

算法歧视的民事责任形态

潘芳芳^{*}

目 次

- 一、无缔约情形下的算法侵权责任
- 二、磋商中的缔约过失责任
- 三、合同履行中的算法违约责任
- 四、结语

摘要 算法决策在社会交往中的应用日趋普遍化,产生了算法歧视这一一般性问题。根据算法歧视发生的不同场域,可以将算法歧视区分为无缔约情形下的算法歧视和缔约过程中及合同履行过程中的算法歧视。在无缔约情形下,算法歧视本质上是对民事主体信用权、平等就业权等人格权的侵害。利用算法决策的行為人如使用敏感数据,则应当推定其存在侵权故意。在其他情形下,由于行為人负有对算法程序的影响评估和合规审计义务以及一定的算法解释和说明义务,违反义务本身即可认定其有过错。在缔约磋商和合同履行过程中,上述保护义务可归入前合同义务及附随义务的范畴,且行為人还负有信息安全管理保护和正当程序保障义务。上述责任形态发生竞合时,应赋予受算法决策影响的一方选择权。

关键词 算法歧视 违约责任 侵权责任 保护义务 通知义务

算法决策技术在人们日常生活中的广泛使用,引发人们对算法歧视危害后果的普遍关注。不同领域的专家和学者从各自不同的视角出发,提出了多元的解决方案。时间维度的规制方案划分中,事前的预防和制止模式,既包括以算法决策根源——数据的合法、合规收集和使用为切入点,来减少和避免算法歧视发生的风险;^[1]也包含设置专门机构、组织和人员,对算法程序进行技术性的审计、评估和测试,通过行政性质的许可和准入等限制措施和手段进行监督和管理,实现算法歧视的风险预防和

* 潘芳芳,中国政法大学人权研究院博士研究生,丹麦哥本哈根大学联合培养博士研究生。本文系中国政法大学博士研究生创新实践项目“智能决策的透明度与人权保护”(项目号2018BSCX53)的阶段性研究成果。

[1] See Michele Wilson, “Algorithms (and the) Everyday”, 20 *Information, Communication & Society* 137-150 (2017).

避免;〔2〕亦有学者提议从技术和法律规制措施融合的角度,要求程序设计者将技术伦理和规范编入算法决策程序。〔3〕与事前的风险预防和规避措施相对应的,是法规范领域传统的事后追责与救济模式,即在算法歧视风险无法避免和消除时,通过法规范体系内相关理念、制度和规则的解释和适用,明确损害后果承担的责任主体和依据,为受害人的损失提供救济的“矫正正义模式”。〔4〕尽管“事前的预防比事后的救济更重要”〔5〕,但在欠缺统一的标准、规则和专业的人员与组织配备的情况下,事前预防模式下各种措施的功能发挥受到诸多限制,不具备基本的现实性和可操作基础。所以,在传统的权益保护制度和框架内,通过被动性作用和功能发挥的事后追责与救济路径的建构,为受害人提供基本的损害赔偿和救济,是目前最具可行性和急迫性的路径,这也是本文意欲探讨的问题。本文的研究目标在于,以《个人信息保护法》为依据,在定型化的民事权益保护与救济框架内,探讨不同场景下算法歧视的责任形态,为受害人寻求司法救济提供理论基础和规范支撑。

一、无缔约情形下的算法侵权责任

明确行为人与受害人之间的法律关系,是确定彼此间权利义务关系的前提。民事规范体系中,存在合同责任和侵权责任两种基本的责任形态,而合同责任则根据责任形成阶段的不同,又细分为发生于接触和磋商阶段的缔约过失责任和发生于合同履行过程中的违约责任。在算法决策程序的行为主体(后文统称“行为人”)和被决策者之间不具有特殊的缔约或合同关系,但被决策者却因为行为人所实施的算法歧视行为而遭受损害时,可通过侵权责任机制作用的发挥,让行为人承担侵权责任,赔偿受害人因遭受不合理的歧视性对待而受到的损失。

(一)算法歧视的侵权行为样态

算法,也可以称之为自动化决策,是指利用个人信息对个人的行为习惯、兴趣爱好或者经济、健康、信用状况等,通过计算机程序自动分析、评估并进行决策的活动。以大数据和计算机程序设计为依托的算法决策程序的运用,无论是数据的收集和挖掘,还是算法程序的设计和使用阶段,都对应产生不同类型的算法歧视:既包括因直接的敏感数据使用和数据挖掘中社会固有刻板印象的再现而导致的“特征选择型”算法歧视,也包括对敏感数据相关联的数据的联合使用而导致的“偏见代理型”算法歧视,以及主要用于商业领域、以类型化和个性化推送为目标的“大数据杀熟型”算法歧视等。〔6〕数据、算法程序和最终的算法决策结果之间,好比食材、食谱和菜肴之间的关系,菜肴是在按照食谱对食材类别、分量和添加程序等指导下做出来的。〔7〕所以,当算法决策结果存在歧视现象时,既有可能是作为食材的变量因素——所使用的“数据”有问题,也可能是算法程序的设计不合适或存在纰漏。

适用广泛的同时伴有极高的风险系数,算法决策技术可能引致的算法歧视危机在世界范围内引

〔2〕 See K. A. Taipale, “Data Mining and Domestic Security: Connecting the Dots to Make Sense of Data”, 5 *Columbia Science and Technology Law Review* 3 (2003).

〔3〕 See B. Custers, “Data Dilemmas in The Information Society: Introduction and Overview”, in B. Custers et al. eds., *Discrimination and Privacy in the Information Society*, Springer, 2013, pp.3-26.

〔4〕 参见郑戈:《算法的法律与法律的算法》,载《中国法律评论》2018年第2期,第66-85页。

〔5〕 参见卢义杰、来扬:《事前预防比事后救济更重要》,载《中国青年报》2012年11月2日,第7版。

〔6〕 参见郑智航、徐昭曦:《大数据时代算法歧视的法律规制与司法审查——以美国法律实践为例》,载《比较法研究》2019年第4期,第111-122页。

〔7〕 See Michele Wilson, “Algorithms (and the) Everyday”, 20 *Information, Communication & Society* 137-150 (2017).

发了广泛的关注和警惕。不同的国家和组织在技术与法律融合的基础上,都专门针对算法歧视风险的避免进行了设定,^[8]内容既涵盖抽象的原则、伦理要求,也有具体的制度和规则建议;既包括事前的预防和制止措施,也有事后的追责与救济机制;既强调内在自律,也重视外在他律。^[9]技术(风险)不分国界。当前,算法安全与算法歧视规制已经成为世界性的难题与挑战。上述不同国家、区域和组织所提出的各种安全标准和行为准则等,已得到我国社会公众、技术专家和规范制定者的广泛吸收和认可,无论是在交易习惯还是新出台的《个人信息保护法》中都有所体现。算法歧视的侵权责任建构,即以上述规范和文件中所创设的交易习惯和法律理念、原则、规则为基础和依据。具体而言,依据主体主观状态和行为样态的不同,可将算法歧视行为划分为通过积极的作为所实施的故意侵权行为和因过失违反作为义务而承担的过失侵权责任。

1. 故意实施侵权行为:敏感数据使用

计算机程序的运行,遵循“GIGO”(Garbage In, Garbage Out)定律,转接到算法歧视场景下则变为“BIBO”(Bias In, Bias Out),即偏见进,偏见出——如果输入的数据是具有歧视性质的,那么输出的算法决策结果就必然带有歧视色彩。^[10]将部分可能造成歧视后果的“敏感数据”排除算法决策体系的自变量范畴,即“技术性脱敏”,已经成为规范领域反算法歧视的基本要求,是最容易操作和实现、操作过程可视化程度最高的一种途径和手段。

关于“敏感数据”的范围界定和类型划分,受不同国家、区域社会特定规范性文化的影响,在界定过程中存在很多差异。具体到算法歧视规制中的“敏感数据”的认定,除了反歧视法规范中所列示的性别、年龄、种族、国籍、残疾、心理的群体保护疾病或心灵能力、性取向、性别、性别认同、性别特征、宗教信仰、信条或个人政治意见等类型之外,针对具体化的算法决策适用场景可能会有所延伸和扩张。典型案例如“大数据杀熟”现象中,经营者根据顾客的经济能力和消费喜好而进行差异化定价。在该场景下,客观上与商品或者服务的价格设定没有关联性的当事人的经济能力和消费喜好,也可被视为与歧视相关的“敏感数据”。《个人信息保护法》中,将“敏感个人信息”界定为“一旦泄露或者非法使用,容易导致自然人的人格尊严受到侵害或者人身、财产安全受到危害的个人信息”,并且在“人格尊严”与“人身、财产安全”的双重保护目标与功能下,立法者将“生物识别、宗教信仰、特定身份、医疗健康、金融账户、行踪轨迹等”以列举的形式,明确纳入个人敏感信息的范畴。^[11]保护当事人免受歧视性待遇是保护当事人“人格尊严”的重要内容。是故,但凡非交往或者交易中所必需,但一旦在决策程序中使用就会招致歧视性待遇和后果的信息,均可经扩张解释,纳入法条中具有兜底性质的“等”字范畴之内。

由此,行为人如若在算法程序运行中使用了歧视性“敏感数据”或可能造成歧视性后果的数据,

[8] 如2015年美国白宫发布的《大数据:把握机遇,维护价值》(Big Data: Seizing Opportunities, Preserving Values)报告;2016年英国下议院科学和技术委员会发布的《机器人与人工智能》;2018年的《多伦多宣言》(The Toronto Declaration: Protecting the Rights to Equality and Non-discrimination in Machine Learning Systems),该宣言明确提出了“将偏见和歧视排除在机器学习算法之外”的指导方针;2018年《欧盟通用数据保护条例》(GDPR: General Data Protection Regulation)和《“可信赖人工智能”伦理指南》(Ethics Guidelines for Trustworthy AI)。

[9] 参见郑智航、徐昭曦:《大数据时代算法歧视的法律规制与司法审查——以美国法律实践为例》,载《比较法研究》2019年第4期,第111-122页。

[10] See A. Phalgune et al., “Garbage in, Garbage out? An Empirical Look at Oracle Mistakes by End-User Programmers”, *IEEE Symposium on Visual Languages and Human-Centric Computing* 40 (2005).

[11] 参见《个人信息保护法》第28条。

便可推定其主观上故意实施了算法歧视的侵权行为。除非行为人能够对所使用“敏感数据”与最终决策之间合理、正当的关联性和数据使用的必然性进行解释和说明，并能够证明自身已经采取适当的措施，避免因敏感数据使用而带来的歧视性后果。否则，在被决策者权益因算法歧视行为而遭受侵害时，行为人要为之承担相应的侵权责任。

2. 过失违反作为义务

侵权法最能够反映特定社会和文化背景下社会关系、经济关系和伦理道德观念。侵权责任机制中，司法者可通过行为人所负担的“作为义务”内容的适度扩张和解释适用，将特定社会背景下人们的现实需求、特定的道德伦理观念、社会交往方式和交易习惯融入侵权责任规制机制中，弥补“规范滞后”的不足。^{〔12〕}按照德国判例与学说中的观点，“作为义务”的产生，源于法律规定、合同约定和先前危险行为；^{〔13〕}“作为义务”内容的扩张理由，则包括“风险社会”下危险控制的需要^{〔14〕}、法律理念由个人本位转向社会本位的必然要求^{〔15〕}和“道德法律化”理论的要求三种^{〔16〕}。其中，作为“作为义务”产生依据之一的“法律规定”，不仅指代民事基本法，还包括国家因保护特殊主体或者权益而制定和实施的、带有公法性质的特别法规范。在算法歧视的私法救济和民事责任规制路径建构中，可由主体的“作为义务”（内容的扩张）发挥“导管”作用，将特别法规范中倡导和设定的价值理念、原则、规则和制度等引入私法救济体系中，通过私主体间权利义务关系的构建，来实现国家干预和调控的目的，维护和保障社会利益和公共秩序。^{〔17〕}《个人信息保护法》便是在数字化时代背景下，国家为了保护处于弱势地位的信息主体的相关权益而制定和实施的，带有公法性质的特别法规范。

数字化时代背景下，无论是数据和信息瑕疵，程序设计的结果导向，还是算法程序“黑箱”属性下的透明度缺失，都容易引发算法歧视危机，被决策者遭受歧视性待遇的风险系数颇高。行为人引入算法决策系统辅助、部分或者全部地代替人类决策，开启了被决策者可能遭受歧视性待遇的危险。“权责一致”的逻辑和理念下，即便没有故意实施算法歧视的主观意图，作为招致风险的源头，行为人也理应为自己这种“先前危险行为”负担一定的作为义务，采取合理必要的、具备期待可能性的保护和防范措施，预防和避免算法歧视现象的发生，并在风险预防和制止义务履行存在瑕疵、被决策者遭受了歧视性待遇时，为被决策者的损失承担侵权赔偿责任。^{〔18〕}

（1）算法程序运行的影响评估与合规审计义务

《个人信息保护法》明确要求，个人信息处理者应当定期对个人信息处理活动是否遵守法律、行政法规的情况进行合规审计。并且，在利用个人信息进行自动化决策过程中，要对个人信息处理活动在

〔12〕 周友军：《我国侵权法上作为义务的扩张》，载《法学》2008年第2期，第92-101页。

〔13〕 周友军：《我国侵权法上作为义务的扩张》，载《法学》2008年第2期，第92-101页。

〔14〕 法律是社会的产物，社会生活的事实在需要是导致法律制度和规则发生变迁的动力和原因。在技术发展与变革的背景下，社会风险和危险也不断增加，“风险社会”中通过对人们作为义务的增加以要求人们“照顾好邻居”，并通过不作为责任的承担将作为义务规范化和法定化。参见邱聪智：《从侵权行为归责原理之变动论危险责任之构成》，中国人民大学出版社2005年版，第405页。

〔15〕 自由主义思想观念下，个人自由越多越好，承担的义务越少越好，但在社会本位的法律思想下，为了社会的公共利益和安全，个人应当承担的义务也有所增加。参见周友军：《我国侵权法上作为义务的扩张》，载《法学》2008年第2期，第92-101页。

〔16〕 社会的现实需求导致新的道德观念和要求的产生，而法律要与社会的变迁和发展相适应，就得允许新的道德内容的“入侵”，通过“道德内容法律化”的路径将社会生活现实的需求在法律体系中予以体现。参见周友军：《我国侵权法上作为义务的扩张》，载《法学》2008年第2期，第92-101页。

〔17〕 参见苏永钦：《私法自治中的国家强制》，载《中外法学》2001年第1期，第92-109页。

〔18〕 林美惠：《侵权行为法上交易安全义务之研究》，台湾大学2000年法律学研究所博士论文，2000年，第37页；转引自李昊：《德国新债法中附随义务的构造》，载《环球法律评论》2009年第5期，第113-123页。

事前进行影响评估。^[19]算法以数据和决策程序为依托而运行,所以,对算法程序运行的影响评估与合规审计,意味着行为人在信息安全与程序公平两方面展开核查。

自动化算法决策过程中,对信息的审查与评估,一方面,是为了保障个人信息的安全;另一方面,也是为了避免关联数据联合使用下“偏见代理型”算法歧视的发生。对于个人信息安全保护的影响评估,内容主要包括信息的处理目的、处理方式等是否合法、正当、必要,对个人的影响及风险程度,以及所采取的安全保护措施是否合法、有效与风险程度相适应三个方面。此外,在有些算法决策过程中,行为人输入和使用的虽然是经过脱敏处理的客观、中立数据,但这些数据可能会与敏感数据存在一定的关联。对这部分较为隐秘的敏感性“关联数据”的采用,最终也可能会产生歧视性的决策结果,这种情况下所发生的算法歧视,被学者称之为“关联歧视”或“偏见代理型”算法歧视。^[20]典型案例,如看似中立、客观的邮政编码信息,却可能会与某特殊种族人们的集中居住地之间存在一定的关联,依此类邮政编码信息所进行的算法决策,结果很可能就会带有种族歧视色彩。^[21]因此,在数据方面进行评估与审计的另外一个重要内容,即为敏感性“关联数据”的检测与排查。即除了明显、可视的典型歧视类敏感数据外,还应结合算法程序的原理和试运行后果,发现并排除这类可能与歧视性敏感数据相关的“关联数据”,确保输入和使用的数据信息与决策结果之间具备合理关联性,且不会对某一特定群体产生不合理的差别对待。此外,若决策程序所使用的数据信息存在不全、错误或者更新不及时等问题时,也会使被决策对象遭受不公平的歧视性待遇。为避免此类歧视发生的可能性,除了类别和范围限制外,行为人还应当保证所使用数据和信息的质量和准确度,确保所使用的数据不存在错误、遗漏或更新不及时等问题。

除数据和信息相关的义务外,算法决策程序自身也应当接受专业的评估、审查与检验,以确保程序的实施和运行不会对某一特定群体产生不合理的差别待遇。算法的广泛开发与利用与社会公共利益密切相关,不合理、不公平的算法程序运行可能会引发大范围的社会风险。所以,针对不同场景下的算法决策技术的使用,“审慎适用”和精细化的“分类监管”已经成为我国算法行业的基本模式和思路。2018年,在中国科学院软件研究所的牵头之下,电子技术标准化研究院发布了《人工智能深度学习算法评估规范》。作为国内发布的首个算法评估标准,该规范以算法技术使用的具体场景为划分依据,确定了不同层级、不同数目的指标,搭建了初步的算法评估指标体系,成功地迈出了尝试和探索场景化、精细化的算法治理的第一步。^[22]此外,其他新近出台的一系列法律、规章、条例中,也都有专门的条款对不同场景中的算法应用进行治理和规制,如《数据安全管理办办法(征求意见稿)》中针对个性化的新闻和广告推送、差异化的商品与服务定价的治理;《电子商务法》针对搜索类算法和商品、服务的竞价排名的规范;《信息安全技术、个人信息安全规范(草案)》中针对个性化展示及退出相关的规制等。行为人在引入算法决策程序辅助或者代替人类进行决策的过程中,应当以应用场景为基础,对照不同的规范性文件的要求进行风险的评估与审计,并根据评估和审计的结果确定下一步的行为。

以现行规范性文件为基础和依据展开算法歧视影响的评估与审计,并非履行作为义务的唯一路

[19] 参见《个人信息保护法》第54、55条。

[20] 郑智航、徐昭曦:《大数据时代算法歧视的法律规制与司法审查——以美国法律实践为例》,载《比较法研究》2019年第4期,第111-122页。

[21] See Sandra G. Mayson, “Bias In, Bias Out”, 128 *Yale Law Journal* 2218 (2019).

[22] 参见张欣:《从算法危机到算法信任:算法治理的多元方案和本土化路径》,载《华东政法大学学报》2019年第6期,第17-30页。

径,行为人还有其他多元的途径和方式可供选择。行为人既可以邀请领域内的专家对程序的设计逻辑和原理进行评估,也可以委托有资格和能力的组织和机构进行检验和认证;在具备现实基础的情况下,为确保评估与审查的效果,行为人还可以采取试运行的方式,接受非领域专家和相关利益群体的监督、检查与批评。但无论采取何种途径和方式,行为人都需要谨慎进行,只有在确保算法程序运行的各个环节都不会产生歧视性后果的基础上,才可正式引入算法决策程序,投入运行和使用。

总的而言,行为人在算法决策系统投入前,对数据和程序所展开的审查与评估,既可以有效避免歧视的发生,又能够满足算法决策审查与监督机制的要求,为后期可能发生的纠纷和争议的处理提供证明和依据,^[23]是行为人理应承担的一项非常重要的义务。

(2) 算法程序的有限透明和可解释性义务

保证算法决策的透明度以及处理结果的公平、合理,是对行为人最基本的要求和其理应遵守的基本原则。为了保障被决策者的权益不受算法决策的侵害,以欧盟为代表的相当部分的国家和区域在相关规制中,都通过赋予被决策者新型权利的模式来化解双方在信息或者权力上的不对称局面。我国针对个人信息保护的规范制定中,立法者针对自动化算法决策技术的使用也明确规定,被决策者认为算法技术的使用会对其权益造成重大影响的,有权拒绝行为人通过算法技术作出决定,或者要求行为人对其个人信息处理规则进行解释和说明。

《个人信息保护法》中,立法者欲通过行为人解释和说明义务的履行,保障算法决策实施的公正和透明,避免被决策者的权益遭受侵害,这与欧盟《通用数据保护条例》中赋予被决策者享有算法解释权的路径在本质上的功能与意义相同,学界将其称之为被决策者所享有的“算法解释权”,或者行为人负担的算法解释义务。算法解释权(义务)的提出,在本质上是为了抑制“算法权力”,矫正权利失衡的格局,保障在实质上处于弱势地位的被决策者合法权益不受侵害。但依据我国民事规范领域的传统,“权利类型”法定,新型权利的设定权只能归立法者所有。所以,在当前《民法典》和《个人信息保护法》中均没有明确赋予信息主体享有“解释权”的情况下,可通过行为人“可解释性义务”负担的路径,使司法者在不突破权利体系与规范传统的前提下,通过灵活、包容、开放的“作为义务”内容的适度扩张,实现或保障算法决策的公开与透明。

如前文所述,行为人算法可解释义务的负担,源于其对算法歧视危险的开启和维持,即损害结果可能性的产生;目的在于打开算法的“黑箱”,帮助人们了解算法决策程序的运行逻辑与机理,实现算法规制中最基本和朴素的追求:算法有限度的透明性和算法可追责性。^[24]解释义务的履行包括算法解释和算法信息披露两个方面;^[25]履行的方式和程度要求,则需要根据具体的使用场景予以确定,考虑因素包括但不限于歧视风险的产生原因和大小、解释对象的专业能力与理解水平、对商业秘密或知识产权的可能影响和“对歧视风险控制的作用大小”等。^[26]此外,不同的场合和适用场景下,解释的内容、侧重点也会有所区别。如经营者对消费者个性化的广告和服务推送算法的解释集中于信息获取来源和范围,而被决策者对信用评价中的算法解释则更看重数据的具体类型、特定类型数据所占的权重,以及算法程序运行的规则和原理。为避免与商业秘密、知识产权等权益的冲突和矛盾,义务主

[23] See Sonia K. Katyal, "Private Accountability in the Age of Artificial Intelligence", 66 U.C.L.A. Law Review 54 (2019).

[24] 参见李婕:《人工智能中的算法与法治公正》,载《人民法院报》2018年5月23日,第2版。

[25] 参见郑智航、徐昭曦:《大数据时代算法歧视的法律规制与司法审查——以美国法律实践为例》,载《比较法研究》2019年第4期,第111-122页。

[26] 李昊:《交易安全义务论:德国侵权行为法结构变迁的一种解读》,北京大学出版社2008年版,第115-118页。

体不需要太过细致地阐明程序的设计细节和运行机理,但对于数据类型、使用数据与决策结果的关联性和基本逻辑等问题需要作出解释,并以被决策者可理解的语言和方式进行解释。^[27]

法官对行为人所负担的作为义务的内容设定,具备灵活性、延展性和流动性特征,能够随着社会的变革和发展而变化——“通过作为义务的扩张,法律秩序可以借助损害赔偿制裁有义务控制危险的人,来防止陷入物或者人的危险范围的公众的期望落空。”^[28]大数据、信息化社会背景下,行为人为了预防和避免算法歧视现象的发生而承担的数据脱敏、算法程序的影响评估与合规审查、算法程序的解释等义务,不仅能够有效避免和减少算法歧视现象的发生,也为算法歧视侵权规制机制的建构提供了实现路径,成为行为人侵权责任承担和被决策者获得私法救济的规范基础和依据。

(二)算法歧视的行为认定与损害后果

“算法歧视”是对因算法决策技术的使用而遭受的不合理的差别待遇的统称,具体到侵权救济机制中遭受侵害的具体权益类型的确立,则需要根据算法决策技术使用的具体场景做出判断。算法存在不同的功能和类型,有用于个人和群体特征分析和类型化的数字画像(profiling)和分类(classification)算法;有针对个性化推送和服务的过滤(filter)和推荐(recommendation)算法;有对未来可能性进行预测的数据挖掘(data mining)和机器学习(machine learning)算法。^[29]功能和类型多元的算法被广泛用于信用评级或评分、职场能力评估、个人喜好预测、广告推送和个性化定价等多场景化的生活中。^[30]算法程序获得的歧视性决策结果,尤其是在结果对外公开和使用的情况下,会对个人的信誉度、交易能力和水平、就业和贷款资格、违法或者违约可能性以及商品和服务购买能力等产生负面影响,使主体的信用权益、劳动者平等(就业)权以及消费者的合法权益等产生侵害。

在算法歧视案件的审理过程中,对于不公平的歧视待遇的认定与审查,主要有区别待遇(disparate treatment)和差异影响(disparate impact)两种模式。^[31]以“区别待遇”模式为基础的歧视认定规则,也被称之为“主观故意歧视”,主要包括两个要素:同样的情况遭受不同的对待;有主观上的歧视意图。而“差异影响”也可称之为“无过错的算法歧视审查模式”,是指所实施的行为或规则表面上看似中立,但对所保护的对象产生了不合比例的不利影响,这种情况下,歧视的认定,无须考量主体实施行为或者规则的时候是否有歧视的故意,只需要审查行为或者规则实施以后,是否针对对方产生差异性的影响即可。^[32]

在行为人存在不同样态的侵权行为模式下,针对前文所述的“主观故意实施算法歧视”的行为的审查与认定,法官可采“区别对待”模式。行为人在明显可知、可察的情况下将歧视性质的敏感数据纳入决策的自变量范畴,导致被决策者因不公平的歧视性决策结果而遭受损失时,为之承担侵权损害赔偿责任。而对于大多数情况下,行为人是以消极性、带有隐秘特性的不作为,违反应当履行的歧视风险预防和制止义务时,法官则可采用“差异影响”的审查模式。在被决策者能够证明算法决策给自

[27] See Andrew. Selbst & Julia Powles, “Meaningful Information and The Right to Explanation”, 7 *International Data Privacy Law* 233–242 (2017).

[28] 李昊:《交易安全义务论:德国侵权行为法结构变迁的一种解读》,北京大学出版社2008年版,第114页。

[29] See Mittelstadt, B. D., et al., “The Ethics of Algorithms: Mapping the Debate”, 2 *Big Data & Society* In press 1-21(2016).

[30] 公共领域如司法系统中法官根据算法对再犯罪概率的预测进行量刑、行政领域依据算法对人力和物力资源进行配置,或依据算法对公共资源和服务进行分配等也存在歧视风险,但属于公共主体的行为规制问题,因不在本文研究范围内,所以不予涉及。

[31] See Solon Barocas & Andrew Selbst, “Big Data’s Disparate Impact”, 104 *California Law Review* 671 (2016).

[32] See Richard Primus, “The Future of Disparate Impact”, 108 *Michigan Law Review* 1341-1387 (2010).

已带来了不利的差异性后果,且行为人不能够证明自身没有过错时,认定行为人违反了歧视风险的预防和制止义务,承担损害赔偿的侵权责任。“差异影响”的算法歧视审查模式,能够发现隐藏的歧视意图,有效规制在形式上非歧视但会造成歧视性后果的算法决策行为,^[33]有效平衡与弥合双方因专业能力、不对称的信息和失衡的风险预防和制止能力等导致的不平等的法律地位,适应算法歧视案件的特殊性,^[34]能够更好地维护和保障受害者基本权益。

二、磋商中的缔约过失责任

缔约过失作为一种学说继受的成果,已经得到我国立法者的认可,并在《民法典》予以全面采纳。“抑制违反诚实信用原则的过失行为”和“保护合理信赖”,是缔约过失责任制度的思想根据,也是民法典设立该责任制度的基本功能和目的。^[35]以诚实信用原则为价值基础,在算法决策的双方当事人为了合同的缔结而进行接触和磋商,而对对方产生了合理信赖之时,彼此间就已互负以协作、照顾和保护等为内容的先合同义务。并且,这种在“具有特别结合关系的当事人之间”基于彼此的信赖关系而产生的注意义务,通常要比侵权法中没有特定关系的主体间的注意义务要求程度更高、更重,与侵权行为法中一般性的注意义务在本质上具有一定的差别。^[36]如果一方当事人在接触和磋商过程中引入了算法决策程序,但却因没有尽到必要的注意义务而使对方遭受歧视性待遇,并导致合同不成立、解除或被撤销时,行为人的行为构成先合同义务的违反,即“缔约上过失”,应为被决策者的信赖利益损失承担“缔约过失责任”。

(一) 先合同义务的内容

行为人先合同义务的负担是其承担缔约过失责任的前提。先合同义务的具体内容,则需要根据具体的案件事实,如双方地位的失衡程度、信息掌握量的差异以及彼此间信赖程度等,结合法政策的价值选择和衡量、社会交易习惯和行业标准及规范等内容进行综合确定。在合同法的相关规范框架内,结合不同专业与领域的技术规范和标准以及域内外的相关规范、政策和文件可知,双方当事人磋商和接触过程中,行为人因引入算法决策程序而承担的先合同义务的具体内容包含以下两个方面。

1. 信息义务——算法程序使用的告知和披露义务

信息,是人们意思形成和作出决策的前提和基础,保障私主体决策的作出要建立在充分了解和掌握信息基础之上才能真正地维护当事人意识自由、实现私法自治。^[37]大数据时代背景下,合同缔结的算法引入者与被决策者之间,无论在专业知识还是信息获取能力和数量上都明显失衡。严重的“信息不对称”和“地位失衡”现象,突破了传统私法行为理念中以“抽象理性人”为基础所预设的民事主体之间法律地位平等的局面。^[38]因而,继续将收集、分析和掌握与交易内容相关的信息视为交易主体自身的“不真正义务”,并要求当事人“买者自负”,自行承担“怠于履行”下信息不全而导致的不利

[33] 崔靖梓:《算法歧视挑战下平等权保护的危机与应对》,载《法律科学(西北政法大学学报)》2019年第3期,第29-42页。

[34] See Gorge Rutherglen, “Disparate Impact Under Title VII: An Objective Theory of Discrimination”, 73 *Virginia Law Review* 1297-1345 (1987).

[35] 孙维飞:《〈合同法〉第42条(缔约过失责任)评注》,载《法学家》2018年第1期,第179-191,196页。

[36] 参见崔建远主编:《合同法》(第五版),法律出版社2010年版,第122页。

[37] 参见彭正穗:《民事法对市场信息不对称的防范》,载《中南财经政法大学学报》2000年第4期,第103-107,128页。

[38] 参见徐明月、林全玲:《信息不对称、附随义务与缔约过失责任》,载《河北法学》2005年第10期,第36-38页。

后果,必然无法实现真正的意思自治,也有违法律最基本的公平正义追求。鉴于协商和洽谈中被决策者处于“结构性劣势”^[39]地位,不利于交易公平和意思自治,有必要给予其特殊的、倾斜性的法律保护,对行为人课以相应的信息披露和告知义务,矫正双方失衡的地位和能力,实现实质平等与正义。

“理解算法运作的机理是监督并抵消处于优势地位的为人‘算法权力’的重要前提,既能够避免行为人的恣意和武断,又可作为保证决策的可信、正当和理性的重要约束机制。”^[40]算法程序运作机理的解释、说明和披露的内容包括数据和程序两个方面。数据信息层面,除了数据收集过程中所要求的“明示收集、使用信息的目的、方式和范围,并经被收集者同意”^[41]等出于个人信息保护目的所要求的内容外,行为人应主动向被决策者告知和说明算法决策结果作出所依据的数据类别,具体来源、不同数据的权重以及所输入数据与最终决策结果之间的关联性等与算法程序运行相关的数据信息。^[42]此外,告知与披露过程中,针对被决策者对数据的相关信息所产生质疑和反馈,行为人也有义务进行进一步的解释和证明,并在发现问题与瑕疵时,作出修正和完善。与算法程序运行逻辑相关的信息包括引入算法程序的功能和目的、算法决策结果在最终决策中的重要性(即是否会对合同最终缔结与否产生实质性影响),决策结果可能存在的风险、会对被决策者产生的影响和后果等。^[43]同数据信息披露一样,在被决策者针对程序运行机理问题尚有疑问的情况下,行为人也需要对相关内容进一步地解释和说明,在有必要的情况下修正和完善。被决策者在了解具体信息后,拒绝使用算法决策结果的,行为人理应支持和尊重。

鉴于专业知识的匮乏和“技术鸿沟”所导致的被决策者理解障碍,行为人信息告知和披露义务的履行方式也需要满足一定的要求:告知需要通过“简单透明、明晰且容易获取的方式,清楚明确的语言,合适的方式”进行;信息的提供需满足“简洁明了、清晰可见、晓畅易读”的要求,^[44]以实现“相对人对告知的内容有明确而清晰的认知和了解”的目的。^[45]披露与告知的最终目的,在于被决策者的“知情且同意”。即,只有被决策者对合同缔结中算法程序的功能、运行机理和参与度等内容均充分知情、理解并予以同意的情况下,行为人才能够在磋商过程中引入算法决策系统,并根据算法决策的结果决定合同缔约与否。

2. 保护义务——歧视风险的评估、审计与预防

从算法程序运行的自身特性出发,无论是数据层面上,如使用了有偏见的、不具有代表性的训练数据,或者数据的体量、精细度不够,还是程序自身(融入了程序员的认知偏见或程序设计目的中的偏见)存在偏见等,都可能导致歧视性后果的产生。所以行为人的评估与审计对象也应当既包括基础数

[39] 张艳:《前契约信息义务的正当性基础及法律适用研究》,载《法学》2019年第7期,第67-76,84页。“结构性劣势”是指,算法决策的使用者是以盈利为目的,而被决策者旨在满足个人最基本的私生活的需求。

[40] 张欣:《算法解释权与算法治理路径研究》,载《中外法学》2019年第6期,第1425-1445页。

[41] 该部分相关信息的告知和披露内容,属于数据收集和使用过程中“知情同意”原则的具体化,与个人信息安全保护相关的规范中都有所规定和体现,典型示例见《欧盟通用数据保护条例》(GDPR)第4条第11款:“数据主体的同意是指,数据主体依照其意愿自由作出的、特定的、知情的、明确的指示。通过以声明或清晰肯定的行为作出的该等指示,数据主体表明其同意处理与其相关的个人数据。”

[42] See Margot E. Kaminski, “The Right to Explanation, Explained”, 34(1) *Berkeley Technology Law Journal* 189-218 (2019).

[43] Sandra Wachter, Brent Mittelstadt & Luciano Floridi, “Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation”, 7 *International Data Privacy Law* 78-79 (2017).

[44] 《欧盟通用数据保护条例》第12条第1款。

[45] 《欧盟通用数据保护条例》第12条第6款。

据和信息,又包括算法程序自身。为避免作为自变量的数据的种类、范围和质量的负面影响,行为人在运用算法程序对缔约对象进行分类、评估或者数字画像时,有必要进行“脱敏”处理,避免带有歧视色彩的数据和歧视数据关联数据的使用,并保证输入的数据和信息来源合法、内容完整、准确。

此外,行为人有义务对所使用的算法程序的歧视风险进行影响评估和审计,避免被决策者因算法程序的使用而遭受歧视性评价和待遇,最终影响缔约的结果。^[46]并非所有的算法程序使用者都具备相关的专业知识,所以程序的歧视风险评估和审计包括“内部分析”与引入“外部力量”两种方式进行。“内部分析”是指,行为人在运用算法程序对缔约对象的履约能力、业务水平和业务能力等进行评估和预测之前,通过程序核检和试运行等方式,在其内部对程序的歧视风险进行评估和审计。使用“外部力量”则主要是因为受专业能力的限制,并非所有行为人都有能力配备专业的人员和设施进行风险预估和审查。此时,行为人可引入或委托具备相应资格的外部力量进行歧视风险分析,确保决策对象受到公平公正的评估和对待。^[47]除了通过专业的人员和机构进行歧视风险的评估与审计外,以“歧视感知数据挖掘技术”为代表的、通过“算法”识别算法程序的运行对公平或者伦理所带来的潜在威胁,也是行为人对算法决策程序进行审计、评估与监督,实现算法公平、透明与可问责时可供参考的方式和路径之一。^[48]

保证算法程序的被决策对象不受歧视性待遇,保障被决策者的人身和财产的完整性利益是设定行为人保护义务的最终目的。所以,本文所探讨的缔约之“过失”的行为认定应采客观标准,^[49]在保护义务的最终目的未达成的情况下,无论行为人主观意愿如何,都需要承担缔约过失责任,避免行为人以缺乏专业性知识和能力为由推卸责任。并且,对行为人课以严格的作为义务,能够提醒其谨慎地引入算法决策程序、谨慎地使用算法决策程序所得出来的结论,减少算法程序歧视现象的发生,降低被决策者遭受损失的可能性。

(二) 缔约过失赔偿责任

当事人因自己的过失致使契约不成立者,对信赖其契约能够有效成立的相对人,应赔偿基于此项信赖而遭受的损失。以耶林该理论为基础,一方当事人在合同的磋商和缔约过程中引入算法程序辅助或替代人类决策,但却未履行或未完整履行上述的信息义务和保护义务,致使被决策者遭受算法的歧视性待遇,影响了合同的缔结,使合同不成立或成立后无效、被撤销时,行为人应承担缔约上过失之责任,赔偿对方因信赖合同成立而遭受的信赖利益损失和人身、财产权益遭受损害导致的完全性利益损失。其中,信赖利益损失,包括所受的损害和所失的利益,如为了合同的签订而合理支出的费用、丧失与第三人订立合同的机会而产生的损失等。^[50]此外,若算法程序对被决策者作出的带有歧视性的评估和预测结果因对外泄露和公布使被决策者人身、财产等完全性利益遭受进一步损害时,行为人还应当对扩大的损害后果承担相应的责任,且这部分损害责任的承担与原本的缔约过失责任承担之间产生责任竞合,受害人既可以选择基于侵权请求权,也可以选择基于缔约过失请求权请求赔偿损失。^[51]

[46] 参见张欣:《从算法危机到算法信任:算法治理的多元方案和本土化路径》,载《华东政法大学学报》2019年第6期,第17-30页。

[47] See Adam Harkens, "Fairness in Algorithmic Decision-making," 1(1) *Amicus Curiae* 84-96(2019).

[48] See Reuben Binns, "Algorithmic Accountability and Public Reason", 31 *Philosophy & Technology* 543-556(2018).

[49] 参见韩世远:《合同法总论》,法律出版社2018年版,第170页。

[50] 参见孙维飞:《〈合同法〉第42条(缔约过失责任)评注》,载《法学家》2018年第1期,第179-191,196页。

[51] 参见韩世远:《合同法总论》,法律出版社2018年版,第187页。

三、合同履行中的算法违约责任

合同当事人之间的权利义务关系由合同内容确定,而合同的内容,除了当事人合意约定之外,法律还会根据交易习惯和特殊原因等进行补充和完善。当事人因法律补充和完善而承担的这部分义务,被称为“附随义务”。作为合同履行的基本原则,“全面履行”的内容既包括当事人约定的义务,也包括因诚实信用、合同性质和目的以及交易习惯所确定的附随义务。任何一种义务未履行或者未完满履行,当事人都需要为之承担违约责任。^[52]

某些算法决策程序使用场景中,行为人与被决策者之间具有合同关系,双方以契约为基础进行商品或服务的交易行为,这种情况下的算法歧视责任承担的规范基础、依据和具体规则也具有自身的特殊性。在一方当事人引入算法决策程序辅助或者代替其自身进行决策时,对方因为信息不对称和技术认知能力不足等原因,在实质上处于弱势地位,需要法律在有限的范围内进行“倾斜性”保护。^[53]对行为人课以与算法决策程序使用关联的附随义务,目的不仅在于维持交易活动公平与正义、预防和避免风险的发生,还能够于算法歧视发生后,为受害人的违约救济提供请求权基础和依据。^[54]

(一) 合同履行中的附随义务

算法歧视的后果一般是决策使用者违反附随义务造成的。这里的“附随义务”取狭义,专指合同关系存续期间和履行过程中的通知、协助和保密等义务。^[55]埃默里希教授认为,具体附随义务的内容,需要结合合同的内容和性质,当事人受保护的必要性,风险控制的成本和难易程度,一方对另一方当事人的信任和依赖程度,合作的方式和规模等内容综合予以确定。^[56]一般情况下,合同关系中交易主体对彼此间信赖程度都比较高,一方的知识优势越大、控制风险的成本越低,对对方可能造成损害的风险就越大,行为人应当承担的附随义务的内容和程度要求也就越高。

以义务功能发挥的主体标准不同为依据,可将附随义务的内容类型化为行为人积极主动承担的“自律型”义务和增加透明度以推进“他律”实施的义务。自律义务主要体现在算法歧视风险预防和避免措施的实施,也可称之为积极性的保护义务。如前文所述,歧视风险的排除,既要保证所使用的数据和信息不包含歧视性的敏感信息(技术性“脱敏”),且数据和信息来源合规、内容准确和完整;又要保证算法程序自身无瑕疵,经过专业人员的风险评估、审查与认可,不会对相对人作出歧视性的决策结果。对于部分歧视性敏感数据的关联数据的处理,原则上也应当禁止和排除,只有在特殊情况下^[57]或者行为人对该类数据使用的合理性、必要性、与决策结果之间的关联性保证该类数据不会导致歧视后果进行了担保、必要说明和阐释的情况下,才能够例外使用。^[58]此外,行为人在合同履行的过程中,通过信息的收集、使用和算法程序的运行辅助或者代替人类决策进行交往和交易的,还应当

[52] 参见最高人民法院民法典贯彻实施工作领导小组主编:《中华人民共和国民法典合同编理解与适用(一)》,人民法院出版社2020年版,第340页。

[53] 参见张艳:《现金贷消费者保护的范式转换及制度构建》,载《法学》2019年第6期,第154-167页。

[54] 参见马长山:《智慧社会背景下的“第四代人权”及其保障》,载《中国法学》2019年第5期,第5-24页。

[55] 参见韩世远:《合同法总论》,法律出版社2018年版,第342页。

[56] 参见李昊:《德国新债法中附随义务的构造》,载《环球法律评论》2009年第5期,第113-123页。

[57] 参见《欧盟通用数据保护条例》第9条第1款,关于“特殊种类的个人数据的处理”中,第1款针对个人“种族或者民族出身,政治观点、宗教或哲学信仰以及……数据的处理应当被禁止”。

[58] See Solon Barocas & Andrew Selbst, “Big Data’s Disparate Impact”, 104 *California Law Review* 671-691 (2016).

承担以下附随义务。

1. 实体:信息安全保障义务

算法技术的展开和运行,以收集、使用和处理个人信息为基础。行为人的一系列数据相关行为,必须要符合个人信息保护相关的法律和法规,履行相关的法律义务。首先,个人信息的收集、使用和处理必须要满足最基本的“知情同意”原则。行为人在处理个人信息前,需要以显著的方式、清晰易懂的语言告诉对方信息处理的目的、方式、种类以及保存期限等,保证对方是在充分知情且同意的前提下,自愿、明确地同意行为人的信息行为。另外,行为人在进行信息处理时,应当采取合法、正当的方式,遵循诚信原则的要求;信息的处理有明确、合理的目的,且符合公开、透明的原则。行为人对对方的信息安全负责,并采取必要的措施保障个人信息的安全。以《个人信息保护法》中的规定为原型可知,行为人应当采取的必要措施包括但不限于以下方面:(1)制定内部管理制度和操作规程;(2)对个人信息进行分级分类管理;(3)采取相应的加密、去标志化等安全技术措施;(4)合理确定个人信息处理的操作权限,并定期对从业人员进行安全教育和培训;(5)制定并组织实施个人信息安全事件应急预案;(6)法律、行政法规规定的其他措施。

2. 程序:提供程序性保障义务

美国计算机学会公众政策委员会针对算法的透明与归责问题制定了算法治理的指导细则,提出了知情、质询和申诉、算法责任认定、数据来源披露以及可审计原则等,算法的规制不仅重视实体权利的维护,而且强调当事人在不同阶段应当享有的各种正当程序。^[59]《欧盟通用数据保护条例》(GDPR)虽然以“赋权”为数据保护和算法治理的基本模式和理念,但也对当事人权利行使的程序性保障进行了规定,如知情同意、数据访问、不准确信息的纠正、删除与限制处理权等。我国与之相关的规范内容自然也不例外。比如,《个人信息安全规范》第 7.10 条中,赋予了个人提出申请的程序性权利;《个人信息保护法》也要求信息处理者保障当事人知情、决定、查阅、复制、更正、补充、拒绝和删除等权益,并且建立了个人权利行使的申请受理和处理机制。

具体而言,行为人程序性保障义务的履行主要体现为相关信息的告知和披露。首先,在合同义务的履行确有必要引入算法决策程序辅助或者替代人类决策时,主动引入算法决策程序的一方有义务告知对方将算法程序引入合同履行中的目的和功能,算法程序发挥作用的阶段和范围,可能会对相对人带来的风险和影响等信息,并在对方明确“知情且同意”的前提下,引入算法决策程序,辅助其履行合同义务。其次,在被决策对象需要了解时,行为人有义务对所引入的算法程序的运行机理(使用的数据、信息和算法程序运行逻辑)向被决策者解释和说明,且内容上“简洁明了、清晰可见、晓畅易读”,^[60]以“简单透明、明晰且容易获取的方式,清楚明确的语言,合适的方式”进行;实现被决策者“对告知的内容有明确而清晰的认知和了解”的目的。^[61]再次,行为人应积极、主动地告知和提醒相对人在算法程序运行中所享有的知情、参与、查询、更正(数据和信息)和选择(参与或退出)的权利,^[62]确保相对人能够随时知晓、参与和控制与自己相关的决策过程,随时享有对算法决策系统的使用提出拒

[59] See Council, ACM US Public Policy, “Statement on Algorithmic Transparency and Accountability”, <https://www.acm.org/articles/bulletins/2017/january/usacm-statement-algorithmic-accountability>, accessed October 20, 2019.

[60] 《欧盟通用数据保护条例》第 12 条第 1 款。

[61] 《欧盟通用数据保护条例》第 12 条第 6 款。

[62] 《欧盟通用数据保护条例》第 7 条数据主体的知情同意权、第 15 条数据访问权、第 16 条的不准确信息的纠正权、第 17 条的删除权、第 18 条的限制处理权。

绝和反对的权利。^[63]最后,在被决策者对相关信息都有了清楚的认知并明确表示同意引入后,行为人有义务严格按照所告知对方的使用目的和范围内以“合法、公正、透明的”^[64]方式操作和使用。

此外,行为人还应赋予对方更多的选择空间以及参与或者退出的自由。个人信息保护规范中立法者也明确要求,通过自动化决策进行商业营销、信息推送等行为的,应当同时提供不针对其个人特征的选项。并且,被决策者拒绝算法解决结果的,行为人应允许拒绝并保障其享有全面的退出机制。

（二）违约责任的认定与承担

如前所述,合同履行过程中,一方因便捷、高效和节省履约成本等原因而引入算法决策程序辅助或代替其自身决策,使相对方处于遭受算法歧视的风险之下的,使用算法程序的行为人就应当为自己的“先前危险行为”承担相应的附随义务,绝对地预防和避免算法歧视现象的发生。且按照合同义务“全面履行”原则的要求,行为人在未履行附随义务或义务履行不充分的情况下,理应为对方因此遭受的损失而承担相应的违约责任。并且,只有在行为人能够证明自己履行了义务、没有过错的情况下,才可以减轻或者免除责任。

上文所述的附随义务,属于“以维持利益为指向的保护义务”,与“以给付利益为指向的给付义务”相对应,设立的初衷在于实现保护功能,即维护合同履行过程中,对方当事人人身和财产上的维持利益和固有利益。^[65]因而,违反附随义务而导致的违约责任的承担,应以歧视后果发生、对方因遭受了歧视性待遇而使人身或财产遭受害为前提。歧视性待遇的认定,采无过错的“差异影响”审查模式,即无须考量主体实施行为或者规则的时候是否有歧视的故意,只需要审查行为或者规则实施以后,是否针对对方产生差异性的影响即可。^[66]关于责任承担方式,原则上而言,违反附随义务而导致的违约责任的承担,只能“就其所受损害,得依不完全履行规定,请求损害赔偿”。^[67]即,因合同附随义务的不履行或不完满履行而违约的,相对方不得要求解除合同,只能够请求对方承担违约责任,赔偿所遭受的损失。^[68]

四、结语

作为民法的基本原则,诚实信用原则功能的发挥,使民法完成了主体地位从抽象平等到具体平等,法律价值从追求形式正义到实质正义的变迁,现代民法从“概念法学”高度抽象的形式主义理性转换到对“伦理性的人格主义”的关注上。^[69]按照拉伦茨的观点,以缔约和合同关系为基础的契约法上的保护义务与侵权法中的交易安全义务在本质上没有区别,都是在“诚实信用”原则的基础上,为了保障当事人的完整利益(持续利益),而要求交往或者交易过程中“开启或维持”某种危险状态源头的人适当地承担一些法定义务,采取必要的、具备期待可能性的风险预防与保护措施,并在法律义

[63] 参见张欣:《算法解释权与算法治理路径研究》,载《中外法学》2019年第6期,第1425-1445页。

[64] 《欧盟通用数据保护条例》第5条“与个人数据处理相关的原则”。

[65] 韩世远:《合同法总论》,法律出版社2018年版,第348页。

[66] See Richard Primus, “The Future of Disparate Impact”, 108 *Michigan Law Review* 1341-1387(2010).

[67] 王泽鉴:《债法原理》(第1册),中国政法大学出版社2001年版,第40页。

[68] 参见韩世远:《合同法总论》,法律出版社2018年版,第343页。实践中亦有法官支持违反附随义务可请求“合同解除”的案例,但韩世远教授认为这个属于“特例”,作为与主给付义务相区分的特征之一,附随义务违反后原则上不能请求解除合同。

[69] 邱雪梅:《试论民法中的保护义务——“两分法”民事责任体系之反思》,载《环球法律评论》2007年第5期,第110-121页。

务未履行或未完满履行的情况下承担相应的缔约过失、违约或者侵权责任。^[70]但二者也有一定的区别:交易安全义务具备侵权法的对世性,是针对“来自公众任何一人”^[71]的义务;而保护义务则存在于特定的具有缔约或者合同关系的相对人之间,义务的主体和内容都要更加明确和具体。^[72]算法歧视责任形态的变化,伴随着侵权责任到缔约过失责任再到合同违约责任的过度,遵循的是主体间法律关系亲密度逐步加深、彼此间互负的保护义务内容和履行程度的要求也逐级递增的逻辑。在行为人的义务违反同时满足保护义务与安全交易义务所要求的内容时,必然会产生“责任竞合”现象。为了能够更好地实现对被决策者的权益保护,应当允许其根据具体情况,在缔约过失请求权和侵权请求权之间进行自由地权衡和选择。

此外,“现阶段的人工智能并不具备内源性行为能力,没有对主体间的权利义务关系造成颠覆性冲击,故有些问题尚可以在现有法律框架内通过条文的解释加以解决”,^[73]因而,以人类行为特征和模式为基础的法律逻辑和规制手段,在大数据和算法的时代背景下尚且能够继续适用。但也不能忽视,部分在传统认知框架中仍被视为“私主体”的公司、组织和平台,因在数据和信息获取能力、专业技术能力和社会资源掌握层面具有绝对的优势,而形成一种“准公权力”的效果,^[74]颠覆了以主体间法律地位平等为基础而建构的私法规范体系,使处于弱势地位的一方主体的人身和财产权益时刻处在被侵犯的边缘。面对这种因技术的变革和应用所导致的“以权力配置为代表的社会关系的变化”,^[75]新出台的《个人信息保护法》通过一系列原则、规则和相关主体权利义务内容的设定,为弱势的信息主体的权益保护提供了“倾斜性”的规范基础和依据,也为算法歧视发生后的事后追责与侵权救济指明了方向和路径。但事后的诉讼救济,不如事前的未雨绸缪。在条件成熟的前提下,算法歧视现象的规制,仍需要来自不同领域的专家和学者携手,探索和建构更加完善的事前性风险防范机制。比如,将技术伦理和规范直接编入算法决策程序中等,使社会在享受科技进步带来的便捷与高效的同时,也能够通过科技手段的运用,减少和避免负面影响和效应。

(责任编辑:陆宇峰)

[70] 参见李昊:《交易安全义务论:德国侵权行为法结构变迁的一种解读》,北京大学出版社2008年版,第79-85页。

[71] 林美惠:《侵权行为法上交易安全义务之研究》,台湾大学2000年法律学研究所博士论文,第37页;转引自李昊:《德国新债法中附随义务的构造》,载《环球法律评论》2009年第5期,第113-123页。

[72] 参见王利明:《合同法新问题研究》,中国社会科学出版社2003年版,第152页。

[73] 涂永前:《规制人工智能:一个原则性法律框架研究》,载《人工智能法学研究》2018年第1期,第119-132,231页。

[74] 郑戈:《算法的法律与法律的算法》,载《中国法律评论》2018年第2期,第66-85页。

[75] 沈伟伟:《算法透明原则的迷思——算法规制理论的批判》,载《环球法律评论》2019年第6期,第20-39页。