策的方式做出决定，明确了个人权利。

《个保法》第 24 条第 3 款从实质上规定个人的算法解释权，对打破“算法黑箱”实现算法透明具有明显的积极意义。但是权利虽然得到明确，但是权利的构成要件仍较为模糊，这可能会引发法律适用的障碍。《互联网信息服务算法推荐管理规定》（征求意见稿）对算法解释权进行细化。但是最终版本相对于征求意见稿有较大幅度的修改。这也表明了算法解释权规范设计的复杂性以及落地的困难。故而对算法解释权的适用机制进行充分讨论十分必要。

#### 4 算法解释权的适用机制

目前各国虽已引入了对自动化决策规制的相关条款，但相关的法律规定较为概括抽象，存在解读的空间。为使算法解释权能够更加顺畅运行，需要明确规范算法解释权的适用范围和启动条件。在此将主要依据《通用数据保护条例》（General Data Protection Regulation, GDPR）和《个保法》为基础对算法解释权进行分析阐述（如表 1 所示）。

表 1 GDPR 与《个保法》关于算法解释权的条文梳理

权利内容	GDPR 序言和正文的相关条文	《个保法》的相关条文
解 释 时 间	第 13 条；第 14 条；第 15 条；第 35 条	第 17 条；第 24 条；第 55 条
	事后序言 71；第 22 条	第 24 条；第 48 条
解 释 主 体	第 2 条；第 4 条；第 22 条	第 2 条；第 24 条；第 48 条；第 73 条
解 释 内 容	第 13 条；第 14 条；第 22 条	第 17 条；第 23 条；第 24 条；第 56 条

##### 4.1 算法解释权的适用主体

在分析算法解释权的适用机制前，首先需要明确算法解释权是请求权。

GDPR 在一开始就强调“本条例旨在保护自然人的基本权利和自由”，明确“个人数据：是指与已识别或可识别的自然人相关的任何数据”；《个保法》第 2 条也强调保护对象是自然人的个人信息权益。虽然法人在法律上被拟制，但从数据主体的角度来看，二者在权利保护制度上有所区别。结合现行的法律法规，可以得出权利主体是自然人，不包括法人。权利主体不仅是自然人，还是自动化决策算法应用的相对人。算法解释权和算法公

开不同，其对算法黑箱的探索符合“掀开最小缝隙”理论<sup>[14]</sup>。算法只对受它不利影响的人打开大门，因此算法解释权的权利主体须是算法应用的相对人。权利主体还是通过自动化决策方式个人权益受到重大影响的自然人，我国《个保法》第 24 条明确“通过自动化决策方式个人权益受到重大影响”是算法解释权的前置条件，而且只要求决定是由自动决策方式做出，并未要求排除不存在人工干预，这与 GDPR 第 22 条要求的决策必须“仅”由算法做出不同。首先关于自动化决策的含义，《个保法》第 73 条第 2 款中给出了明确规定。GDPR 中也有规定：“自动化决策是指，包括识别分析在内的，在没有人为直接参与的情况下自动作出决策的模型。”其次关于重大影响的具体标准，《个保法》并没有给出详细的解释。GDPR 第 22 条中有与此类似的描述“legal effects concerning him or her or similarly significantly affects him or her”。在理解此处时，可以参考欧盟第 29 条工作组发布的《关于自动化个人决策目的和识别分析目的准则》（下文简称为《准则》），《准则》点明法律影响是法律权利受到影响，如法律地位或者是合同下的权利受到影响；并且列举了自动化决策算法在信贷和广告领域应用的例子阐明何为近似重大的影响，最后提到这种近似重大的影响既可能是积极的，也可能是消极的。

《个保法》第 24 条规定，有解释义务的是“通过自动化决策方式作出对个人权益有重大影响的决定”的个人信息处理者；GDPR 第 2 条点明其调整对象是个人信息处理行为，第 22 条明确算法解释权仅适用于通过自动方式处理数据的情形。因此算法解释权的义务主体是通过自动化决策方式处理个人数据的个人信息处理者。义务主体既包括法人，也包括自然人。结合社会实践分析，自然人实际参与算法自动化决策多是职务行为，个人很难单独通过算法进行自动化决策，因此自然人很难成为义务主体。但是这并不意味着将自然人从算法解释权的义务主体中剔除，自然人仍然存在成为算法解释权义务主体的可能。诚如前文所述，自动化决策算法的商业化应用十分广泛，已经形成相对完善的产业链，在这条产业链上，存在单纯的数据拥有者，算法开发设计者和算法使用者等身份，这些身份或相互独立，或相互重合。虽然《个保法》第 4 条第 2 款规定“个人信息的处理包括个人信息的收集、存储……”，但单纯的收集和储存行为很难满足“通过自动化决策方式作出对个人权益有重大影响的决定”这一条件，因此单纯的数据拥有者不能成为算法解释权的义务主体。当算法决策的使用者与算法的开发者身份相重合时，自动化决策算法的使用者身份便是普通数据控制者、处理者和算法开发者的限缩。

但是存在算法决策的使用者与自动化决策算法的开发者身份相互独立的情况，若自动化决策算法的直接使用者没有能力对算法进行解释，此时算法的开发设计者就应当与算法决策的使用者共同承担算法解释义务，以保证算法解释权条款的正常运行。

#### 4.2 算法解释权的解释内容

《个保法》第24条仅仅给出了“个人有权要求个人信息处理者予以说明”的要求，对需要说明的具体内容并没有明确规定，此处便存在对解释内容明确的必要性。自动化决策算法本身是计算机指令系统的形式存在的，对于绝大多数非专业的权利主体来说，算法的数学模型和技术知识是难以理解的也是不需要理解的，权利主体真正需要了解的是“决策”。因此算法解释权要求解释的内容不是算法本身，而是做出决策的逻辑，也就是算法运行的逻辑路径。从《个保法》第24条第3款的法律文本可见，个人信息处理者要说明的是“决定”。结合《个保法》第24条第1款的规定可知，立法者强调的是决策的透明度而不是算法本身的透明度。因此义务主体应找到合适的方式将算法运行的逻辑或决策逻辑告诉权利主体，而无需总是尝试解释其使用的复杂算法或者披露出全部算法。

算法解释权要求义务主体解释的内容是运用算法作出的决定，但是理解算法决策过程需要解释算法的技术帮助。首先，算法决策所适用的数据集、输入和输出的特征参数和个案预测的结果等方面可以描述决策算法相对静止的状态；其次决策算法开发过程中有算法模型的关键信息，如模型开发的初始需求和任务、数据标签的体系和模型优化记录等；最后面对不同的解释对象，需要准备不同的解释工具。采用自然语言、可视化图标、规则型解释器、示例等工具对决策过程进行解释，有助于缺乏技术背景的用户理解算法决策逻辑。而对于有一定技术基础，具备一定专业知识的用户来说，给出模型的算法公式、深度学习的优化方法和注意力机制、模型判断的依据权重等可能更有效果。除去在解释技术上的优化，也可参考《个保法》第17条第1款第2项、第23条的规定要求处理者告知“个人信息的处理目的、处理方式，处理的个人信息种类、保存期限”，第56条的“个人信息的处理目的、处理方式等是否合法、正当、必要”规定，算法解释权的解释也可以决策目的、决策逻辑、决策后果为详细内容。

#### 4.3 算法解释权的适用时间

对于算法解释权适用的时间要求，《个保法》和GDPR都没有给出明确详实的规定。因为在不同的解释时间，解释的内容标准存在差异，解释的目的也存在差别，

所以算法解释权的适用时间需要明确。以自动化决策的发生时间为界限，可将算法解释分为事前解释和事后解释两种情况。

《个保法》第17条第1款第2项“在处理个人信息前”的规定，第55条“事前”的规定，《算法推荐管理规定》的第16条规定，GDPR第13条“at the time when personal data are obtained”，第13、14、15条中的“the envisaged consequence”，第35条的“prior to”的规定，均可以被解读为，个人信息处理者在准备通过自动化决策方式处理个人数据时，要对可能受到算法决策效果影响的相对人进行事前解释，以保障自动化决策算法运行的透明度。事前解释发生在自动化决策算法开始运行前，所以此时所做出的说明仅仅是系统性的说明。此时需要解释的内容包括“自动化决策系统的逻辑、意义、预期后果和一般功能”<sup>[15]</sup>。

《个保法》第24条的规定更偏重决策结果的公平公正，并且根据法律条文的具体文义可以分析出“个人有权要求说明”是发生在“作出对个人权益有重大影响的决定”之后的；GDPR序言第71段中的“the decision reached after such assessment”和GDPR第22条关于拒绝自动化处理的权利规定，均可以被解读为，个人信息处理者在通过自动化决策方式处理个人数据后，要对遭受算法决策的不利影响的相对人进行事后解释，以保障自动化决策算法运行的结果公正。另外在《准则》对GDPR第22条的解释中明确提到，个体拥有“对决定提出质疑的权利”。《准则》虽没有法律效力，但在一定程度上代表欧盟对于GDPR的官方立场<sup>[16]</sup>。事后解释的起点在于自动化决策算法做出了对个人权益有重大影响的决定，因此事后决策聚焦在对“具体决策”进行解释，这种情况下解释的内容包括算法模型的基本原理和个体情况，如在制定出决策时所考虑的主要特征、依据的信息的来源和相关性等。

### 5 结论

作为自动化决策算法发展应用的伴生产物，黑箱效应引发的法律风险始终被重点关注。算法解释权作为一种算法新型权利，肩负着增强算法透明度和算法可理解性的重担。算法解释权从学理创设到步入法律文本，其内涵也越发清晰。算法解释权在提高算法透明度和增强算法可责性上可大有作为。《个保法》中第24条和48条已经对算法解释权的启动条件和权利主体做出了初步的制度画像，但算法解释权的解释内容、解释时间依旧模糊。对算法解释权的主体、内容和时间标准进行探讨，对算法解释权的适用具有重要价值。

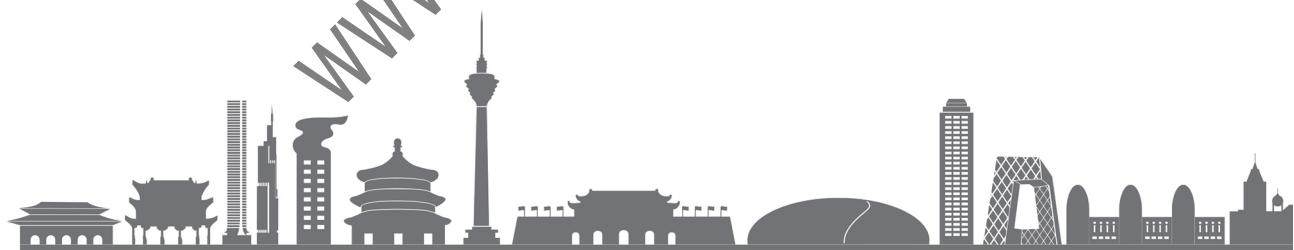
## 参考文献

- [1] MATTHIAS A. The responsibility gap: ascribing responsibility for the actions of learning automata [J]. Ethics and Information Technology, 2004, 6 (3): 175 – 183.
- [2] 北京市消费者协会网站 [EB/OL] (2022-03-01) [2023-09-15]. [http://www.bj315.org/xfdc/202203/t20220301\\_32335.shtml](http://www.bj315.org/xfdc/202203/t20220301_32335.shtml).
- [3] 刘耀华. 治理算法黑箱：建构算法科技与新闻伦理的道德智能体 [J]. 大众文艺, 2023 (18): 59 – 61.
- [4] 马丁·艾泊斯, 玛尔塔·坎泰罗·伽米托. 算法治理：法律和道德挑战 [M]. 姚前, 冯蕾, 译. 北京: 中国金融出版社, 2022.
- [5] ANANNY M, CRAWFORD K. Seeing without knowing: limitations of the transparency ideal and its application to algorithmic accountability [J]. New Media & Society, 2018 (20): 973 – 989.
- [6] 辛巧巧. 算法解释权质疑 [J]. 求是学刊, 2021, 48 (3): 100 – 109.
- [7] 尚荣华, 等. 智能算法导论 [M]. 北京: 清华大学出版社, 2021.
- [8] 何新新. 算法解释权的证成与限定 [J]. 大连理工大学学报 (社会科学版), 2023, 44 (3): 65 – 73.
- [9] 孔祥维, 唐鑫泽, 王子明. 人工智能决策可解释性的研究综述 [J]. 系统工程理论与实践, 2021, 41 (2): 524 – 536.
- [10] 张志铭. 法律解释原理 (上) [J]. 国家检察官学院学报, 2007 (6): 50 – 61.
- [11] [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.
- [12] 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 – 132.
- [13] 董少平, 马丽欣. 算法性别歧视视角下算法权力异化的治理路径 [J/OL]. 南京邮电大学学报 (社会科学版), 1 – 11 [2023-12-24]. <http://doi.org.libproxy.v.zzu.edu.cn/10.14132/j.cnki.nysk.20230926.001>.
- [14] 刘琳. 算法解释权与商业秘密保护的冲突化解 [J]. 行政法学研究, 2023 (2): 168 – 176.
- [15] 张凌寒. 商业自动化决策的算法解释权研究 [J]. 法律科学 (西北政法大学学报), 2018, 36 (3): 65 – 74.
- [16] 丁晓东. 基于信任的自动化决策：算法解释权的原理反思与制度重构 [J]. 中国法学, 2022 (1): 99 – 118.

(收稿日期: 2023-12-26)

## 作者简介:

本熙清 (1999-), 女, 硕士研究生, 主要研究方向: 法学理论、计算法学。



## 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