

商业算法自动化决策的私权构建与实现

孙跃元 许建峰

摘要: 商业算法自动化决策的私权建构是算法私人治理中的关键一环,个体赋权的核心在于保护“人”在技术威胁中的主体地位。算法权利能够保护决策相对人的正当利益,捍卫人的独立性与自治,其为概念标准之满足;具体的算法权利或被明文规定于法规中,或可用权利推定的方法从基础权利中派生出,并具有实现可能性,其为实证标准之满足。用“权利束”作为分析权利的实用主义方法,算法权利可被视为决策相对人享有的一系列旨在增强个体抵御算法风险能力的权利类型,具有开放特征和规范弹性。然而,权利理论若向实践转化,需借助算法“权利束”中各个权利的制度化实现,从算法知情权、算法解释权 and 算法决策拒绝权出发,构建权利实现路径。

关键词: 算法自动化决策;权利证成;算法权利束;私权构建

中图分类号: D923.99 **文献标识码:** A **文章编号:** 1003-0751(2024)02-0070-09

2021年1月英国竞争和市场管理局CMA发布的《算法:如何减少竞争和伤害消费者》报告指出,算法对消费者能产生直接危害。算法个性化可能以极其隐蔽的方式攻击弱势消费者或带来不公平的分配效果,形成算法操纵和算法合谋,影响竞争秩序^[1]。商业领域算法自动化决策的应用风险可类型化为内部风险和外部风险。内部风险有具体损害结果,可落入侵权责任、信义义务乃至刑事责任范畴^①。外部风险侧重算法行为端,在技术运用过程中向多数不特定群体扩散,具有模糊性、聚合性和累积性^②。目前世界多数国家均重点关注算法风险治理领域,形成以欧盟为代表的严格赋权模式和以美国为代表的算法问责模式。但因拘囿于技术鸿沟,这两种模式均存在一定的问题,如欧盟赋予个体严格算法解释请求权的模式存在治理效率低、增加算法服务提供者合规成本等局限性,美国算法问责制存在治理主体单一化、实操性差等缺陷。

有鉴于此,我国应在综合考虑公民权利、技术发展、市场活力等多种因素的基础上,提出符合我国国

情的算法治理方案。当下,我国现行法律偏重对算法的公共治理,对算法权利的保护则明显不足。学界也更多从功利主义角度考察政府涉算法公共治理政策的功能性问题,系统论述个体赋权的研究较少^③。本文聚焦算法自动化决策的私权建构,在澄清权利构建必要性和正当性的基础上,对权利定位、性质等本体论问题进行探讨,对比域外方案,探寻权利制度化路径,提炼权利建构的核心机理,为我国算法风险治理提供有益借鉴。

一、算法自动化决策私权构建的正当性证成

在互联网时代,新型权利主张不断涌现,如数据权利、被遗忘权等。新型权利的证成关涉是否存有被强制实施的可能,故而需要审慎对待。学界已有关于构建算法权利的主张,但更多是一种现象或策略性申明,并未从概念标准和实证标准予以论证。这里将在既有研究基础上进一步论证算法自动化决策

收稿日期:2023-08-20

作者简介:孙跃元,女,中国政法大学民商经济法学院博士生(北京 100088)。许建峰,男,最高人民法院信息中心主任,中国政法大学博士生导师(北京 100745)。

私权构建的正当性。

(一) 概念标准:被保护的合理性

1. 所保护利益的正当性

权利由社会经济状况和文化水平所制约,是社会关系中个体行为的自由和限度,能够激励个体积极性、创造性的发挥^[2]。权利是通往利益的手段,其目的是对物质的、精神的利益的享有^[3]。而享有利益的前提是利益本身存在保护必要性和正当性,并涉及保护目的及侵害防范等问题^[4]。对个体而言,公民的正当利益包括自由、平等和尊严,当自身合法权益受到算法自动化决策的侵犯时,有权提出救济。对整体社会而言,平等是社会的基本价值^[5]。由此,算法自动化决策领域的基本价值是公平和正义。

正义也是法律的基本价值追求,算法决策的正义价值是法律在社会实践中的目标。根据罗尔斯正义空间理论,正义包含正义价值和正义环境要素^[6]。依靠平台扩散权力趋势的算法决策可与几亿用户产生深度互动,任何微小的改变都能深刻影响用户的行为模式。因此,算法自动化决策创造了新的正义空间,表现为算法技术的社会意义和法律对算法价值分析的方式^[7]。算法正义的要求意味着需要从价值角度切入,解决算法决策应用中的风险,保护公民正当利益。算法决策内嵌于网络社会整体架构,成为事实上的网络空间规则输出者。个人遵守其规则的动力并非来源于国家权力的控制,而是广泛而深入的算法技术统治。此前,只有国家主权才会有如此普遍的决策权力^[8]。美国管理和预算办公室(OMB)2021年11月发布的《人工智能应用的监管指南》备忘录中提出了10条原则,供各机构在考虑“人工智能应用的设计、开发、部署和操作的监管和非监管方法”时加以权衡,其中的关键原则之一就是“公平和不歧视”^[9]。与之类似,欧盟《可信人工智能伦理指南》提出人工智能系统的开发、部署和使用需要满足“多样性,非歧视和公平”“环境和社会福祉”等7项关键要求才能成为值得信赖的人工智能^[10]。

在信息社会背景下,算法自动化决策的独特性加剧了技术与法律的冲突。技术本质上是使事物暴露其本质特征的解蔽手段。技术遵循一定的客观规律和发展逻辑,服务于人类改造客观世界的活动,但并非具有完全的客观性,而彻底独立于人类的主观世界。有学者认为,技术和科学都有一定的意识形态属性^[11]。在互联网世界,技术中立意味着网络

中立。网络中立意味着网络平台等商业主体应无差别地对待用户,保持中立的立场提供信息服务^[12]。这与社会实践显然不符。法学的使命便是监督和审查技术可能带来的非理性后果,以及如何通过法治减损技术发展可能带来的风险^[13]。当下,基于机器学习的算法自动化决策正逐渐向基于神经网络的超级自动化决策演进^[14]。大数据和机器学习模型结合,使得自动化决策主动探索规律和识别模式,替代人类实现基础性的预测和鉴别工作。机器学习算法在训练数据的帮助下,可以自主学习规则并建立决策模型。由此,确立清晰的问责机制有极大的实践难度。黑箱决策的模糊性可以使恶意行为者通过商业自动化决策有效地隐藏偏见和歧视,从而追求非法利润或维护特殊利益,造成个人正当利益和社会实质正义受损。综上所述,算法自动化决策保护的利益,如个人的平等、自由和尊严等,具有正当性和必要性。

2. 保护个人选择的重要性

新型权利在满足利益正当性的初始条件后,个人的独立性以及对个人自治的保护则是更为重要的条件^[7]。在算法社会中,算法自动化决策替代人类决策广泛出现在社会生产生活中,个人尊严和个人自治正面临着严重威胁。一方面,个人尊严正面临前所未有的危机。个体正逐渐丧失对其产生重要影响的决策的理解和控制能力^[15]。大型互联网平台等算法服务提供者以难以察觉的方法,如针对行为评分、设定奖惩制度等,持续、单向地对用户个人监视。有学者认为,算法机器优势和架构优势必然会带来监控资本主义^[16]。当自动化决策开始刻画勾勒个体的品质、状况、行为举止时,人的尊严将受到技术的重大挑战,个体成为一个个由数据组成的标签。商业主体便捷而低成本地获得海量数据,建立起完整的评价体系,达到管理约束平台和利益最大化的目的^[17]。人的尊严是现代国家宪法的精神基础,捍卫尊严成为现代社会的普遍共识。这便要求,无论社会发展到何种阶段,国家、社会中的任何组织和个体都不可以将人视为工具,人是一切的目的。有学者认为算法将独立的个体视为分类节点,否认个人的真正含义,可能会造成“自我危机”^[18]。技术的扩张不能以侵犯人的自由和尊严为代价。如果仅认为个体是技术做决策时的工具和信息源,社会的根本价值取向会不断弱化。

另一方面,数字经济下的商业模式以个性化推荐为主要特征,隐含着算法对个人选择的操纵,个人

的自治能力和主体性地位都在遭受严重侵蚀。私法的价值理念强调意思自治和契约自由,侧重于保护个体的意思表示自由。当下,商业主体部署商业自动化决策的过程中,个体纯粹的意思自治已经很难实现。早在2016年美国白宫发布的《大数据报告:算法系统、机会与公民权利》中便对算法影响个人自治表示担忧,认为个性化服务难以发现的缺陷是缩小而不是扩大用户选择^④。算法辅助商业主体预测和调整个体的行为轨迹,以获取收益并控制市场。电商平台、社交媒体等成为第三主体,以浸入型的架构特点重塑社会关系。通过烦琐的平台规则和流于形式的知情同意机制,算法服务提供者实质上越来越体现着单方意志^[19]。因此,不可否认的是,超国家机构、跨国集团和大型互联网平台等私权利主体已经拥有部分公共权力^[20]。在我国《电子商务法》起草过程中,起草组也认可“电子商务平台逐渐具有了准公共产品的特征”^[21]。鉴于此,为了防止个体的主体性不断被算法技术所弱化和侵蚀,需要通过赋权来保护人的个体尊严和个体自治。

(二) 实证标准:可容纳性和可实现性

1. 法律体系的容纳性

新型权利除了满足权利的概念标准成为道德权利,还必须被既有的法律体系所容纳,具备实在法意义上的存在的基础和依据。新型法律权利是否存在,主要取决于它是否被法律体系中的特定规则所规定^[22]。有两种被规定的途径:一是明文规定于法律规则之中,二是以司法续造为基础渐进入法,也即从包含基础权利的法律规则中推导出新型权利。算法自动化决策中的个人权利已然被立法机构所认可,或被明文规定于法律规则中,或可用权利推定的方法从基础权利中派生出。具体而言,算法决策拒绝权(也称脱离自动化决策权、完全自动化决策拒绝权)和算法解释权已在我国《个人信息保护法》第24条第3款中明文规定。《个人信息保护法》赋予个人针对可能对自身权益有重大影响的自动化决策,有要求信息处理者说明解释和予以拒绝的权利。上述两种权利均属于事后救济的权利,能有效提高相对弱势的个体对算法决策的控制力,增强人的主体性。算法知情权(有学者称之为个人透明度权利^[23])属于事前要求处理者对特定决策进行具体而有意义解释的权利,被认为是实现算法透明度的高度体现,派生于《个人信息保护法》第24条第1款中对信息处理者保证自动化决策透明度的要求。2022年我国正式实施的《互联网信息服务算法推荐

管理规定》进一步要求算法推荐服务提供者应当以告知和公示的方式向用户说明算法基本原理、目的意图和主要运行机制等^⑤。这也是国家对确立新型算法相关权利的进一步肯定。有学者认为算法知情权等权利来源于隐私权、知情权、平等权等现有权利^[24],笔者认为该权利推定存在逻辑纰漏。我国《宪法》并没有规定知情权,它来自对附属规范的归纳性建构,是从人民主权原则、公民的批评建议权等宪法规定中归纳出公民享有知悉、获取公共信息的权利。同理,算法知情权也可用归纳性建构的方法,从《宪法》《个人信息保护法》和《互联网信息服务算法推荐管理规定》中归纳而来。综上所述,包含算法知情权、算法解释权和算法决策拒绝权在内的算法自动化决策权利体系不仅满足道德权利的条件,也能够被证明为现有法律体系所容纳。

2. 被实现的可能性

权利并非生存于真空中,若特定现实条件下没有实现权利的可能性,则不具有法律上的权利地位,并会减损法律的权威性。具体影响权利能否实现的因素主要包括社会成本和政治现实,目前争议较大的是算法解释权被实现的可能性。国内外的法学学者已逐渐达成共识,即商业自动化决策应当具备可解释性,决策相对人有了了解决策如何得出的权利^⑥。反对的观点认为,普通用户无法理解掌握算法的基础知识和技术,即使被详细告知决策技术原理,也会因技术鸿沟而使权利流于形式,因此算法解释权并无实际存在的意义。在实践中,过多地解释和披露算法可能会泄露商业秘密,加重平台负担。笔者认为,算法解释权能否具有现实意义,关键在于对算法解释的对象、标准等进行细化与区分,从而将其从技术方案转为可行的法律制度。在现有的用户私权保护路径中,商业自动化决策的算法解释权的功能在于修正具体的自动化决策。修正意味着能够以最小的成本获得正确的决策,这对于数字经济具有极高的制度价值。

同理,算法知情权也具有被实现的可能性和意义。透明度规则作为监管和系统设计的工具而非最终目标,不是揭示算法技术的工作原理,而是确保用户了解算法参考何种数值权重和占比而作出决定的。个人如果获得负面且消极的决策结果,并可能对自身权利产生重大影响,其就应当有权知道算法作出决策的理由。但决策系统能公开的设计内容涉及算法透明的边界。公开意味着风险,商业主体一般以防止泄露核心技术和保护商业秘密为由,拒绝

公开决策过程、决策目的等关键信息。若不对用户的算法知情权加以保护,会加剧用户对算法自动化决策的不信任,用户可能会拒绝接受商业主体收集其数据用于开展自动化决策运营活动,这显然不利于数字经济的良性循环。综上所述,算法自动化决策权利体系具有被实现的可能性,但需要解决的关键问题是如何将相关技术方案纳入法律规制的范围,这便需要准确可行的制度设计作为权利的实现路径。

二、算法自动化决策的权利属性

商业算法自动化决策的私权构建,离不开对其法律属性的判断。在法理学中,法律的介入意味着将具有争议的实体问题“转译成”法律上的问题^[25]。因此,需要在厘清算法自动化决策法律性质的基础上,对算法自动化决策权利进行准确定位,探讨最合理的法律介入方式。

(一) 算法自动化决策的权利定位

历史演进表明,权利是用来对抗权力的。权利以物为客体,是所有法律关系的终极目的^[26]。权利是一种正当的价值判断,具有正当性和有效性。本文所提出的算法权利的核心目的在于通过赋权增强个体抵御和控制技术风险的能力,增强个体在技术威胁中的主体地位。与算法的公共治理和监管路径不同,赋权并不是为了解决算法决策的功能性问题(如算法黑箱、算法妨害)。换言之,赋权并非通过监管激励或引导商业主体改进算法偏见和错误等功能性问题。个体赋权更强调自我决定和积极防御算法权力化的风险侵害。

韦伯认为,权力是排除他人意志的现实的支配力,能够影响他人。并非只有国家机器才具有支配力^[27],在互联网社会中,头部平台、大型跨国公司依托算法技术,以强大的组织力量对个体形成压倒性优势,深刻影响用户行为模式,已然具备支配力。有学者认为,算法技术统治是一种“准国家权力”性质的公共力量^[28]。不断迭代更新的算法技术所带来的控制力甚至超越了国家。国家制定的规则可能会被改变和推翻,但网络基本的运算法则和运行逻辑拘囿于商业最大化利益目标却很难改变。算法权力随着商业主体信息不对称的优势地位而有被异化的风险,从而削弱主体对数字的信任,弱化人的主体性地位。以算法权利体系中的算法决策拒绝权为例证,算法决策拒绝权的理论可追溯至“个人信息自

决权”,最早由德国学者施泰姆勒提出^[29]。自决权属于人类自由发展权的范畴,以自主决定个体的行为并有权选择是否由他人知悉为核心,对世界范围内个人数据保护立法影响深远。我国现行法律虽未将自决权规定为公民的法定权利,但从《个人信息保护法》第4章构建的个人信息权利体系中可推导出法律对个人自决利益的保护。在算法自动化决策领域,自决权更多指向对算法决策的使用和控制,要求应遵循个人的自由意志。为了实现自决,个人有权知悉决策的流程和原理。同样,个人获得不公正或者错误的决策结果时,应当被赋予允许拒绝的权利和获得合理的救济途径。是故,算法权利的核心目的在于防范算法权力对人的客体化和异化,增强个体对算法决策的控制力和自决力。

(二) 算法自动化决策的权利性质:基于“权利束”的构想

算法权利是指算法决策相对人在算法决策过程中享有的正当利益,它是一系列权利的总称,主要包括算法知情权、算法解释权和算法决策拒绝权。由于技术发展日新月异,未来可能有更多元的权利诉求,而权利之间的差异化必然导致冲突。如何在平衡各权利的利益冲突的同时,确保各权利从整体上保持统一的范式 and 一贯的价值内涵,需要寻求新的研究范式。笔者认为可以“权利束”为观察视角,分析算法自动化决策的权利属性。

“权利束”是指将一宗财产上发生的基于共同利益束点的多重权利关系集合在一起,像花束一样捆绑成束体^[30]。“权利束”理论来自于制度经济学,最初应用于无形财产权的保护问题,目的是对产权下的多项民事权利进行统合收纳^[31]。然而,制度经济学以效率优先和经济理性人为价值标准,与法学追求的公平正义价值不相融合。霍菲尔德提出的4种权利元形式,即主张(claim)、自由(privilege)、权利(power)、豁免(immunity)增加类推适用的前提,补强了“权利束”理论的合理性。作为新的权利理论研究范式,“权利束”能有效平衡权利间的冲突,主张权利的完整和独立。同时,可根据具体的场景对权利间的位阶进行动态调整^[32]。过去“权利束”主要描述财产权,前提是同一资源上存在多元的利益可能,且权利间具有利益性、享有性、排他性的特点。当下,学界引入“权利束”理论构建个人信息权益和数据权益,个人信息和数据之上存有复杂交织的人格利益、财产利益和公共利益,不完全是财产权。

鉴于此,尽管算法权利不完全是财产权,同样可以“权利束”理论为考察视角^⑦。“权利束”理论有助于解释不具有物理排他性的各种权益的集合,算法“权利束”高度涵盖多元权利,这些权利都趋向于促成“权利束”的中心价值,即“权利束”的束点。有学者认为,算法“权利束”的束点在于对算法权力的规制,属于功能相似性的经验描述^[33]。笔者认为该论点失之偏颇。算法权利的中心价值在于通过约束算法行为,增强个人抵御算法决策风险的能力。算法知情权要求算法服务提供者事前提供决策的一般信息,算法解释权要求运营商针对具体决策进行具体且有意义的解释,算法决策拒绝权赋予个人拒绝通过自动化决策的方式作出对个人权益有重大影响的决定。这些权利从对个人的人格保护出发,其共同目标价值都是通过赋予个人采取必要措施以对抗算法风险,保护人的主体性地位。这与个人信息“权利束”相类似,其束点在于保证主体个人信息自决。综上所述,从“权利束”理论角度考察,算法权利是决策相对人享有的一系列旨在增强个人对算法决策控制力的权利类型,具有开放结构和规范弹性。引入“权利束”作为分析权利的实用主义方法,既能够为算法权利中的各具体权利提供中心价值,同时也可以为具体权利的制度化设计提供更开放的空间。

三、算法权利的制度化实现

算法权利是用户启动的算法治理路径,私权利保护路径具有不可替代的作用^[34]。我国《互联网信息服务算法推荐管理规定》也通过权利路径,针对算法决策应用的不同场景配置相应的行为规则,赋予个人对抗算法决策的权利^⑧。这里拟从私人治理的视角,针对技术应用中的算法知情权、算法解释权和算法决策拒绝权,提出相对可行的体系化的制度化设计。

(一) 算法知情权的制度化实现:以利益平衡为原则

1. 算法解释权的独立价值与逻辑定位

学界有观点认为,算法解释不仅包括事后的解释、异议、更正等权利,也包含事前的告知。换言之,算法知情权可以被以算法系统功能为中心的解释权所吸收^[35]。算法知情权和算法解释权容易存在概念混淆,已有域外立法例将事前告知义务作为解释权的实现路径^⑨。笔者认为,从触发权利义务的时间节点、权利内容及价值导向等方面考量,算法知情权与算法解释权有一定功能暗合,但总体有逻辑分

殊的特点。

算法知情权属于传统上获取信息事实的知情权体系,是公民知情权在算法决策领域的自然延伸。法律约束以义务履行为支点,算法知情权也首先通过算法服务提供者的告知义务来体现。告知义务履行具体的时间节点是在算法自动化决策处理前,个体行使算法知情权对可能的风险及后果作出预判。实践中,算法服务提供者主要以用户协议的方式履行告知义务。这意味着告知对象是某类服务的受众群体,内容涉及系统的一般性信息等。《互联网信息服务算法推荐管理规定》提出特殊的算法推荐服务告知义务,告知内容包括算法推荐服务的基本原理、目的意图、运行机制等^⑩。与之不同的是,算法解释权属于事后解释,行使时间节点在自动化决策完成后。决策相对人有权要求对某项特定决策提供具体的信息,权利行使更像是“个案分析”的后果。解释涉及的问题也更为复杂,包括解释算法结果是否公平不歧视、数据比重是否需校准等主观性评价,还需对算法模型与特定决策结果之间的因果关系展开论证^[36]。由是观之,算法知情权有独立价值和功用,无法被算法解释权所吸收,为了算法自动化决策权利体系的完整,有必要将其独立成权。

2. 算法知情权的实践应用维度

实践中,应当将对决策相对人的算法知情权保护定位于提高算法透明度的要求中。算法透明度在学界一直存有争议。有学者认为,透明度在可行性和必要性上存在缺陷,是算法治理的必要而不充分的手段^[37]。也有学者认为,过于依赖透明处理原则会陷入“透明陷阱”(transparency fallacy),公民缺乏必要的专业知识来有意义地行使算法知情权^⑪。笔者认为,个体通过行使算法知情权了解决策内部程序和关键信息,以便数据主体后续提出质疑或发表意见,具有不可替代的积极作用。对算法服务提供者来说,知情权可以督促和规范其内部算法运作规则的标准化。有学者甚至将算法知情权称为公民的一项基本权利^⑫。如果自动化决策会对个体产生法律影响或近似重大影响,那么个体可以通过行使算法知情权要求算法服务提供者提供三方面的信息:算法活动的引入逻辑(logic involved)、数据处理的意义(significance)和设想的后续结果(envisage consequences)。告知引入逻辑(logic involved)的信息,意味着算法服务提供者应通过简单的方式,提供决策原理或标准、数据来源、处理逻辑等必要背景信息。例如,使用信用评分评估贷款申请时,金融机构

收集相关数据或由信用咨询机构直接提供,无论哪种来源,都必须告知相对人决策时所依据的表现特征、数据来源和相关性^⑬。意义(significance)和设想结果(envisaged consequences)表明个体有权知悉决策可能会对自身产生的影响。欧盟出台的《关于保护与自动化处理个人数据有关的个人的惯例》对其进行解释,要求实践中信息主体有权了解“数据处理的内在原理和后续结果(即可以导出的任何结论)”。算法服务提供者须重点告知算法对个人可能的不利影响,使得个人能够决定是否允许算法处理个人信息^[38]。

此外,为了便于个体在权益遭受侵害后及时获得救济,告知的内容需要包含用户的行为指引。例如实践中自动化决策算法应用场景有金融领域的信用评分等,算法决策直接决定相对人的财产利益。为了使用户可以预测到当事人双方的行为模式及后果,告知内容中应当具备行为指引。若经查明不当的决策结果产生原因是数据错误,则可以指引用户进行数据修改。

(二) 算法解释权的实现路径

算法技术黑箱使得决策相对人和决策部署者之间的技术障碍逐渐加大。许多监管政策的协调中心是要求完善算法解释权的实现路径,以提高决策透明度。被誉为“算法治理制度的核心”的算法解释权作为商业自动化决策事后问责体系中的重要环节,应从以下两个方面进行制度设计。

1. 明确商业自动化决策的算法解释内容

笔者认为,自动化决策算法解释内容的确定应首先依据算法分级分类将决策纵向分为一般影响和重大影响,再横向分为内部解释和外部解释。纵向方面,对于个案影响轻微的算法决策,应允许运营商自行设定个案解释的规则,避免公权力过多的强制性监管;对个体造成重大影响的算法决策,应设置更高的解释标准,以保护个人的核心权益^[39]。横向方面,内部解释主要负责向内部工作人员解释技术类问题;外部解释是从系统设计层面向用户和监管部门解释其运行机制,而非披露源代码^[40]。

总的来说,决策输出结果由算法和输入数据决定。因此,可解释性应当包括算法和数据两个方面。决策的内在逻辑、决策树、预测模型和标的权重等,属于算法层面;解释数据是否准确、完整且具有时效性,则属于数据层面。美国《平等信贷机会法》(The Equal Credit Opportunity Act)也同样做了规定,如果申请人的信贷申请被拒绝,信贷局必须提供作出决

定的关键理由^⑭。此时算法解释的内容不仅指法律和计算机科学的内部逻辑,而且意味着披露数据的收集方式、数据的清理和规范化、分析中使用的方法、超参数的选择和其他阈值等。当解释内容涉及商业秘密时,学界有不同的观点。有学者认为,算法解释权制度建构应以道义论权利哲学为基本遵循,应坚持个人数据基本权利优先于商业秘密保护的原则^[41]。笔者则认为,为了促进行业发展和技术创新,若解释内容涉及商业秘密,应优先保护企业合法的竞争利益。

2. 规范自动化决策算法解释权的行使程序

为了避免个人滥用解释权而给平台造成过重负担,需要明确算法解释权的适用前提,即对用户具有法律效力或重大影响。是否对当事人具有法律上的直接影响,根据法律规定即可判定。重大影响包括信用评分、就业申请、订立合同等。关于算法解释的标准,需要满足程序性要求和实质理性的要求。算法解释的提供程序也须满足程序标准与实质标准,包括解释的客体完整^[40]。商业自动化决策的算法解释应是一种要式行为,以书面形式作出。实践中,可以由监管部门统一提供格式解释文本,保证解释客体完整,为用户提供解释留存证据。在举证责任方面,决策相对人应举证商业自动化决策对自身有法律效力或重大影响,自动化决策部署者应承担该决策结果正确的证明责任。若证明决策结果错误,需收集大量自动化决策结果统计,对决策相对人的举证难度过高,造成维权成本过大,不利于权利的行使。有学者认为,应组织建立“算法安全委员会”,统一代替请求人接受算法解释,并对算法集中审查和问责^[42]。笔者认为,算法解释权作为事后问责体系的关键组成部分,不可让渡于其他组织。但考虑到我国目前并无对算法决策监管和审查的专门机构,可由具备掌握算法验证原理的专业人员组成第三方机构,设立行业标准和准入制度等,负责对算法决策的运行进行监管。

(三) 算法决策拒绝权的理论构建

1. 算法决策拒绝权的权利性质

我国《个人信息保护法》第24条第3款明确了算法决策拒绝权的请求权基础^⑮。算法决策拒绝权赋予个体对完全自动化决策的结果提出质疑和拒绝接受的权利,达到“人在环路”(human in the loop)的混合决策效果。算法决策拒绝权在欧盟一直存有“权利”还是“禁令”的争议。若将欧盟《通用数据保护条例》(下文简称GDPR)第22条^⑯解释为禁令,

意味着只要不属于第22条第2款的三种例外情形时,完全的自动化决策被一般性禁止。解释为权利意味着符合该条第2款的三种例外情形,则请求权不适用。显然权利路线对数据主体要求更多,禁令解释更能保护个体的数据权益^[43]。但GDPR以义务为本位,倾向于对数据权利的严格保护,同时也阻碍欧盟范围内人工智能技术的创新和行业发展^[44]。与之相反,英国采取了“一般允许、特殊禁止”的权利模式^⑩。我国《个人信息保护法》确定的反自动化决策明确表示为一项权利,这意味着采取了与欧盟相反的价值立场,包含着对完全的自动化决策的默认允许。故而,该法权的理论构造应体现出我国鼓励人工智能产业和刺激数字经济发展的政策立场。

2. 算法决策拒绝权的适用范围与适用主体

若算法决策拒绝权的范围不明确,个体在任何情况下均可提出算法决策拒绝权,会使企业不堪承受数据合规的成本,阻碍数字经济的发展。笔者认为,可从以下两方面对算法决策拒绝权的适用范围进行限定:一是该决策必须是完全的自动化决策系统的输出结果,而非人工干预的结果。若在最终输出结果时有人为发表评论或采取产生实质影响的措施,则不属于完全的自动化处理。但需警惕商业主体为了避免适用算法决策拒绝权,伪造人为干预的事实。例如,某惯常的人工行为对自动生成的主体画像结果未产生任何实质影响,则仍应被视作完全自动化处理^⑪。因此,商业主体必须确保对决策的干预是有实质意义的,并非象征性干预。具体表现为信息处理者在信息收集、信息分析及决策选择三个环节中有人工干预的义务。二是需要对决策相对人的合法权益有重大影响。我国《个人信息保护法》中的“重大影响”的内涵和外延并不明确,需要进行个案衡量^[45]。欧盟GDPR提出的“法律意义”和“近似重大意义”是指具有长期影响,或导致对个人的排斥或歧视^⑫。并在“序言”71中提供了构成近似重大影响的典型示例:“在线申请信用卡时被自动拒绝”或“无任何人为干预的在线招聘活动”。类似场景化的分析方法可以为我国所借鉴。

针对算法决策拒绝权的适用主体,笔者认为,应仅限于自然人。从保护对象来看,主要保护自然人的基本权利和自由,主体理应是自然人。我国2019年10月发布的《个人信息保护法(专家建议稿)》中明确,“信息主体指被识别出的特定生存自然人”^⑬,并且特别强调“特定生存自然人”,即意味着

享有权利和履行义务的主体是有出生事实的自然人,没有生命的死者自然不属于信息主体。根据国家标准《信息安全技术个人信息安全规范》中的规定,“个人信息主体(personal information subject)是个人信息所标识的自然人”^⑭。根据传统民法理论,关于人的自由发展的权利,其主体自然是自然人。法人是否可以成为算法决策拒绝权的主体,虽然目前各国法律没有直接规定,但根据上述我国立法和国家标准,法人成为该权利的主体缺乏立法目的和价值的支撑。

3. 算法决策拒绝权的内容

算法决策拒绝权的内容包括决策相对人有权不受对其自身产生法律或重大影响的商业自动化决策的限制,并自主决定是否从自动化处理中退出。行使拒绝权意味着允许决策相对人查阅算法、提出质疑和要求人工介入。因此,算法决策拒绝权的内容包括上述三个环节。拒绝一旦作出,对决策部署者即产生相应的法律后果:若该处理尚未开始,则不应开始;若已经开始自动化处理,应立即停止,并经决策相对人的要求进行人工干预,形成新的混合决策。

综上所述,算法决策拒绝权的制度化实现路径可遵循上述权利本体论的规范构造而展开,并且以平衡人工智能产业发展和个人权益保护为立场。尽管该法权来源于欧盟,但中国化的构建不可一味照抄欧盟禁令式立法,需要立足我国独特国情采取符合经济发展规律的建构模式。

结 语

商业平台利用自动化决策系统处理数据和管理资源,属于生产环节,需要适当的顶层法律设计介入规范和调整。我国现行法偏重对算法的公共治理,对算法权利的关注度明显不足。学界也更多从功利主义角度考察政府涉算法公共治理政策的功能性问题。在泛权利化表述的时代背景下,本文聚焦算法自动化决策的私权建构,系统论述个体赋权的理论证成、权利属性和制度实现。算法权利作为一种新型权利,满足权利的概念标准和实证标准。该新型权利的核心目的在于防范算法权力对人的客体化和异化,增强个体对算法决策的控制力和自决力。引入“权利束”作为分析权利的实用主义方法,算法权利可被视为决策相对人享有的一系列旨在增强个人控制力的“权利束”。针对具体权利的制度设计,本文从目前技术应用中的算法知情权、算法解释权 and

算法决策拒绝权出发,体系化提出相对可行的权利实现路径。法律蕴含着最低的道德尺度和价值判断,能够通过价值指引作用,弱化技术治理中过度技术化带来的非理性和非人道的问题。算法“权利束”具有开放结构和规范弹性,可以用温和的态度容纳更多新兴权利,以正确的立场来应对新技术带来的风险。

注释

①内部风险如因样本偏见、程序错误等原因致使算法系统输出错误结果等,可直接落入传统部门法既定的权益保护框架。②外部风险如社交媒体利用推荐算法造成受众“信息茧房”,征信评分体系用算法决策进行贷款、投资等金融方面的信用状况评估,隐含算法歧视,造成对公众群体的危害。③系统化论述算法权利的文献较少,主要有如下论文:谢琳、曾俊森:《个体赋权在算法决策治理中的定位》,《法治社会》2022年第1期;宋保振:《数字技术差异赋权风险的法律规控》,《郑州大学学报》(哲学社会科学版)2022年第4期;范玉吉、李宇昕:《从权力到权利:算法治理的路径》,《西南政法大学学报》2022年第1期。④See Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights, <https://www.benton.org/headlines/big-data-report-algorithmic-systems-opportunity-and-civil-rights>。⑤⑩参见《互联网信息服务算法推荐管理规定》第16条。⑥See Association for Computing Machinery Us Public Policy Council (Usacm): Statement on Algorithmic Transparency and Accountability, Jan. 12, 2017。⑦认同此类观点的文献有:苏宇:《算法规制的谱系》,《中国法学》2020年第3期;范玉吉:《基于“权力—权利”二元结构的算法治理研究》,《青年记者》2021年第7期;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》,《华中科技大学学报》(社会科学版)2022年第1期。⑧参见《互联网信息服务算法推荐管理规定》第19条至第21条。⑨See Danielle Vanderzanden: *The Artificial Intelligence Video Interview Act: Privacy Implications of Illinois AI Statute*, <https://ogletree.com/insights/the-artificial-intelligence-video-interview-act-privacy-implications-of-illinois-ai-statute/>。⑩See Frank Pasquale: *The Black Box Society: The Secret Algorithm That Control Money and Information*, Harvard University Press, 2015, pp. 75-76。⑪See Felzmann H., Fosch-Villaronga E., Lutz C. et al., *Towards Transparency by Design for Artificial Intelligence*, Science Engineering Ethics, Vol. 26/2020, p. 3334。⑫GDPR第14条第(2)款f项规定:“个人信息尚未从数据主体处取得的,关于数据来源的信息必须提供给数据主体。”⑬The Equal Credit Opportunity Act, <https://www.justice.gov/crt/equal-credit-opportunity-act-3>。⑭我国《个人信息保护法》第24条第3款规定:“通过自动化决策方式作出对个人权益有重大影响的决定,个人有权要求个人信息处理者予以说明,并有权拒绝个人信息处理者仅通过自动化决策的方式作出决定。”⑮欧盟GDPR第22条规定:“数据主体有权反对此类决策;完全依靠自动化处理(包括用户画像等);三是自动化处理的决策结果应当对个人有法律效果或其他类似重大影响。”⑯See Sandra Wachter, Brent Mittelstad, Luciano Floridi: *Why a Right to Explanation of Automated Decision-making Does Not Exist in the General Data Protection Regulation*, International Data Privacy Law, 2017, No. 2, p. 95。⑰⑱Article 29 Data Protection Working Party, Guidelines on Automated Individual Decision-making and

Profiling for the Purposes of Regulation, adopted on 3 October 2017, as last Revised and Adopted on 6 February 2018, WP251 rev. 01, <https://ec.europa.eu/newsroom/article29/items/612053/en>。⑲《个人信息保护法》(专家建议稿)附则第103条第1款规定:“信息主体,指姓名、出生日期、身份证件号码、个人生物识别信息、住址、电话号码等个人身份代码和记述等被识别出的特定生存自然人。”⑳国家标准《信息安全技术个人信息安全规范》3.1项规定:“个人信息(personal information)指以电子或者其他方式记录的能够单独或者与其他信息结合识别特定自然人身份或者反映特定自然人活动情况的各种信息。”其3.3项规定:“个人信息主体(personal information subject)指个人信息所标识的自然人。”

参考文献

- [1] CMA. Algorithms. How they can reduce competition and harm consumers. [EB/OL]. (2021-01-09) [2022-06-23]. <https://www.gov.uk/government/publications/algorithms-how-they-can-reduce-competition-and-harm-consumers/algorithms-how-they-can-reduce-competition-and-harm-consumers>.
- [2] 闫立东. 以“权利束”视角探究数据权利[J]. 东方法学, 2019(2): 57-67.
- [3] 武步云. 马克思主义法哲学引论[M]. 西安: 陕西人民出版社, 1992: 199-200.
- [4] 商希雪. 个人信息隐私利益与自决利益的权利实现路径[J]. 法律科学(西北政法大学学报), 2020(3): 71-85.
- [5] 丁晓东. 论算法的法律规制[J]. 中国社会科学, 2020(12): 138-159.
- [6] 罗尔斯. 正义论[M]. 何怀宏, 何包钢, 廖申白, 译. 北京: 中国社会科学出版社, 1988: 14.
- [7] 郑玉双. 计算正义: 算法与法律之关系的法理建构[J]. 政治与法律, 2021(11): 91-104.
- [8] 梁志文. 论算法排他权: 破除算法偏见的路径选择[J]. 政治与法律, 2020(8): 94-106.
- [9] White House Guidance for Regulation of Artificial Intelligence Applications [EB/OL]. (2021-7-24) [2022-09-23]. <https://www.ai.gov/white-house-guidance-for-regulation-of-artificial-intelligence-applications/>.
- [10] European Commission. Ethics Guidelines for Trustworthy AI [EB/OL]. (2019-04-08) [2022-09-14]. <https://ec.europa.eu/digital-single-market/en/news/ethicsguidelines-trustworthy-ai>.
- [11] 哈贝马斯. 作为“意识形态”的技术与科学[M]. 李黎, 郭官义, 译. 北京: 学林出版社, 1999: 83.
- [12] 郑玉双. 破解技术中立难题: 法律与科技之关系的法理学再思[J]. 华东政法大学学报, 2018(1): 85-97.
- [13] 韩大元. 当代科技发展的宪法界限[J]. 法治现代化研究, 2018(5): 1-12.
- [14] 唐林垚. “脱离算法自动化决策权”的虚幻承诺[J]. 东方法学, 2020(6): 18-33.
- [15] 萨普特. 被算法操控的生活: 重新定义精准广告、大数据和 AI [M]. 易文波, 译. 长沙: 湖南科学技术出版社, 2020: 154-163.
- [16] 张凌寒. 算法权力的兴起、异化及法律规制[J]. 法商研究, 2019(4): 63-75.
- [17] 胡凌. 数字社会权力的来源: 评分、算法与规范的再生产[J]. 交大法学, 2019(1): 21-34.

- [18]多梅尔.算法时代:新经济的新引擎[M].胡小锐,钟毅,译.北京:中信出版社,2016:28.
- [19]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(2):42-56.
- [20]齐延平.论人工智能时代法律场景的变迁[J].法律科学(西北政法大学学报),2018(4):37-46.
- [21]全国人大财经委员会电子商务法起草组.中华人民共和国电子商务法条文释义[M].北京:法律出版社,2018:98-100.
- [22]萨姆纳.权利的道德基础[M].北京:中国人民大学出版社,2011:58.
- [23]谢琳,曾俊森.个体赋权在算法决策治理中的定位[J].法治社会,2022(1):86-99.
- [24]范玉吉,李宇昕.从权力到权利:算法治理的路径[J].西南政法大学学报,2022(1):86-95.
- [25]陈景辉.算法的法律性质:言论、商业秘密还是正当程序?[J].比较法研究,2020(2):120-132.
- [26]刘杨.基本法律概念的构建与诠释:以权利与权力的关系为重心[J].中国社会科学,2018(9):112-135.
- [27]吕世伦,宋光明.权利与权力关系研究[J].学习与探索,2007(4):99-106.
- [28]许晓东,邝岩.算法权力的形成与风险治理[J].郑州大学学报(哲学社会科学版),2022(3):18-24.
- [29]张建文,李锦华.欧盟个人数据保护法上的反自动化决策权研究[J].重庆邮电大学学报(社会科学版),2019(2):23-29.
- [30]王利明.论数据权益:以“权利束”为视角[J].政治与法律,2022(7):99-113.
- [31]袁俊宇.个人信息的民事法律保护:以霍菲尔德权利理论为起点[J].江苏社会科学,2022(2):91-104.
- [32]许可.数据权利:范式统合与规范分殊[J].政法论坛,2021(4):86-96.
- [33]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):54-63.
- [34]张凌寒.算法治理应首先立足保障算法权利[N].中国社会科学报,2023-07-21(4).
- [35]张恩典.大数据时代的算法解释权:背景、逻辑与构造[J].法学论坛,2019(4):152-160.
- [36]万方.算法告知义务在知情权体系中的适用[J].政法论坛,2021(6):84-95.
- [37]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(6):20-39.
- [38]林涸民.《个人信息保护法》中的算法解释权:兼顾公私场景的区分规范策略[J].法治研究,2022(5):48-58.
- [39]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.
- [40]张凌寒.商业自动化决策算法解释权的功能定位与实现路径[J].苏州大学学报,2020(2):51-60.
- [41]张恩典.超越算法知情权:算法解释权理论模式的反思与建构[J].东南法学,2022(1):1-17.
- [42]孙建丽.算法自动化决策风险的法律规制研究[J].法治研究,2019(4):108-117.
- [43]王苑.完全自动化决策拒绝权之正当性及其实现路径:以《个人信息保护法》第24条第3款为中心[J].法学家,2022(5):72-86.
- [44]张欣.免受自动化决策约束权的制度逻辑与本土构建[J].华东政法大学学报,2021(5):27-40.
- [45]程啸.个人信息保护法理解与适用[M].北京:中国法制出版社,2021:233.

Private Rights Construction and Implementation of Business Algorithmic Automated Decision-making

Sun Yueyan Xu Jianfeng

Abstract: The construction of private rights in automated decision-making of business algorithms is a crucial part of algorithmic private governance, and the core of individual empowerment lies in protecting the subject's position of people in technological threats. Algorithm rights can protect the legitimate interests of decision-makers, defend their independence and autonomy, and satisfy conceptual standards. Specific algorithmic rights may be explicitly stipulated in legal rules, or can be derived from basic rights through the method of right inference, and have the possibility of realization, which is the satisfaction of the empirical standards. Using the "bundle of rights" as a pragmatic approach to analyzing rights, algorithmic rights can be seen as a series of rights enjoyed by decision-makers aimed at enhancing individuals' ability to resist algorithmic risks, with open characteristics and normative flexibility. However, if the theory of rights is to be transformed into practice, it is necessary to rely on the institutionalized implementation of various rights in the algorithmic "bundle of rights", starting from the algorithmic right to know, the algorithmic interpretation right, and the algorithmic decision rejection right, to construct a path for the realization of rights.

Key words: algorithmic automated decision-making; right proofing; algorithmic rights bundle; private rights construction

责任编辑:一鸣