

算法歧视视域下妇女平等就业权的保护

董羽涵

青岛大学法学院, 山东 青岛

收稿日期: 2025年9月22日; 录用日期: 2025年10月15日; 发布日期: 2025年10月24日

摘要

本文以人工智能时代算法决策对妇女就业权益的影响为研究对象, 聚焦算法性别歧视对妇女平等就业权的侵害问题。通过分析算法性别歧视的内涵界定与特征, 揭示其与传统就业性别歧视在成因、影响范围及法律认定上的差异。当前在算法就业歧视领域面临认定标准模糊、归责原则不统一、举证责任分配不合理及责任主体缺位等法律困境, 进而提出优化路径, 包括厘清算法歧视的认定标准、明确归责原则、完善举证责任倒置制度及合理分配算法主体法律责任, 以期为保障妇女平等就业权提供法律与实践层面的参考。

关键词

算法歧视, 平等就业权, 性别歧视, 算法性别歧视

Safeguarding Women's Equal Employment Rights in the Context of Algorithmic Discrimination

Yuhan Dong

Law School, Qingdao University, Qingdao Shandong

Received: September 22, 2025; accepted: October 15, 2025; published: October 24, 2025

Abstract

This paper examines the impact of algorithmic decision-making in the era of artificial intelligence on women's employment rights, focusing specifically on how algorithmic gender discrimination infringes upon women's right to equal employment. By analyzing the definition and characteristics of algorithmic gender discrimination, it reveals the differences between such discrimination and

文章引用: 董羽涵. 算法歧视视域下妇女平等就业权的保护[J]. 争议解决, 2025, 11(10): 281-287.

DOI: 10.12677/ds.2025.1110335

traditional employment gender discrimination in terms of causes, scope of impact, and legal recognition. Currently, the field of algorithmic employment discrimination faces legal challenges such as ambiguous criteria for identifying discrimination, inconsistent principles of imputation, unreasonable allocation of the burden of proof, and the absence of accountable entities. Consequently, this paper proposes optimized pathways, including clarifying the criteria for identifying algorithmic discrimination, defining principles of imputation, improving the system of inverted burden of proof, and reasonably allocating legal responsibilities among algorithmic entities. The aim is to provide references at both legal and practical levels for safeguarding women's right to equal employment.

Keywords

Algorithmic Discrimination, Equal Employment Rights, Gender Discrimination, Algorithmic Gender Discrimination

Copyright © 2025 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. 引言

在传统就业模式下，妇女就业本就面临诸多挑战，全国人大代表张祖涛受访时表示，2025 届全国高校普通毕业生规模预计达 1222 万人，较 2024 年同比增加 43 万人。调查显示男性求职者接受面试邀请的次数是女性的 1.42 倍，男性获得 offer 的比例比女性高 13.1 个百分点，男性的教育回报率明显大于女性，且随着学历的升高回报率差距越明显^[1]。《中华人民共和国劳动法》(以下简称《劳动法》)第 12 条¹与《中华人民共和国劳动就业促进法》(以下简称《就业促进法》)第 3 条²均规定不得就“性别”实施就业歧视。《中华人民共和国妇女权益保障法》(以下简称《妇女权益保障法》)第 41 条³规定妇女享有与男子平等的劳动权利。

随着人工智能技术在招聘、绩效评估、薪酬管理等就业领域的深度应用，算法自动化决策是否扩大了在就业市场的不公，是一个现实存在且备受关注的问题。AI 招聘系统能够向潜在求职者推送招聘广告，对海量简历进行解析、筛选和匹配，为企业降低了招聘成本，提高了招聘效率。然而，算法歧视如同隐藏在黑暗中的“幽灵”，对妇女平等就业权构成了严重威胁。那么算法就可能会从这些历史数据中识别出用人单位对男性职工的偏好，并复制和固化这种偏见，做出不利于女性职工的算法雇佣决策，从而形成了就业性别歧视。本文的目标并非追求算法时代下男女之间的绝对无差别对待，而是旨在算法自动化决策主导的社会劳动生产下，女性劳动者的处境不至于恶化，并能够得到有效的法律救济。

2. 算法性别歧视的概念界定与特征分析

(一) 算法性别歧视的内涵界定

算法歧视诞生于算法决策的过程中。算法决策运作的基本原理是利用算法对大数据中的特征、关系、类别进行挖掘识别，而后借助算法模型来预测对象的行为，并以此为基础作出相应的决策^[2]。具体到算法雇佣决策的语境下，首先需要由用人单位制定决策目标并收集合适的训练数据集，算法程序员根据问题的类型选择并优化算法，并将用人单位的招聘需求转译为具有相关性的可测量特征或称“结果变量”，

¹ 《中华人民共和国劳动法》第十二条：“劳动者就业，不因民族、种族、性别、宗教信仰不同而受歧视。”

² 《中华人民共和国劳动就业促进法》第三条：“劳动者依法享有平等就业和自主择业的权利。劳动者就业，不因民族、种族、性别、宗教信仰等不同而受歧视。”

³ 《中华人民共和国妇女权益保障法》第四十一条：“国家保障妇女享有与男子平等的劳动权利和社会保障权利。”

与现有员工情况等数据一起输入到算法中。算法模型利用深度学习的方式自主筛选数据特性并学习其中的决策模式，并通过反复的训练和调适进而获得无限接近需求目标的数据模型，从而实现对求职者数据的分析并做出筛选。

在劳动市场中，劳动者与用人单位相比本就处于天然弱势地位，《劳动法》对于女性劳动者权益的保护采用的是特殊保护模式。算法技术在劳动领域的运用，通过大量的训练数据对劳动者进行数据分析，产生了算法性别歧视的潜在风险。这种歧视在劳动领域通常表现为对女性平等就业权、劳动权以及隐私权的侵犯。在杜萌、广东粤海丽江房地产发展有限公司平等就业权纠纷案⁴中，法院对就业性别歧视作出界定，是用人单位在招聘、工作过程中存在差别对待的行为，在晋职、晋级、考核评定、报酬、社会保险、生活福利等方面对女性劳动者存在歧视性差别。除此之外，在联合国《消除对妇女一切形式歧视公约》中也有关于性别歧视定义的厘清。⁵劳动领域的算法性别歧视会涉及多方主体，包括用人单位、平台等，客体为女性劳动者的平等就业权，内容为因算法深度学习的特性作出不合理的差别对待行为。算法技术的运用提高了用人单位人才招聘的效率，但由于在算法雇佣决策中可能嵌入了算法开发者以及雇佣决策者的价值判断和固有的就业偏见意识，造成就业歧视，影响劳动者的公平就业。综上对算法性别歧视进行定义，即是指用人单位或者平台运用算法技术，以算法的语言输入和训练数据为基础对女性作出的不合理差别对待行为，造成了女性合法权益受损的结果。

（二）与传统就业性别歧视的核心差异

算法性别歧视与传统的就业性别歧视具有显著区别，主要表现在性别歧视的成因和影响范围方面。人工智能的设计目的体现了类似人类的智力特质，例如概括推理或从过去的经历中总结，导致算法性别歧视与传统就业性别歧视存在本质差异。传统的就业性别歧视往往基于人类自身的偏见理念或刻板印象，导致女性劳动者在就业方面受到有意的不公平对待。算法性别歧视一般是由技术本身的“缺陷”引发的，其源于算法参数不健全，导致女性劳动者在就业方面受到无意的不公平对待。其次，二者的影响范围存在差异。传统就业性别歧视的影响范围相对有限，通常局限于特定的用人单位。算法性别歧视是系统性的，其影响范围更为广泛，涉及所有使用算法进行用工决策的企业和机构。

另一方面，算法性别歧视还具有隐蔽性、间接性和难以解释性的特点，进一步增加了劳动者举证的难度。传统就业性别歧视往往在招聘等环节中直接体现为显性的偏见，算法可以通过看似中立的数据和算法，间接地实施歧视行为，使得歧视更加难以察觉。“算法黑箱”造就了算法性别歧视的隐蔽性，在做出决策时的内部逻辑和计算过程对用户来说是不可见的，这意味着即使算法在决策中嵌入了就业性别歧视，被决策者和外界也很难发现到这一过程。算法性别歧视的间接性体现在，不直接使用性别作为考量因素，而是通过其他多样性的特征来表达就业性别歧视，其通常在算法生成的最终决策中表现为一种“合理”的选择。传统的就业性别歧视是显性的或可通过行为和言辞直接观察到的，这一行为在法律上容易被识别和解释，用人单位只需对自己做出的行为进行解释^[3]。“算法黑箱”的出现让劳动者难以证明决策者在运用这些算法时作出了何种价值取向的选择。劳动者若想证明其被歧视，需要获得算法的相关信息，而这些通常被视为商业秘密而不对外公开。

3. 算法性别歧视下女性平等就业权实现的法律困境

（一）算法就业歧视的认定标准模糊

当前，针对算法歧视的法律规范仍相对匮乏，法律框架也尚未健全。尽管一些国家和地区已开始重

⁴ 杜萌、广东粤海丽江房地产发展有限公司平等就业权纠纷案，广东广州市中级人民法院(2020)粤01民终219号民事判决书。

⁵ 《消除对妇女一切形式歧视公约》第一条：“为本公约的目的，‘对妇女的歧视’一词指基于性别而作的任何区别、排斥或限制，其影响或其目的均足以妨碍或否认妇女不论已婚未婚在男女平等的基础上认识、享有或行使在政治、经济、社会、文化、公民或任何其他方面的人权和基本自由。”

视这一问题,但至今尚未形成统一且明确的认定标准。这使得在实际操作中,难以依据现有法律精确判定算法就业歧视的具体情况,难以清晰界定算法歧视的范围和程度,进而增加了认定的复杂性。此外,由于算法技术的迅猛发展,法律规范的滞后性可能会使算法歧视的认定变得更加困难。

与此同时,在应对算法歧视方面,我国常通过反歧视的途径来实施治理。在现行法律框架下,平等权保护机制的有效运作依赖于对歧视事实的明确界定。然而,在现行的法律规范体系中,我国在职场性别歧视的法律界定上存在显著的制度缺失。《妇女权益保障法》第41条虽确立了就业平等权原则,却未对歧视性就业行为进行具体类型的界定。这种立法上的模糊性导致了司法裁判标准存在明显的地域差异。对于什么是就业性别歧视及其具体表现形式,目前缺乏明确规定,仅停留在学术理论界的探讨阶段。在司法实践中,法官需要根据个案情况进行权衡。但由于专业知识和个人价值偏好的影响,不同法官对这一问题的认识可能存在差异,容易导致“同案不同判”的现象。

(二) 责任归责原则尚未形成统一标准

目前,关于算法性别歧视行为归责原则的适用,理论与实务尚未形成统一标准,而且理论界仍存在较大争议。首先,将过错责任原则应用于劳动领域的算法歧视纠纷中,难以有效保障女性劳动者的权益。而且就我国目前的司法现状而言,对歧视行为的判断标准和平等就业权的界定仍相对模糊。其次,鉴于算法歧视行为中“算法黑箱”的存在,使得隐性歧视更为普遍。为有效保障女性的合法权益,应尽可能提升算法的透明度,避免“算法黑箱”带来的不透明性和隐匿性,因此主张采用无过错责任原则^[4]。另一方面,适用无过错责任原则虽然可以对平台和用人单位使用算法技术提出更高的要求,即他们需要承担更多的注意义务和风险防控义务,尽可能地防止算法歧视行为的发生。但算法技术并不都是可控的,如果将所有算法歧视归责原则都适用无过错责任原则将不利于社会发展和进步,在实践中也难以立足。

(三) 举证责任分配不合理

在就业歧视纠纷中,劳动者还面临举证不能的困境,现行的证明责任分配机制无法很好地保护劳动者。在我国反就业歧视的法律体系中,侵权责任法律框架占据着核心地位。根据《就业促进法》第62条的规定,受歧视者有权向人民法院提起民事诉讼,以求权利救济。实际上是将就业歧视归类为一般侵权行为,要求原告承担双重举证责任:首先,需证明存在不合理的差别对待行为;其次,需证明雇主方面存在主观过错。这一规范体系沿袭了《民法典》第1165条⁶所确立的过错责任原则,要求职场歧视行为同时满足四个构成要件:行为的违法性、实际存在的损害后果、主观过错的可归责性及因果关系的直接关联性。在“谁主张,谁举证”的诉讼模式下,劳动者的举证责任过重,一旦无法举证,就可能随时面临着败诉的风险,进而被法院以证据不足为由驳回全部或者部分诉讼请求。

然而,就业歧视纠纷的特殊性在于劳资双方力量对比的显著失衡。在已建立的劳动关系中,劳动者往往处于相对弱势的地位,特别是在算法背景下,雇主掌握算法决策的技术解释权,劳动者难以获取歧视性证据;其次,劳动仲裁程序要求原告承担初步举证责任,这与算法黑箱特性形成了制度冲突;最后,司法实践中对“主观过错”的认定标准存在认知分歧^[5]。现行的举证责任制度对于劳动者群体明显不利,缺乏合理性,胜诉机率十分渺茫。虽然劳动者遭受就业歧视的侵害已经是不争的事实,但是苦于举证困难只能被迫接受最终的审判结果,即使劳动者提起上诉进行二审程序甚至是最后的再审程序,最终结果也难以发生改变,大多数法院都会做出驳回上诉、维持原判的决定。

(四) 劳动者寻求救济时责任主体缺位

传统的招聘模式中,一般情况下只有求职者和招聘单位双方主体,招聘单位做出拒绝录用决定时,往往需要向求职者说明理由,招聘主体的歧视意图和歧视行为相对来说比较容易被求职者感知。此时,

⁶ 《中华人民共和国民法典》第1165条:“行为人因过错侵害他人民事权益造成损害的,应当承担侵权责任。”

求职者可以对实施就业歧视行为的招聘单位向人民法院提起诉讼，将来承担法律责任的主体也是招聘单位，责任主体相对来说比较明确、清晰。

随着算法技术越来越多地参与招聘过程，很多情况下算法的设计主体和使用主体并不是同一主体。算法程序可能是第三方算法设计机构自主开发的，可能是具有研发能力的用人单位独立设计的，还有可能是二者合作研发的。第三方算法机构设计的算法程序本身存在问题会造成歧视性后果；用人单位受经济利益驱动擅自改编算法、注入歧视性数据也会引发歧视问题；第三方算法设计机构和用人单位相互勾结，第三方算法设计机构在用人单位的要求下故意推出带有歧视性色彩的算法操作系统同样也会带来歧视风险[6]。此时，第三方算法设计机构和用人单位都有可能成为算法就业歧视的责任主体，用人单位很有可能以自己只是算法程序的使用者、算法程序产生歧视性后果的责任在于第三方算法设计机构为由进行抗辩，而真正的算法设计主体劳动者也无从知晓。另一方面算法深度学习模型的非线性决策机制可能导致歧视性后果与开发者和使用者无关，而是数据训练过程中形成的算法偏见，导致在发生算法歧视时难以准确确定归责主体。

4. 算法性别歧视下女性平等就业权实现的优化路径

（一）厘清算法性别歧视的认定标准

考虑到算法决策的技术复杂性与歧视形态的隐蔽性，我国现行法律体系在自动化决策歧视的司法裁判中面临规范供给不足的困境。为了在数字时代有效保护劳动权利，需要构建一个主客观相统一的审查标准体系，进一步明确并细化算法环境下就业性别歧视的构成要件。

在客观认定标准方面，直接歧视案件中用人单位的歧视故意可以通过明示或推定方式证明，司法认定相对清晰。对就业歧视的判定需要侧重于客观层面，并据此推断用人单位的主观态度。具体来说，需要重点审查以下四个构成要件：第一，识别差别对待行为，即判断 AI 招聘系统是否对求职者进行与岗位核心职能无实质关联的分类处理；第二，验证法定禁止性事由，包括核查算法决策是否涉及《就业促进法》第 3 条明确列举的禁止性特征；第三，量化实际损害结果，既包括显性权益损失如薪酬差异，也包括隐性发展障碍如职业路径受限等；第四，构建因果关系链，通过技术手段证明算法决策与损害结果的关联性。

在主观认定标准方面存在理论争议。算法歧视性后果可能源于三种不同情形：企业故意歧视、第三方技术瑕疵和算法内生偏见。对于后两种情形，要求企业承担主观过错责任存在合理性争议。在直接歧视案例中，歧视意图常直接显现于招聘公告或拒聘通知中，通过鲜明的歧视性语言可以确认或逻辑推断雇主的主观歧视意图，判定难度相对较低。但在间接歧视案件中，用人单位的歧视性动机往往更具隐匿性。由于算法决策的技术黑箱特性，歧视性意图可能通过数据标注偏差、特征选择偏好等隐性方式实现，导致间接歧视的判定颇为棘手。因此，为了评估歧视意图，必须结合一系列客观情境进行综合分析。

（二）明确算法性别歧视的归责原则

在算法时代，反就业歧视需遵循利益平衡和必要正当原则。为保护劳动者免受算法歧视侵害，应坚持劳动法宗旨，提供法律保障和救济途径。同时，避免过度规制，限制算法技术应用，阻碍产业发展。利益平衡要求兼顾劳动者与算法开发企业的利益，牺牲任何一方都不明智。只有实现最大利益平衡，才能达到公平可持续。这是规制算法就业歧视的必要性和正当性体现。算法就业歧视给劳动者和反歧视法律带来挑战，需采取必要规制措施，但力度需适中。规制不足或过度都非理想效果。因此，应结合我国人工智能发展现状，采取适当规制措施，坚持必要适当原则。这不仅能有效规制算法歧视，还能推动人工智能创新发展，提高国际竞争力，实现经济高质量发展。需在保障劳动者权益的同时，为人工智能提供良好政策环境，实现科技与社会和谐共进。

(三) 完善举证责任倒置制度

我国传统的就业歧视纠纷中,普遍遵循“谁主张,谁举证”的证明责任分配原则。然而,这一机制在实际操作中暴露出明显的局限性,致使遭受歧视的劳动者往往难以获得应有的法律保护,常因举证困难而面临败诉的风险。因此,为了更好地保护劳动者的平等就业权益,并增强用人单位的责任意识,我们有必要适当减轻劳动者的举证责任,合理分配劳资双方在就业歧视纠纷中的证明责任。在算法就业歧视纠纷中实行举证责任倒置,具有其合理性和可行性。算法程序由用人单位掌控,其运行机制和决策过程主要为用人单位所知。劳动者在算法技术中的参与度较低,只是按照用人单位的要求提供相关数据和信息。当劳动者怀疑自己遭受算法就业歧视时,他们能够提供的证据非常有限,而用人单位则因技术优势掌握了绝大部分相关证据。

我们应在充分考量我国就业市场现状的基础上,积极探索并逐步建立举证责任倒置的分配机制。可以考虑只要求遭受算法就业歧视的劳动者提供足以证明歧视行为可能存在以及自身遭受不公平对待的证据,达到使法院相信可能发生或已经发生算法就业歧视的程度即可。接下来的举证责任则转移至用人单位,由其提供证据证明自己的清白。这一机制不仅有助于减轻劳动者的举证负担,还能有效促进用人单位在使用算法技术时的规范和约束,从而实现劳资双方权益的更公平保护。

(四) 合理分配算法主体法律责任

在应对算法就业歧视的行为时,法律需对其责任人进行公正的审判,使其承担相应的法律责任。然而,在某些情况下,承担此法律责任的主体可能并不一致。一些用人单位具备独立开发算法雇佣决策系统的能力,若由此导致算法就业歧视,则用人单位需自行承担相关法律责任^[7]。而部分用人单位采用的是第三方机构开发的算法雇佣决策系统,当算法的开发与设计主体和实际使用主体不一致时,便需要在用人单位与第三方平台之间合理分配法律责任。

对于责任应由算法开发设计主体承担,还是由后续更改算法程序的用人单位承担,抑或是两者共同承担责任。可考虑以下情形:如果歧视源于算法开发设计主体的失误,且作为使用方的用人单位已尽到合理注意义务,则法律责任由算法开发主体独自承担,用人单位可在事后向算法开发主体追偿相关损失;若歧视由算法开发主体的失误引起,而用人单位未能尽到注意义务便投入使用,则用人单位承担主要责任;若歧视是由于用人单位擅自更改算法程序所致,则用人单位承担全部责任;若算法歧视是因用人单位与算法开发主体共谋而故意设计,则两者需承担连带责任,并按各自的过错程度进行内部责任分配。

5. 结语

在人工智能深度介入劳动市场的背景下,算法性别歧视对妇女平等就业权的侵害已成为数字时代劳动权益保护的新挑战。算法技术黑箱特性使得传统反歧视法律体系在认定标准、归责原则及举证责任等方面遭遇困境,而责任主体的多元性与算法偏见的隐蔽性则进一步加剧了妇女就业维权的难度。保护妇女平等就业权,需在技术创新与权利保障之间寻求平衡。通过完善法律认定标准、优化责任归责机制、改革举证责任分配、明晰责任主体,构建适应算法时代的反歧视法律框架,推动反就业歧视立法。同时,应推动企业建立算法伦理审查制度,加强政府监管与技术审计,形成多方共治格局。唯有如此,才能真正实现数字时代的性别平等,让算法成为赋能而非阻碍妇女发展的力量。

参考文献

- [1] 赵为. 保障高校女毕业生平等就业权——访全国人大代表张祖涛[J]. 中国人力资源社会保障, 2025(3): 33-34.
- [2] 胡萧力. 算法决策场景中就业性别歧视判定的挑战及应对[J]. 现代法学, 2023, 45(4): 59-74.
- [3] 罗熠琛, 周新衡. 算法自动化决策中就业性别歧视的类型与规范路径[J]. 中国人力资源开发, 2025, 42(1): 92-103.

-
- [4] 崔慧慧. 算法性别歧视下女性平等就业权实现的困境与路径[J]. 中国价格监管与反垄断, 2024(12): 62-64.
 - [5] 杨云霞, 谢敏. 算法引致就业性别歧视的规制研究[J]. 陕西行政学院学报, 2025, 39(2): 35-41.
 - [6] 徐素萍. 算法就业歧视法治应对的困境与出路[J]. 汕头大学学报(人文社会科学版), 2024, 40(12): 17-26+94.
 - [7] 唐芳. 对妇女的就业性别歧视界定与《妇女权益保障法》相关立法完善[J]. 中华女子学院学报, 2021, 33(6): 32-39.