

算法背景下智能招聘歧视的法律规制

蒋 浦

澳门科技大学法学院，澳门

收稿日期：2023年1月4日；录用日期：2023年1月13日；发布日期：2023年3月3日

摘 要

在算法时代下，有时候不是人在主动的歧视，而是一个自动系统在编程的时候有意或无意的进行歧视。这类的歧视是在计算机系统的人智能算法决策中产生的，虽然计算机系统本身没有意识，但它产生出的决策却构成了歧视。在就业市场招聘程序通过人工智能算法决策参与到招聘的过程中，其虽简化了招聘的流程提升了招聘效率，但在某种情况下也可能构成对应聘者的就业歧视。本文将涉及这些问题，提出应建立算法歧视审查机制应设立专门的算法审查部门对算法问题进行规制。

关键词

人工智能招聘，算法歧视，算法规制

Legal Regulation of Intelligent Recruitment Discrimination in the Context of Algorithms

Pu Jiang

Faculty of Law, Macau University of Science and Technology, Macao

Received: Jan. 4th, 2023; accepted: Jan. 13th, 2023; published: Mar. 3rd, 2023

Abstract

In the algorithmic era, sometimes it is not a human being who is actively discriminating, but an automated system that is programmed to discriminate, either consciously or unconsciously. This type of discrimination is generated in the human-intelligent algorithmic decision making of a computer system, and although the computer system itself is not conscious, the decisions it generates constitute discrimination. In the employment marketplace recruitment processes are involved in the recruitment process through artificial intelligence algorithmic decisions, which streamline the recruitment process and improve recruitment efficiency, but in some cases may also constitute employment discrimination against candidates. This paper will address these issues and propose that

a review mechanism for algorithmic discrimination should be established and a special algorithmic review department should be set up to regulate algorithmic issues.

Keywords

Artificial Intelligence Recruitment, Algorithmic Discrimination, Algorithmic Law Regulation

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 问题的提出

我国是劳动力资源最丰富的国家，在过去有超九成的企业招聘决策完全依赖于人工完成，不仅需要求职者花费大量时间寻找合适企业和岗位，企业的 HR 每天要耗费大量精力筛选简历。传统招聘的模式会花掉企业 HR 数周甚至数月的时间。随着人工智能时代的来临，自动化决策算法被应用于劳动市场的各个阶段，在劳动关系建立的初步阶段，人工智能技术的飞速发展使得 BOSS 直聘、智联招聘、前程无忧等以算法利用为核心的网络招聘平台成为各大企业人才招聘的主要路径之一。目前，在 1000 人以上规模的企业中，招聘环节应用 AI 技术的占比约 35%，1000 人以下规模的企业中，招聘环节应用 AI 技术的企业占比约 26%，它们应用 AI 技术最多的场景则是通过外呼机器人与求职者沟通求职意向[1]。随着时代的发展，用人单位在雇佣决策上对算法技术的依赖性不断提高。用 AI 进行雇佣决策看似排除了的用人单位主观偏见，但是招聘算法技术真正做到无偏见、无歧视了吗？人工智能技术应用引发的就业歧视相较传统就业歧视而言更加隐蔽，在现实实践的招聘环节中可能将妇女、残疾人、老年人等弱势群体置于不利的就业地位。虽然目前我国还未出现人工智能算法产生的国内就业歧视案例，但随着人工智能技术、算法技术在企业中应用，我们不难预测算法技术参与招聘决策所引发的就业歧视将会成为未来招聘歧视争议的焦点。

2. 人工智能及算法概念

2.1. 人工智能的概念

1950 年，阿兰·图灵在《计算机与智能》《机器人能思考吗？》中提出了“图灵测试”，从此，展开关于机器人思维的问题讨论。1956，达特茅斯会议的召开，标着这“人工智能”这一词的正式诞生[2]“人工智能”这一词一种受到各种定义的影响，强调智能的不同方向。只有被人类普遍认为是智能的，及所由的智能定义都与人类有关；有些是侧重人类的理性思维，有些则围绕着人类行为。即使目前人工智能技术还未达到与人类认知思维相类似的能力，但近年来使用人工智能代理的尝试已经获得了相当大的发展，即通过人工“智能代理”，通过编程来实现某些目的和结果。所谓的机器学习是指这些智能代理从数据中进行学习，在这些数据集的基础上运行一个规定的算法来完成某个目标，在一个不熟悉的环境情况中，智能代理将利用这些数据从而获得出一个最佳的结果，并随着时间的推移根据其所积累到的结果继续调整它们的行为。与传统的统计技术行为不同的是，机器学习是非参数化的，不涉及任何的特定数学公式，而是将结果变量表达为可以解释的变量函数并应用到数据上。由于机器学习不需要现有的知识和对变量之间的关系的仔细识别，它可以适用于更为广泛的问题，其与人类判断或基于统计的方法相比，它可以提供更好的预测，机器学习系统也能快速的适应时间的变化，当使用者给予机器新的数据

时，算法系统会立刻开始建立新的模式并完善以前的预测。近年来，机器学习技术已经表现出令人印象深刻的进步和日益普及，在过去，机器学习是一种成本较为高昂的一种技术，而如今，随着技术和资金降低，谷歌、微软、亚马逊和苹果等大型公司以及无数规模的小型企业都可以使用机器学习技术。

2.2. 算法的概念

算法(Algorithmic)是用于解决问题或执行计算的过程。算法充当指令的精确列表，在基于硬件或软件的例程中逐步执行指定的操作。算法广泛应用于 IT 的所有领域。在数学和计算机科学中，算法通常是指一个可以解决反复出现的问题的一个小过程。算法也被用作执行数据处理的规范，并在自动化系统中发挥重要作用。算法可用于对数字集进行排序或更复杂的任务，例如在社交媒体上推荐用户内容。算法通常从描述特定计算的初始输入和指令开始。执行计算时，进程将生成输出。

2.3. 算法偏差的概念

算法偏差(Algorithmic Bias)是算法是指计算机系统上的系统性和可重复性产生的错误，这些错误可能会产生“不公平”的结果，例如，以与算法预期功能不同的方式“优先”一个类别而不是另一个类别。算法的偏差可能来自许多原因，包括但不限于算法的设计或与数据编码、收集、选择或用于训练算法的程序相关的意外或意外的决策。例如，在搜索引擎结果和社交媒体平台中观察到算法偏差。这种偏见可能会产生各种影响，在无意中侵犯隐私种族、性别、性和民族的社会偏见。它还出现在刑事司法、医疗保健和招聘方面，加剧了现有的种族、社会经济和性别偏见。

3. 人工智能技术在招聘中的应用

人工智能招聘是将机器学习或预测算法等人工智能运用到招聘过程中的做法，其根本的目的是招聘人员和团队希望运用自动化招聘流程，使得招聘的过程更快，招聘的成本更便宜，更容易的管理大量的候选人。

3.1. 人工智能招聘兴起的原因

人工智能是提高招聘效率的有力工具，通过渐少用于筛查的时间、技能测试、采访。使得人力资源团队可以专注于吸引和留着员工的方法。根据美国行业分析师分析，超过一半的人才招聘负责人表示，“招聘中最困难的环节是从大量的申请者中找到合适的候选人。”更困难的是，56%的受访者预计招聘规模会增加，而 66%的招聘团队要么保持原有规模，要么缩小规模。简而言之，公司用更少的人力资源雇佣更多的人。¹通过人工智能技术可以帮公司的招聘团队提高他们的招聘质量节省时间和金钱，并尽量减少一些隐藏的人类偏见，阻止少数候选人获得公平评估。算法在就业方面可以说是发挥不同的作用。在求职者开始申请之前，由算法决定的针对特殊人群的广告，将求职者引向某些职位和职业类型。一旦求职者开始申请某个特定的工作，公司就可以使用求职者追踪系统(ATS)来扫描简历中的关键字，或者将求职者的简历样本与之前被雇佣员工的简历进行比较[3]。上述系统可以根据简历放弃某些申请人，而将其他候选人推向面试，与此同时在面试过程中，公司还可以使用视频面试算法，如 HireVue，来分析申请者在面试中的声音和面部反映，从而给申请者的面试打分。2020 年 5 月，网络求职平台“智联招聘”推出“AI 易面”这一功能，旨在通过对求职者进行语义分析、神态分析、智力评测、性格测试等一系列结构化测试，智能评判求职者是否符合预设的候选人画像，借此快捷完成人岗匹配[4]。

3.2. 人工智能技术在就业招聘中的歧视风险

在如今的招聘中，候选人在申请工作的时候越来越有可能面对算法，因为超过 95%的财富 500 强排

¹人工智能招聘：人力资源专业人员的权威指南 <https://ideal.com/ai-recruiting/>, 2022 年 4 月 7 日访问。

行榜的公司正在招聘中使用 ATS, 算法正在决定雇用谁。²以希尔顿酒店为例。它使用聊天机器人对数千名呼叫中心候选人进行初步评估。那些进入下一步的人将接受视频面试, 在那里他们回答一系列问题。面试者面部表情的反映通过招聘算法进行分析, 从而进行打分。美国希尔顿大量需求雇用呼叫中心工作的员工, 尤其是在假期前后。该岗位招聘 1200 人, 其收到了超过 30,000 名申请人的申请, 据希尔顿全球招聘副总裁 Sarah Smart 称。招聘算法带来好处: 包括降低招聘成本和加快招聘时间, 这一切要归功于自动化, 希尔顿使用位于加利福尼亚州帕洛阿尔托的人工智能招聘工具供应商 AllyO 的聊天机器人进行初步候选人评估。例如, 聊天机器人会询问候选人可以工作的时间, 并且由于这些呼叫中心职位是远程的, 候选人可以通过互联网工作。如果候选人通过了初步评估, 候选人可以进行视频面试, 希尔顿采用 HireVue 的技术筛选面试者, 该技术能够分析候选人如何回答问题, 以及在几分之一秒内发生的微表情或快速面部表情。³在希尔顿的案例中, 招聘算法将招聘人员处理呼叫中心招聘的时间减少了 23%。该公司将剩余的时间转移到其他任务上。希尔顿使用的技术不会将人力资源从招聘流程中移除, 但招聘算法会提出建议, 并可以优先考虑最适合该工作的申请人。在线零售商亚马逊(Amazon)的全球员工中有 60% 是男性, 男性占公司管理职位的 74%, 最近在发现性别偏见后停止使用招聘算法。⁴工程师用于创建算法的数据来自提交给亚马逊的 10 年简历, 这些简历主要来自白人男性。该算法被教导识别简历中的单词模式, 而不是相关的技能组合, 并且这些数据与公司以男性为主的工程部门进行基准测试, 以确定申请人的合适性。结果, 人工智能软件惩罚了任何文本中包含“女性”一词的简历, 并降低了就读女子学院的女性的简历, 导致性别偏见。⁵上述情况在我国也常有发生, 当一些用人单位在招聘软件如: 智联招聘、BOSS 直聘、猎聘等, 输入“程序员”时, 男性求职者的简历会优先被推送出来; 当输入“前台”时, 系统则会优先将女性求职者的简历推送出来[5]。

3.3. 人工智能简历筛选中算法歧视的特性

3.3.1. 算法歧视隐蔽性

随着人工智能技术的迅速发展, 人工智能技术造成的就业歧视的司法认定也随之难度增大。首先, 部分招聘软件存在定向推送的情况, 该种定向的推送不仅很难被发觉而且其难以被认定为就业歧视。再者, 定向招聘的推送中并没有出现对任何群体歧视的字眼, 只是在定向推送的背后, 通过深度算法对现有的数据样本进行批量分析后生成的自动决策结果, 该决策会导致部分招聘信息只推送给特定人群从而产生歧视性后果(例如仅仅向适龄的男性人群推送特定岗位)。那些未能接受到推送信息的其他求职者(例如能力相仿的适龄女性人群)实际上被剥夺了知悉该就业岗位的权利。用人企业通过以深度算法为核心的自动决策方式对大量应聘者简历审核, 看起来是依赖于这种中立性技术通过智能的方式对求职者工作能力、学历层次、年龄、身高、相貌、婚姻状况、健康情况等做出的客观分析。从最终结果来看, 是否录取的决策并不是用人单位主观偏好的结果, 而是通过人工智能算法技术决定的, 因此用人单位很容易以“工作相关性”和“与岗位职责一致性”为理由对未获得面试机会的应聘者答复。

² 参见 Kerri Anne Renzulli. (2019 February 28). 75% of resumes are never read by a human—here's how to make sure your resume beats the bots. CNBC Make It. <https://www.cnbc.com/2019/02/28/resume-how-yours-can-beat-the-applicant-tracking-system.html>, 2022 年 12 月 1 日访问。

³ 参见: Hiring algorithms prove beneficial, but also raise ethical questions <https://www.techtarget.com/searchhrsoftware/news/252471753/Hiring-algorithms-prove-beneficial-but-also-raise-ethical-questions>, 2022 年 12 月 10 日访问。

⁴ 参见 Why it's totally unsurprising that Amazon's recruitment AI was biased against women <https://www.businessinsider.com/amazon-ai-biased-against-women-no-surprise-sandra-wachter-2018-10>, 2022 年 12 月 10 日访问。

⁵ 参见 Amazon reportedly scraps internal AI recruiting tool that was biased against women <https://www.theverge.com/2018/10/10/17958784/ai-recruiting-tool-bias-amazon-report>, 2022 年 12 月 10 日访问。

3.3.2. 算法歧视复杂性

即使雇主不直接依靠算法来做决定,一种被称之为自动化偏差(as automation bias)的现象也会为算法创造机会。当算法预测的数字或者分数排名被准确客观的呈现出来时,负责评判的人可能给予它们超出实际的分量或者青睐,但这其实是不是雇主想要的。自动化偏差是危险的,因为随着算法变得越来越复杂和广泛应用,人们就更倾向于遵循和使算法决策,而不是真正的了解算法决策的细节。正如法律学者 Danielle Citron 所解释的:“自动化偏见有效的将计算机程序的建议转化为可以信任的最终决定。”[6]此外人们不仅仅可能会无法审查计算机算法的偏见,而且随着时间的推移,电脑的使用者可能会失去检查电脑系统建议的技能。人工智能是算法“黑箱”(Black box)。人工智能对大数据的深度学习,不断的通过计算和学习从而产生预测或决策,最后可能连该算法程序的设计者都不知道系统最终会做出什么样的决策。算法最终可能通过现有的数据和以往的招聘人员自动做出歧视性招聘决策。这并非自动化系统开发商和使用方的故意行为,因此很难清晰地追究责任[7]。

4. 简历筛选算法歧视规制的原则

4.1. 基于平等权原则

平等权是一项兼具原则性的基本权利,其既可以作为一项具体的基本权利,也可以作为关于基本权利的原则,因此平等权会涉及到各种类型的基本权利,例如,人身层面的平等权、政治层面的平等权,社会经济层面的平等权等等[8]。劳动法第一次规定了平等就业权,是当代中国反就业歧视法的正源[9]。“平等”与歧视相对立。《劳动法》第12条规定:“劳动者就业,不因民族、种族、性别、宗教信仰不同而受歧视。”根据立法者的解释,该条的意义是:“每个劳动者在劳动市场求职的竞争中享有同等的权利,站在同一条起跑线上,享受同等的机会,意思就是平等竞争。”[10]由此可见就业的平等权是指竞争机会平等,平等竞争是将法律所禁止的事项从人为的决策中排除出去。然而算法的出现该平等理论面临巨大挑战,原本人为的决策现在将有一个个计算机程序根据其设定的算法做出非人为的决策从而通过算法歧视损害求职者的基本平等权利。另外,平等权作为宪法上的基本权利,近现代随着宪法的颁布进而有效的实施,世界上多数民主国家已经将这项权利宪法化。平等权也是我国宪法规定的基本权利体系中的重要组成部分。我国宪法第33条第2款规定了平等权条款:“中华人民共和国公民在法律面前一律平等。”但平等权作为宪法上的一项基本权利,其自身并无具体的内容,这也就意味着平等权只有在权利主体行使其他权利时才能实现,如通过实现政治、经济、文化、社会、家庭的平等来体现其权利价值。例如,在劳动市场在大规模招聘软件出现前,人们大多数是通过参加线下招聘会或者通过亲友介绍等方式获取招聘信息,但随着招聘软件的兴起,越来越多的人开始通过线上的方式寻求工作机会。应聘者将自己的简历信息上传到招聘平台,而招聘平台和用人单位基于大数据分析算法为不同的应聘者提供个性化的信息和推送,形成针对特定人的特殊规则,可能使得应聘者获取单一的招聘信息、工作岗位等。从而可能导致人工智能算法歧视侵犯公民在宪法上的平等权这一结果的出现。

4.2. 算法程序决策正当化原则

算法自动化决策对公民权利影响巨大,却可以规避正当程序审查,当大规模的自动化决策算法被引入到智能招聘当中,当企业树立了以用人绩效为主流的价值观,认为使用自动化算法技术筛选应聘者就是自发的选择同等薪酬下的资质最佳的劳动者,属于正当性的“择优录取”,是企业用人的自主权,企业招聘时,传统招聘流程需要相关人事部门进行会议讨论,对其将要招聘的人进行观点表达,也有直接的相关标准和公司制度,但是在算法进行简历筛选时,求职者无法明确得知自己的求职结果有无遭受到算法决策时针对不同因素采取不同权重的侵权行为。算法自动化决策是披着算法程序知识产权的外衣,

从私法的层面上来说,确实不宜公开透明,否则会产生侵权甚至阻碍创新的问题。但从法律规制的层面来说,如果对算法歧视的法律规制不能对算法产生质疑时进行公开的解释,一直拒绝对利害关系人公开。应用于美国社区管理的“真知”算法设计者查德·柏克坦陈到:“不透明的算法可以帮我摆脱一切的麻烦,我在做出对人不利的预测时,从来不会说明理由,我唯一的目标就是做出正确的预测。”^[11]此外,很多算法的使用者兼具公共性特征与私主体特征。在互联网企业兴起的今天,很多企业的公共性属性与商业性属性已经变得很难区分。对于阿里、腾讯、百度这样的龙头企业来说,很难以纯粹的市场私主体看待它们,已经与实际情况不符。从数据、资源、算法到服务,这些企业已经与人们的日常生活捆绑,从而具有部分准公权力的性质或者公权力的一些属性^[12]。当公权力机关对当事人做出不利决定时,当事人有权利知道决定的内容和理由,同时给予了当事人申请陈述和救济的途径^[13]。我国司法程序中也在实际运用正当程序原则审查具体的行政行为^[14]。随着上述大型机构企业私主体和公权力的界限日益模糊,“私行政”的理念也应要求国家对其组织形式和程序负有监督的义务^[15]。正当程序就是我们每个人都要受到法律的个别对待,在正当程序下,受不利决策影响的人应当有权知晓决定产生的理由和决定的内容,并且享有诉讼的权利和抗辩的机会。

5. 针对 AI 招聘算法歧视的应对措施

5.1. 我国目前对算法歧视问题立法现状

目前从立法层面上来看,算法招聘歧视陷入因果关系和责任认定困难的局面,传统的侵权责任适用陷入困境。目前,我们现在并没有一般化、可以普遍适用的人工智能和算法的法律,我国法律法规对算法招聘歧视的规制仅仅来源于综合性法律规制。《个人信息保护法》、《网络安全法》等法律提出透明性原则、最小侵害原则等数据处理原则以确保数据挖掘与利用的合法性与正确性,同时赋予个人信息所有者访问权、更正权、删除权等数据处理的权力,对个人信息处理者的行为进行事前规制^[16]。我国《互联网信息服务算法推荐管理规定》⁶制定更具有针对性的算法推荐规定,明确了算法推荐服务提供者的主体责任。其中第4条明确了算法推荐服务的公平公正原则,第20条“算法推荐服务提供者向劳动者提供工作调度服务的,应当保护劳动者取得劳动报酬、休息休假等合法权益,建立完善平台订单分配、报酬构成及支付、工作时间、奖惩等相关算法”是对劳动用工领域的精准规制,但局限于劳动关系存续期间平台用工中对劳动者劳动报酬、休息休假等权益的保护,并未提及招聘阶段的就业歧视问题。根据国家互联网信息办公室关于发布互联网信息服务算法备案信息的布境内互联网信息服务算法名称及备案编号⁷,在2022年10月公布的数据中,涉及对家互联网企业对其的算法进行备案,总计70条算法备案信息。其中涉及招聘平台的有智联招聘其备案了社区首页推荐算法,主要应用于招聘资讯推荐,根据用户属性信息和行为信息,推荐面试经验和职业规划。猎聘备案信息搜索算法,主要应用于搜索功能,根据用户输入的关键词,为用户提供候选职位信息。⁸

5.2. 学术界对算法归责问题的探讨

而与事前规制相比,我国对事后规制的研究与实践尚未成熟,法律未对招聘中所发生的算法歧视的责任承担予以明确,劳动者在受到算法招聘歧视后将陷入求法无门的困境。在理论学界,学者对于算法歧视的归责问题也未统一意见。中国社会科学院大学法学院冯珏教授认为可以适用如产品责任等传统责任框架的规制路径对其进行规制^[17]。中国人民公安大学法学院副教授苏宇认为应将产品责任中的谨慎标

⁶ 参见互联网信息服务算法推荐管理规定 http://www.cac.gov.cn/2022-01/04/c_1642894606364259.htm。

⁷ 参见国家互联网信息办公室关于发布互联网信息服务算法备案信息的公告 - 中共中央网络安全和信息化委员会办公室) http://www.cac.gov.cn/2022-08/12/c_1661927474338504.htm, 2022年12月21日访问。

⁸ 以上数据由国家互联网信息办公室关于发布互联网信息服务算法备案信息的公告的附件2:境内互联网信息服务算法备案清单(2022年10月)整理得出。

准延伸至人工智能产品领域,其提议健全完善谨慎算法标准从而对算法进行归责[18]。同时,也有学者认为侵权行为的责任主体应该是人工智能,而不是具体操作者。华东政法大学教授刘宪权认为人工智能具有自我辨认能力与控制能力,人工智能拥有做出行为的选择权,故应该独立承担责任,若将侵害他人的行为完全归责于研发者与使用者实为不妥[19]。中国政法大学法学院张凌寒教授认为《个人信息保护法》第24条规定“个人信息处理者利用个人信息进行自动化决策,应当保证决策的透明度和结果公平、公正”体现立法者的技术治理思路,直接将平台算法作为法律监管对象,而不再纠结主观过错与违法结果等传统法律归责的认知,确立了平台算法问责制与以平台治理为核心的算法自动决策治理框架[20]。综上所述目前我国立法实践和理论学界关注点多聚焦于算法技术在应用过程中的事前规制层面,对于算法技术事后规制层面有所不足,对于算法归责问题也各抒己见,未达成统一的结论。因此必要结合算法技术在招聘过程中的应用场景与技术特征对传统侵权责任规制机制进行研究与探讨,在现有法律框架之下寻找到算法招聘歧视归责的新路径。

5.3. 域外针对算法歧视的立法保护

在全世界范围内如何解决就业中的歧视都是监管难度较大的难题,为解决这一问题,美国专门建立了平等就业机会委员会(Equal Employment Opportunity Commission)。该机构《1964年民权法案》实施机制的核心,作为重要的独立联邦的行政执法机构其在美国反就业歧视立法和执法方面起着重要的作用[21]。另一面,欧盟在其《人工智能法案》(Artificial Intelligence Act)中提出了算法审查机制,法案对人工智能技术划分风险等级,分别是不可接受的风险、高风险、有限风险和最低风险,对以上不同风险级别的人工智能系统使用采取不同的限制措施。该法案建议对存在高风险的人工智能产品和服务引进 CE (Conformité Européenne) 认证,为获取认证 AI 开发者和使用者需严格履行七种义务,包括完备的风险评估系统;向系统提供高质量的数据集,将风险和歧视性结果降至最低;记录活动以确保结果的可追溯性;提供有关系统及其目的的所有必要信息,以供当局评估其合规性;向用户提供清晰、充分的信息;采取适当的人为监督措施,将风险降至最低;高水平的安全性和准确性。除了严格的前置管理程序外,法案还要求对涉及高风险的人工智能系统进行全过程监督和合规评估。对于有限风险和最低风险的人工智能系统,法案没有设置前置审查程序。⁹总体来说,目前世界各个范围内的算法监管部门逐渐明晰,并开始走向专业化、扩张化的道路,由于算法的专业性问题,成立专门的部门对其进行规制和审查也是算法监管的一个新趋势。在处理有偏见的结果时,监管机构大多依靠标准的反歧视立法。只要有人可以对有问题的决定负责,这是可行的。但随着人工智能的日益普及,个人责任感受到削弱。更糟糕的是,人工智能增加了潜在的偏见规模:任何缺陷都可能影响数百万人,使公司面临历史性的集体诉讼,并使他们的声誉面临风险。

就像人类的判断一样,人工智能并非万无一失。算法、人工智能必然有偏见,但这不是算法的错误,因为算法的偏见是反映人的偏见。人工智能或者算法都是用数据“喂养”起来的,若本身的数据含有偏见的成分,那么算法将不可避免地做出一些不公平甚至不安全的决定。当人们犯错时,通常会进行调查和分配责任,这可能会对决策者施加法律处罚。这有助于组织或社区理解和纠正不公平的决定,并与利益相关者建立信任。那么,我们是否应该要求——我们甚至可以期待——人工智能来解释它的决定吗?

监管机构肯定正在朝着这个方向前进。GDPR 已经描述了“正确的通过算法获得所达成决定的解释”,欧盟在其白皮书和人工智能监管提案中将可解释性确定为增加对人工智能信任的关键因素。但是,对自动决策的解释意味着什么,我们对因果关系的了解往往是不完整的?在这种情况下,解释结果如何得出的能力可能不如重现结果和经验验证其准确性的能力重要——公司可以通过将人工智能的预测与结果进

⁹参见全球首部人工智能立法:创新和规范之间的艰难平衡 <http://www.rmlt.com.cn/2022/0302/641113.shtml>。

行比较来做到这一点。为算法提供全局解释似乎很简单：您所要做的就是分享其公式。然而，大多数人缺乏理解这样一个公式所需的数学或计算机科学方面的高级技能，更不用说确定其中指定的关系是否合适了。在机器学习的情况下，人工智能软件创建算法来描述训练数据中变量之间的明显关系，数据中的缺陷或偏差，而不是算法，可能是任何问题的最终原因。此外，公司甚至可能无法直接了解其算法的运作情况，并且响应监管限制的解釋可能需要他们超越数据和计算机部门，甚至可能着眼于外部专家。

6. 结论

随着算法技术的不断成熟和应用在生活中各个领域，算法歧视也随之产生在越来越多的地方，价格歧视、就业歧视、信用歧视等字眼也不断的出现在各种新闻和新浪微博热搜中。相对于前文中学者提出算法事前规制或者算法公开的方式，对于事前禁止算法的规制方式而言，一方面，如何在事前定义算法“本身违法”是一个有待解决的难题，在事前预估其可能造成的后果在事实上可能具有一定的实践难度，并且伤害程度如何衡量也需要进一步的探讨。尤其在算法决策过程存在不透明以及不可解释性时，排除或确定可能的危害和危害程度更是难上加难。另一方面，这种监管方式的有效性可能受限于算法结果的不可预测性。算法输出的结果不仅依赖于算法本身的性质和设计思路，也在很大程度上依赖于“喂入”的算法的数据，算法在被投放于市场或投入使用之后，其所处理的数据可能较为庞杂并且本身存在一定的偏差，因而造成输出结果不符合预期的情况。此外随着人工智能技术的不断发展，各个企业也随之会对其的算法进行更新和优化，其更新的算法可能与之备案或者公布的有不同的地方，这也会导致事前规制的难度加大。

算法歧视相对传统的歧视不同，其具有隐蔽性、复杂性等特点，如果想要对不同领域产生的算法歧视进行审查和规制，不仅存在规制成本较高且短时间难以实施的困境，还存在着难以发现和认定等问题，基于算法的监管和审查是一门既跨学科，又跨部门的流程体系。因此建立一个全国性的算法审查机关是一个值得考虑的选择。其次，支持公民提起算法歧视诉讼，算法歧视仍然属于传统歧视的范畴。针对歧视，我国目前并无专门的反歧视行政执法机构，只在部门法中规定了歧视的司法救济，如《就业促进法》规定，实施就业歧视的，劳动者可以向人民法院提起诉讼。虽然实践中存在受害人因歧视提起诉讼维权成功的案例，但由于我国反歧视立法的不健全，人民法院认定构成歧视还是非常艰难的。并且，在算法歧视案件中若个人提起民事诉讼，由于对被歧视者存在对算法技术缺乏了解、同时又难以获得企业内部算法运行程序的数据、从而产生受到的歧视难以证明等问题，受歧视者往往在调查取证方面处于弱势地位，导致其胜诉可能性较低；加上司法诉讼的时间和金钱成本相对较高，现实中很多受歧视者会选择放弃救济。在美国，国会最初只授予了平等就业机会委员会调查和和解的权限，但是后来国会扩大了其职能赋予其针对非法雇佣行为提起民事诉讼的权力，其可以选择以自己的名义代表群体受害人提起诉讼，也可以加入由原告提起的民事诉讼。针对简历筛选中的可能遭受算法歧视的应聘者面对着信息不对称、举证麻烦、集体维权困难、起诉成本过高、起诉时间长、胜诉几率过低等问题，专门建立专门的算法审查部门将有利于根据受歧视者的具体情况有针对性地采取相应维权措施，降低维权成本和解决举证困难等问题，从而实现应聘者平等权的保护。不过公权力机构行使或代为行使的方式也会面临诸多挑战，例如公共机构代为请求赔偿是否意味着该等算法的损害要达到一定的额度或造成公共利益损害时才可以起诉？如何分配举证责任？上述这些问题也需要在未来进一步的论证。

在人工智能时代，数据如同奔涌的江流向我们席卷而来，各式各样的算法在呈现在我们每天的生活当中，不难预见未来利用人工智能算法进行简历筛选甚至决策招聘结果是一种常态化的方式，同时使用各式各样的智能招聘平台也在潜移默化的改变人们求职方式尤其是刚出校门的大学生们。因此，在未来要怎么样减少和降低应聘者因算法而产生的不公正风险呢？算法要怎么样保持中立和做到公平正义？人

们要怎么防止用人单位利用算法进行直接或间接的歧视呢？这些都是未来人工智能招聘发展所面临的必然问题。我国现有的算法治理与数字平台的治理，数据保护立法才刚刚起步，目前规范算法技术的法律法规体系主要聚焦于事前的规制，对于个人信息数据保护与个人信息权益维护的思维难以覆盖招聘时算法技术广泛利用个人信息数据以及本身设计偏差所导致的算法歧视风险，以及后续对算法歧视的法律救济。《个人信息保护法》已经为算法歧视发生后的事后追责与侵权救济指明了方向与路径。域外算法的规制经验也可以为我国算法规制以及建立算法保护平台提供借鉴思路。法律总是在为科学技术进步所产生的影响做最后一道防线，希望社会在追求技术进步的同时也不忘为那些被技术进步所影响的人提供法律救济，兼顾高新技术发展与人类权益保障，将数据保护和算法规制同时推进，促进我国数据法治治理体系的建立。

参考文献

- [1] 徐语聪. 王露颐: 招聘进入 AI 时代解决传统招聘诸多痛点[J]. 数据, 2021(4): 30-36.
- [2] 张成岗. 人工智能时代: 技术发展、风险挑战与秩序重构[J]. 南京社会科学, 2018(5): 42-52.
- [3] Ajunwa, I. (2019) An Auditing Imperative for Automated Hiring Systems. *Harvard Journal of Law & Technology*, 34.
- [4] 蓝亭亭. 招聘算法歧视的法律规制[J]. 武汉交通职业学院学报, 2022, 24(4): 80-85+108.
- [5] 程建华, 刘菲菲. AI-HR 招聘的算法歧视及应对[J]. 常州工学院学报(社会科学版), 2022, 40(4): 123-128.
- [6] Citron, D.K. (2007) Technological due Process. *Washington University Law Review*, 85, 1249-1313.
- [7] 侯玲玲, 王超. 人工智能: 就业歧视法律规制的新挑战及其应对[J]. 华东理工大学学报(社会科学版), 2021, 36(1): 1-16+33.
- [8] 韩大元, 王建学. 基本权利与宪法判例[M]. 第 2 版. 北京: 中国人民大学出版社, 2021: 35
- [9] 阎天. 重思中国反就业歧视法的当代兴起[J]. 中外法学, 2012, 24(3): 560-577.
- [10] 全国人大委员会法制工作委员会国家行政法室. 中华人民共和国劳动法释义[M]. 北京: 中国工人出版社, 1994.
- [11] Angwin, J., Larson, J., Mattu, S. and Kirchner, L. (2016) Machine Bias. In: Martin, K., Ed., *Ethics of Data and Analytics*, Auerbach Publications, New York, 254-264. <https://doi.org/10.1201/9781003278290-37>
- [12] 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, 36(4): 37-46. <https://doi.org/10.16290/j.cnki.1674-5205.2018.04.026>
- [13] 周佑勇. 行政法的正当程序原则[J]. 中国社会科学, 2004(4): 115-124.
- [14] 何海波. 司法判决中的正当程序原则[J]. 法学研究, 2009, 31(1): 124-146.
- [15] 杜仪方. 公私协作中国家责任理论的新发展——以日本判决为中心的考察[J]. 当代法学, 2015, 29(3): 39-49.
- [16] 王莹. 算法侵害类型化研究与法律应对: 以《个人信息保护法》为基点的算法规制扩展构想[J]. 法制与社会发展, 2021, 27(6): 133-153
- [17] 冯珏. 自动驾驶汽车致损的民事侵权责任[J]. 中国法学, 2018(6): 109-132
- [18] 苏宇. 算法规制的谱系[J]. 中国法学, 2020(3): 165-184.
- [19] 刘宪权. 人工智能时代刑事责任与刑罚体系的重构[J]. 政治与法律, 2018(3): 89-99.
- [20] 张凌寒. 《个人信息保护法(草案)》中的平台算法问责制及其完善[J]. 经贸法律评论, 2021(1): 36-46
- [21] Lynch, M.K. (1990) The Equal Employment Opportunity Commission: Comments on The Agency And Its Role In Employment Discrimination Law. *Georgia Journal of International & Comparative Law*, 20, 89-104.