

生成式 AI 视阈下数字弱势群体权益保护的法理证成 与制度构建

冯春阳¹, 绳正阳^{*2}

1 重庆大学法学院, 1104502983@qq.com

2 澳门科技大学法学院, yyszeng@163.com

摘要: 生成式人工智能技术的迭代升级不仅在悄然重塑社会运行逻辑, 同时也加剧了数字弱势群体的边缘化困境。对于数字弱势群体权益加以保护, 既是消弭数字鸿沟的现实之需, 也是社会文明转型的内在要求, 更是实现共同富裕的必要之举。保护数字弱势群体权益应秉持以人为本原则、倾斜保护原则和多元协同原则, 以人本主义遏制技术理性对人类主体性的僭越, 以倾斜保护矫正数字红利的非平等分配, 以多元协同破解算法权力分散导致的治理困局。此外, 为助益数字弱势群体权益保护制度的有效实现, 还应从顶层制度着手, 通过科学化的主体界定、体系化的内容设计以及多元化的保障机制来丰盈数字弱势群体权益保护的灵魂, 进而在法治化轨道上实现生成式人工智能技术高质量发展与数字弱势群体高水平保护之间的良性互动。

关键词: 生成式 AI, 数字弱势群体, 权益保护

1 引言

党的二十届三中全会强调, 要“加强关键共性技术、前沿引领技术、现代工程技术、颠覆性技术创新^[1]”。从早期的 Sora 与 o1 大模型到后来的 ChatGPT 与 Deepseek 的兴起, 以人工智能 (Artificial Intelligence, 以下简称 AI)、大数据、云计算和智能机器人等技术为核心的数字社会雏形已经形成^[2]。然而, 正如习近平总书记所指出: “古往今来, 很多技术都是‘双刃剑’, 一方面可以造福社会, 造福人民, 另一方面也可以被一些人用来损害社会公共利益和民众利益^[1]。”生成式 AI 在推动社会运行方式深刻变革的同时, 也在不断重塑权利分配结构与社会分层格局。换言之, 数字红利的普惠性并未均衡释放, 反而在一定程度上加剧了特定群体的边缘化处境, “它像一道无形的屏障隔离、抛弃乃至淘汰新技术的落伍者”^[15]。由此, 数字弱势群体问题逐渐凸显, 并成为数字法治体系中不可回避的核心议题。

所谓数字弱势群体, 是指因经济、技术、社会地位及学习能力差异, 叠加信息社会分化与数字技术不均衡传导等结构性障碍, 导致在数据信息获取、理解和应用层面处于劣势, 进而导致资源匮乏、能力缺失、社会边缘化及正当权益受损的特定群体^[17]。从发生学角度审视, 数字弱势群体是既有社会弱势群体在数字化进程中的延续性, 与数字技术发展所催生的新型弱势形态交互作用的结果。前者体现为传统不平等在数字空间的再生产与强化, 后者则源于数据与算法所构建的技术权力结构对主体能力的重新分层。在生成式 AI 背景下, 人类弱势形态呈现出叠加与放大效应, 从而进一步加剧其结构性不利地位。围绕数字弱势群体问题, 现有研究已从多个维度展开:

作者简介: 冯春阳, 男, 重庆大学法学院博士研究生。

绳正阳, 男, 澳门科技大学法学院博士研究生。

本课题得到[2025年度中央高校基本科研业务费项目“生命共同体理念下自然资源一体化治理研究”(2025CDJSKPT11)]资助。

*通讯作者

有学者直接聚焦该群体的概念界定、权利内容及国家保护义务^[16]；有学者在研究数据人权时将数字弱势群体权益保障作为重要组成部分^[34]；还有学者在研究其他相关问题时间接指明数字弱势群体的利益诉求^[35]。上述研究为理解数字弱势群体的权利结构及其制度回应提供了重要基础，但总体上仍呈现出碎片化特征，缺乏面向生成式 AI 情境的法理证成与制度建构。有鉴于此，本文首先从理论层面证成数字弱势群体权益保护嵌入生成式 AI 法律规制体系的必要性；其次从价值理念层面阐明生成式 AI 对数字弱势群体权益保护应当秉持的基本原则；最后在顶层设计层面提出具有统筹性、整体性的数字弱势群体权益保护制度构想，避免数字弱势群体权益保护流于主观恣态。

2 生成式 AI 背景下数字弱势群体保护之必要性省察

社会的发展塑造了技术，但同时也被技术所塑造^[44]。在数字技术迅速发展的背景下，数字弱势群体的权益保护问题，不能简单归结为技术竞争中的优胜劣汰，而应理解为 AI 推动社会结构调整过程中所产生的结构性矛盾。在以技术研发者与平台主体为主导的治理框架中，个体参与能力相对不足，由此形成技术参与和社会参与之间的结构性失衡，并具体表现为权利实现机会的不均等，即数字不平等^[27]。在此背景下，弥合生成式 AI 与数字弱势群体之间的能力差距、强化相关主体的权利保障，已成为数字人权保护与社会可持续发展的重要议题。与此同时，对数字弱势群体的保护仍需满足利益保护的正当性、与既有法律体系的协调性以及制度实现的可行性等要求。因此，有必要从现实基础、社会基础与政策基础三个层面，对其加以系统证成。

2.1 现实维度：弥合数字鸿沟的必要之举

相较于传统互联网技术主要围绕信息接入与传播不平等展开，生成式 AI 通过重构知识生产机制与人机交互方式，使数字不平等从“接入差异”进一步演化为“认知能力差异”，从而在更深层次上重塑了数字弱势群体的生成逻辑。

一方面，生成式 AI 在重塑社会生产方式与知识获取路径的同时，也加速了数字弱势群体权利受损的风险，使数字鸿沟由“接入差异”转向“能力差异”。传统数字鸿沟主要体现为基础设施与基本技能的不均衡，而生成式 AI 的发展通过提升算法复杂度、强化数据依赖并加快技术迭代，将不平等由物理接入层延伸至认知与应用层。具体而言，技术优势主体借助生成式工具显著提升知识生产与决策效率，而数字弱势群体因数字素养不足与资源受限，难以有效理解与使用相关技术，并在算法机制影响下处于不利地位。此种差异不仅表现为技术使用能力的不足，还可能因推荐机制的“信息茧房”而进一步加剧认知偏差，形成能力弱化与信息受限的叠加效应。从结构层面看，生成式 AI 的发展反映出智能时代权力配置的不均衡：一方面，模型训练所依赖的数据可能内嵌既有社会偏好；另一方面，算力与技术资源的分布具有明显的经济与区域差异。在此背景下，生成式 AI 可能放大既有不平等，使数字弱势群体在技术应用、信息获取与利益表达方面处于不利地位，其相关诉求亦难以进入有效的制度回应范围，由此引发新的数字歧视风险^[13]。

另一方面，与传统互联网基于关键词检索的交互模式不同，生成式 AI 依托自然语言处理 (Natural Language Processing, NLP) 与 Transformer 架构构建的以对话为核心的人机交互模式。该模式对用户提出更高的认知要求，具体表现为需要具备语义理解、上下文把握及多轮交互能力，从而使人机交互由操作性活动转向认知性活动。相关统计数据显示，生成式 AI 用户中 50-59 岁群体占比仅 10.7%，而大专以下学历的网民对于生成式 AI 的使用率低至 12.6%^[7]。这表明，不同年龄与教育背景群体在技术使用能力上的差距，正在通过交互门槛被进一步放大：其一，交互复杂性的提升使原本具有普惠属性的技术在实际运行中转化为筛选机制，部分用户因认知能力不足而被排除在有效使用范围之外；其二，算法在持续优化过程中对高频使用群体形成偏好，进而影响模型训练与服务供给，使数字弱势群体的需求难以得到充分体现。

在此背景下，交互能力差异逐渐转化为社会参与能力差异，并通过能力与机会的循环机制不

断强化：高能力群体通过持续使用积累技术优势，而弱势群体则因使用障碍进一步边缘化，进而加剧数字不平等^[39]。因此，对数字弱势群体权益的保护，不仅是对技术发展所引发不平等后果的回应，也是实现生成式人工智能普惠目标的必要条件。

2.2 社会维度：社会文明发展的内在要求

数字弱势群体权益保护除了要符合现实层面的真实性、重要性和紧迫性，还需要从社会文明维度进行考量。生成式 AI 的广泛应用，正在推动生产方式与生活方式发生结构性变化，并加速社会形态向数字化转型^[19]。在此文明转型的临界点上，将数字弱势群体权益保护纳入制度体系，不仅是回应技术发展带来不平等风险的必要举措，也是维持社会运行稳定性与制度正当性的基本要求。

生成式 AI 技术背后所隐藏的“算法利维坦^[25]”正在对传统文明赖以维系的道德主体性基础产生颠覆性冲击。在此过程中，数字弱势群体的文化表达与价值诉求，可能因数据采集与模型训练机制的偏向而被弱化，甚至难以进入主流信息体系。在人类文明演进的历史坐标系中，社会对弱势群体的保护程度始终构成衡量社会文明高度的核心指标。从雅典城邦对异邦人权利的逐步承认，到启蒙运动对“天赋人权”的价值确认，再到现代国家对社会经济权利的制度化保障，其核心均在于通过制度安排防止弱势群体被排除在公共生活之外。然而，在数字技术澎湃发展的今日，生成式 AI 非但未能凸显数字弱势群体在社会文明中的主体地位，反而加剧其边缘化^[23]。当大语言模型通过语义生成重构公共对话模式时，“沟通理性”所需的主体间性对称关系已被算法的不对称结构彻底瓦解^[4]。生成式 AI 通过海量数据训练形成的“算法霸权”，正在将边缘群体的叙事方式与认知模式排除在主流语义网络之外，其影响程度并不亚于传统的物质性不平等。例如，当老年群体的交互需求未被纳入算法设计与优化过程，其在数字空间中的使用体验与参与能力将受到限制，从而在事实上处于不利地位。从价值层面看，以技术机制为中介的能力差异扩展，将威胁到“人是目的”的文明根基^[31]，即人类刚从自我招致的不成熟状态中解放而出，却又正在面临被技术权力重新奴役的可能^[36]。

此外，数字社会的形成以数据、算法与算力为关键要素，并由此重塑资源配置与社会运行方式，即以数据为核心要素、算法为组织原则、算力为权力基础。在此背景下，若技术发展缺乏对多样性与包容性的制度回应，数字弱势群体所承载的文明多样性极有可能在数据处理中被清洗，从而影响其在数字空间中的可见性与延续性，最终导致人类陷入哈贝马斯所警示的“系统对生活世界的殖民”困境^[4]。为此，唯有将数字弱势群体权益保护锚定于社会文明演进轨道之中，才能彰显社会文明“和而不同”的精神内核。

2.3 政策维度：实现共同富裕的应有之义

伴随社会转型，社会发展对个体发展权与人格尊严的保障功能不断强化，并逐步体现为通过制度安排实现更为普遍的社会福利^[40]。从政策层面来讲，生成式 AI 背景下对数字弱势群体权益的保护，已成为实现共同富裕目标的重要内容。若经济发展以弱势群体发展机会的受限为代价，将有悖于公平分配的基本要求。

从时空维度考察，生成式 AI 发展与数字弱势群体权益保护在实现共同富裕目标中具有结构性关联。从时间维度看，生成式 AI 的技术演进与我国共同富裕目标的阶段性推进具有同步性。一方面，在“十四五”规划及 2035 年远景目标框架下，数字经济被确立为推动现代化发展的重要动力^[6]，生成式 AI 作为关键技术形态，其发展路径与数字化转型的整体部署相契合；另一方面，共同富裕强调分阶段推进社会公平与发展成果共享，这与生成式 AI 逐步扩展应用、释放技术红利的过程在节奏上具有一致性。因此，在技术扩散过程中同步强化对数字弱势群体的保护，具有现实紧迫性。从空间维度看，共同富裕所关注的发展不平衡与不充分问题，在数字社会中表现为不同区域、群体之间的技术获取与使用能力差异。生成式 AI 的应用，可能在不同地区和群

体之间产生差异化影响。因此,将数字弱势群体权益保护纳入技术治理与政策设计之中,有助于在数字化转型过程中缩小差距,推动发展成果的均衡分配。

从数字技术赋权与平权的角度来看,技术赋权以人与技术的深度融合为基础,通过能力补充提升个体参与数字社会的可能性。生成式AI通过自然语言交互、多模态感知等技术特性,为数字弱势群体提供了突破传统交流壁垒的工具,例如,残障群体可借助语音合成与识别技术实现更为便利的信息获取与沟通。然而,赋权主要体现在使用机会层面,尚未必然转化为实质意义上的平等。从规范内涵看,“平权”中的“平”既是宪法当中平等的平,也是公平的平,蕴含从“我”到“我们”的伦理追求^[24]。因此,在共同富裕目标下,不仅应关注数字弱势群体使用生成式AI的覆盖范围,还应重视其使用效果与实际收益。进一步而言,生成式AI的运行依赖数据与算力资源,其分配方式在一定程度上影响技术收益的分布。如果缺乏相应的制度调节,可能导致技术可及性与实际受益之间出现差距,使部分群体在能力与资源获取方面处于不利地位。因此,在多场景应用中强化对数字弱势群体的制度保障,不仅有助于改善其技术参与条件,也契合通过制度安排提升不利地位群体利益的基本正义要求。

3 生成式AI背景下数字弱势群体保护之基本原则

韦伯曾忧虑人们沉迷于技术力量带来的支配感中,没有适切的价值指引,科技的力量让人看到了支配现实世界的可能,却忘记了支配的方法^[18]。事实证明,生成式AI的发展在一定程度上加剧了不同群体之间的能力差异,使既有的权利结构面临调整压力,并呈现出由公权力、平台私权力与个体权利共同作用的复合结构^[33]。基于数字弱势群体在资源获取与能力结构上的不利地位,在推进生成式AI价值对齐与制度设计时,有必要充分考虑技术机制与权利保护之间的内在关联。由此,数字弱势群体权益保护不应仅停留于理论层面,而应通过明确的基本原则加以规范与支撑。

3.1 价值回归:以人为基本原则

生成式AI的技术演进,本质上映射出社会生产力与生产方式迭代更新的历史逻辑,其发展不仅契合AI的科技本质属性^[8],亦再次印证了“生产工具革命性突破驱动社会形态跃迁”的一般规律。然而,技术进步并不必然导向价值正当性,正如有学者所指出,“维护人的价值与尊严始终是现代法治的主导精神和优良传统^[3]。”据此,在生成式AI语境下,数字弱势群体权益保护的规范基础,应当回归“以人为本”的价值立场,将个体权利与人格尊严作为制度设计与技术治理的基本出发点。

一方面,技术理性以数字结构解释世界,将技术实用价值与应当遵循的伦理道德严格区分开来。在此框架下,人类所具备的独特性被理解为一种特定的生物结构,是一个进化意义上的短暂事件^[42]。世界人权宣言以及国际人权公约皆宣称人权源于人固有的尊严,在AI时代,若过度强调技术效率,可能导致技术理性对价值判断产生过度影响,从而对以人为中心的规范立场形成冲击^[9]。在此过程中,人的主体性逐步被消解,即从技术改善人类生活,到技术重新设计人类,再到可能迈向人类的灭绝^[43]。进一步而言,在生成式AI语境下,算法系统在决策与内容生成过程中逐渐呈现出“准主体化”特征,即技术不再仅作为工具存在,而是在一定程度上参与价值判断与意义建构。这一变化直接动摇了“人作为价值主体”的规范前提,使得传统以人类为中心的权利体系面临内在冲击。基于此,有必要在生成式AI发展过程中强化“以人为本”的规范立场,通过明确人的价值优先性,界定技术应用的边界,防止技术机制对主体地位产生不当影响,从而在技术与权利保障之间形成必要的规范约束。

另一方面,数字技术的运行往往嵌入既有利益结构,难以保持完全中立,其发展与应用在一定程度上更易向资源优势群体倾斜,并与性别、阶层、年龄等因素相互关联。然而,并不能据此否认数字弱势群体作为权利主体的地位,亦不影响其依法享有数字人权。原则上,所有个体均应

成为生成式 AI 技术发展的受益者。从政策层面看,《中国关于全球数字治理有关问题的立场》明确提出各国应致力于提升不同群体的数字素养与技能水平,尤其应加强对老年人、妇女、未成年人、残疾人及贫困人口等社会弱势群体的数字技能培训,以切实维护并促进其发展权的实现。这表明,数字弱势群体权益保护已具有明确的制度导向。数字弱势群体具有动态性特征,个体的技术能力与适应能力可能随年龄增长或技术迭代而发生变化,从而使不同群体在不同阶段处于相对不利地位。换言之,数字弱势并非固定身份,而是一种可能普遍发生的结构性状态。在此背景下,若忽视相关群体的利益诉求,可能导致既有社会不平等在数字空间中的延续与强化。因此,应当在生成式 AI 的发展与治理中,坚持以人为本的价值立场,恪守“人是地球上其他一切自然物都与之相关的那个自然最后目的^[37]”的基本前提。

3.2 实质平衡: 倾斜保护原则

倾斜保护原则以承认个体在资源、能力及社会地位等方面的差异为前提,强调对处于不利地位的主体给予制度性保护^[29],其理论内涵主要体现在两个方面:一是基于社会正义的要求,通过差异化的权利义务配置强化对弱势群体的保障;二是通过制度调节矫正利益分配失衡,以实现实质意义上的平等。

一方面,从数字正义的角度来讲,“正义具有一张普罗透斯的面^[38]”,无论是结合 AI 特性重构正义内涵,还是将一般正义原则适用于生成式 AI 治理,均应坚持在保障多数人利益的同时,兼顾对少数群体的保护。数字弱势群体的形成,在一定程度上源于技术发展过程中对效率与工具理性的偏重,其结果是部分群体在资源获取与技术适配方面处于不利地位。在数字社会中,生成式 AI 通过标准化与自动化机制,有助于提升规则适用的一致性,从而在形式层面促进公平,但其并不能当然实现实质正义。具体而言,数字弱势群体的权利诉求往往会被有意识或无意识地屏蔽,譬如,生成式 AI 产品的提供者在面对不同用户之间存在能力差异时,其往往以自身利益为导向,通过调控数字技术应用的包容程度,将一小部分人群排斥在应用之外。即使这些数字弱势群体符合基础门槛和准入要求,但生成式 AI 在自动化决策仍可能基于统一标准作出判断,从而忽视个体差异,使该群体在权利实现与救济能力方面处于弱势。“社会公平正义作为一种价值准则,成为合理地抑强扶弱的法律尺度^[28]。”因此,有必要在生成式 AI 治理中引入倾斜保护原则,通过差异化的权利配置与制度安排,合理分配数字权利与义务,以及技术发展所带来的利益与负担,以弥合技术接入与使用中的实质差距,促进数字弱势群体在参与机会与结果层面的实质正义。

另一方面,从数字平等的角度看,生成式 AI 的引入与应用并非在均等条件下展开。资本投入差异、技术门槛以及数字马太效应,使“智能鸿沟”在数字社会中进一步扩展,呈现出新的不平等形态。从制度逻辑看,平等是调整社会利益关系的重要原则。“迄今为止,所有改革型社会的进程都是一场从身份到契约的运动^[32]。”传统契约关系主要以行为能力为判断基础,对当事人的知识结构与能力差异关注有限,从而在形式上实现主体之间的平等。然而,在数字技术服务情境中,用户的数字素养与技术使用能力直接影响其权利实现效果,使原本形式平等的契约关系在实质上呈现不对等状态。因此,单纯依赖形式平等,难以回应数字时代主体之间的能力差异,甚至可能导致对弱势群体的“拟制平等”。在此背景下,应当将形式平等与实质平等相结合,通过制度安排关注权利实现的实际效果。对数字弱势群体实施倾斜保护,有助于弥合能力差距,使其摆脱仅具形式意义的平等状态。此外,在生成式 AI 背景下,倾斜保护原则亦面临“家长主义”风险,即过度干预可能以保护之名限制数字弱势群体的自主选择空间,甚至强化其依附地位。因此,倾斜保护原则亦应转向以能力提升与条件改善为导向的“赋权式保护”,在保障基本权益的同时,促进其在数字社会中的自主参与能力。

3.3 协同共治: 多元协同原则

党的十九大以来,党中央明确提出“打造共建共治共享的社会治理格局”,其不仅构成数字

社会治理的重要原则,也应成为数字弱势群体权益保护的基本指导。生成式AI的发展在重塑社会运行方式的同时,也带来了新的治理问题,对既有制度框架提出了更高要求。数字弱势群体的权益保护绝非单一主体所能承载,其内在复杂性要求保护框架必须突破单向维度或市场自发调节的窠臼,转而寻求多元主体的动态协同。

从认识论层面审视,数字弱势群体本质上是一个“未完成的概念”。生成式AI的持续迭代,使“弱势”的边界与内涵不断调整,个体在不同阶段可能因技术适应能力变化而处于不同位置。主体的流动性表明,基于静态分类的保护模式难以有效回应现实需求,有必要构建兼具技术嵌入性与社会适应性的协同保护框架。从权力配置模式来看,数字弱势群体权益失衡与权力分散密切相关。AI技术将“广泛性权力和深入性权力,权威性权力和弥散性权力以前所未有的方式结合起来^[26]”,以往国家单极权力格局转变为技术研发者、平台服务提供者和国家的三元分立结构。这就决定了“没有任何单一主体具有能够单独实现目标的知识 and 资源^[12]”。因此,数字弱势群体权益保护的理想状态应为技术开发者将伦理考量嵌入算法设计的底层逻辑,平台运营者在商业利益与公共利益间构建平衡支点,政府以制度供给确立规则边界。

然而,权力的分散并未给数字弱势群体权益提供应有的保护,反而使得数字弱势群体权益受损呈现出复合性特征。一方面,技术研发主体掌握算法设计与数据筛选权,在效率优先的导向下,其技术框架往往以主流用户特征为参照,导致数字弱势群体的需求难以进入模型设定与规则生成过程,从而在技术表达体系中处于边缘地位。另一方面,平台服务提供者掌握用户入口与运行机制,其决策逻辑通常以效率与收益为导向,在产品设计及准入规则设定中,可能对弱势群体的使用需求关注不足,进而影响其实际可及性与受益程度。同时,国家虽具备制度供给与规范制定能力,但在识别生成式AI所带来的隐蔽性风险及其具体影响方面,仍面临信息与能力约束,导致在相关权益保护中存在一定滞后性。上述因素共同作用,使数字弱势群体权益受损呈现出复合性与隐蔽性特征。由此,单一主体主导的治理模式难以有效应对相关问题,有必要通过制度设计明确各主体责任,构建技术开发者、平台主体、社会组织与政府之间的协同治理机制,以提升数字弱势群体权益保护的整体效果。

4 生成式AI背景下数字弱势群体权益保护之制度设计

任何一项新兴权益的研究都不能只是进行理论或者价值上的规范求证,还应当回归到法学研究的理论自觉^[30]。在生成式AI发展的背景下,数字弱势群体权益保护有必要转向法律层面的制度化表达,使相关理念转化为可操作的规范体系。总体而言,《中华人民共和国个人信息保护法》《中华人民共和国数据安全法》《中华人民共和国未成年人保护法》《中华人民共和国老年人权益保障法》等规范,已从个人信息保护、数据安全与特定群体保护等维度,对数字弱势群体权益提供了一定程度的间接保障。然而,上述制度多以一般性权利保护或特定主体保护为导向,“数字弱势群体”问题尚缺乏体系化回应。基于此,下文将从主体范围、权利内容与保障机制三个方面展开,旨在对现有法律框架进行结构性补充与完善,以弥补其在新技术情境下的规范不足。

4.1 数字弱势群体的范畴识别标准

在生成式AI语境下,数字弱势群体的界定不宜依赖封闭式主体列举,而应转向基于“能力缺乏”与“社会排斥”的动态识别标准。具体而言,当主体在数据获取、算法理解或技术应用过程中,因能力缺失或制度性排斥而处于结构性不利地位时,即可被纳入数字弱势群体的范畴。

一方面,从自然人维度审视,上述结构性弱势通常表现为认知能力、交互能力与资源获取能力的不足,其典型形态包括老年群体在智能设备操作中的适应障碍、残障群体在信息获取中的可及性受限,以及偏远地区居民与低收入群体在数字基础设施与数字服务获取上的现实约束。上述问题在生成式AI背景下呈现出进一步强化的趋势:其一,算法推荐与用户画像机制可能在信息分配过程中固化既有差异,从而在数字空间中放大不平等;其二,数据的大规模采集与利用,使

个体在不完全知情的情况下持续参与数据供给,其在信息生产与利用过程中的主体地位相对弱化。因此,自然人的数字弱势不仅表现为技术使用层面的不利地位,还涉及其在数字环境中的权利实现与主体地位问题。在此意义上,将老年人、残障群体等纳入数字弱势群体范围,具有明确的现实基础与规范正当性。

另一方面,自然人与法人的数字弱势具有同构性本质。二者虽在法律形态上存在差异,但其弱势地位均源于对“数据、算法、算力”这一新型生产资料的控制缺失:自然人表现为数据被动供给与算法规训下的主体性弱化,法人则体现为数据依附、算法受制与算力受限所导致的发展空间受压。由此,自然人与法人虽处于不同制度位置,却在数字生产关系中共同处于被支配地位,从而呈现出一致的结构性的弱势。在法人维度,这种结构性不利地位集中体现为企业在平台准入、数据获取与算法依赖中的系统性劣势。以中小微企业为例,其弱势生成机制可从三个层面加以把握:其一,在技术接入层面,可能面临平台准入条件与数据获取渠道的限制;其二,在能力发展层面,受技术迭代与算法不透明性的影响,存在认知与适应能力不足的问题;其三,在权利保障层面,平台规则制定与算法决策机制中,企业的话语权与救济渠道相对有限。在此背景下,大型平台通过数据与算法资源的集中,形成较强的结构性优势,中小微企业在平台生态中往往呈现出依赖性特征。此外,数据资源的集中在提供技术支持的同时,也可能加深中小微企业对平台的依赖,使其在数字经济价值链中的地位相对固化。因此,在生成式 AI 背景下,将部分处于不利地位的法人主体纳入数字弱势群体的分析框架,具有一定的现实基础与制度意义。

4.2 数字弱势群体权益的主要内容

数字弱势群体权益并非单一权利形态,而是由多项具体权利构成的权利集合。因此,有必要在区分自然人与法人两类主体的基础上,对相关权利内容加以具体化与类型化,以明确其规范结构与实现路径。

4.2.1 自然人维度

在自然人维度,应重点关注数字弱势群体作为公民所享有的社会保障与社会福利等具体利益。在此基础上,可将其相关权益概括为以下三种类型:

第一,数字平等权。任何个体都应得到公正对待,这是我国宪法平等权规范的基本意涵。在生成式 AI 技术快速渗透社会各领域的背景下,相较于一般主体,数字弱势群体在该权利的实现上面临更为严峻的结构性障碍:其一,能力差异使其难以有效识别与对抗算法偏见;其二,数据贫困导致其在训练数据中的表征不足,进一步加剧算法对其不利的预测结果;其三,在遭遇数字歧视时难以通过法律或技术路径实现纠正。对此,阿玛蒂亚·森提出的“权利贫困理论”在此语境下获得延伸:当生成式 AI 重塑信息生产与传播方式时,权利贫困的症结已从单纯的分配不均转向为具体权利内容的缺失^[20]。当算法模型通过数据偏好固化社会偏见时,法律应要求技术设计者主动识别并消除训练数据中的结构性偏差,确保数字服务不因年龄、地域、教育水平等差异而区别对待。因此,有必要通过算法纠偏义务、差异化界面设计以及强化平台责任等方式,实现面向能力差异的实质平等保障。

第二,数字自由权。生成式 AI 通过数据训练与模型建构,对个体行为与偏好进行预测与引导,若缺乏必要约束,可能对个体选择空间与意志形成影响。数字自由权旨在保障个体对自身数据及其使用过程的控制能力,但对数字弱势群体而言,该权利在现实中往往受到能力与技术结构的限制,难以充分实现。基于此逻辑,数字自由权可从四个层面加以界定:一是获取层面,即个体享有不受不当限制地接入信息与数据资源的权利;二是利用层面,即能够有效理解并使用数据与相关技术服务;三是处分层面,即对涉及自身的数据享有更正、删除及限制使用的权利;四是认知层面,即在信息获取与交流过程中,避免因算法机制导致的信息单一化,从而保障认知的多样性与自主性。基于此,有必要通过降低技术使用门槛、完善无障碍设计以及引入必要的辅助机

制等方式,促进数字弱势群体在上述各层面的权利实现,从而实现数字自由的实质保障。

第三,数字安全权。传统个人信息安全主要关注数据的非法获取与不当使用,相关法律规范如《中华人民共和国电信条例》《互联网信息服务管理办法》《中华人民共和国网络安全法》等已从多个层面对信息安全提供基础性保护。但在生成式AI技术广泛应用的当下,数字弱势群体因风险识别与防范能力不足,更易成为数据滥用与数字欺诈的受害者,其安全风险具有更高的集中性与脆弱性。因此,数字安全权的内涵有必要在传统信息安全基础上加以扩展,主要包括两个方面:一是数据隐私保护。数字弱势群体在日常生活中需要频繁向平台提供个人信息,其在使用各类数字服务过程中所形成的数据轨迹,可能被持续收集与分析,从而增加隐私暴露风险^[1];二是数据滥用与伪造风险。相关主体在对数据进行处理与再利用时,若缺乏有效约束,可能通过数据分析与内容生成实施针对性误导或欺诈,加剧对弱势群体的损害。基于此,有必要通过制度安排强化对数字弱势群体的安全保障,包括完善平台的默认保护机制、提升风险识别与预警能力,并建立面向弱势群体的专门救济渠道,以实现其数字安全权的有效保护。

4.2.2 法人维度

法人的数字弱势地位,本质上源于技术资源集中对市场结构的影响,其权利体系有必要突破传统以财产权为中心的分析框架,转而关注在技术权力结构中的生存权与发展权。具体可从以下方面加以展开:

第一,数字资源公平获取权。该权利要求在符合法律与安全边界的前提下,实现数据与相关数字资源的合理开放与公平接入,防止平台主体通过数据壁垒形成排他性竞争结构。对中小微企业而言,其在数据获取与平台接入中往往面临“形式开放而实质受限”的问题,进而影响其市场参与和发展能力。据此,有必要在制度层面确立公共数据合理开放与非排他利用的基本原则,通过规范数据共享范围与条件,推动数据接口与技术标准的统一,降低中小微企业接入数字基础设施的制度性成本。同时,算力资源作为关键生产要素,其配置亦应纳入公平获取的规范框架,通过反垄断与竞争法规制相关主体滥用市场优势地位的行为,防止算力供给与定价机制对中小企业形成不当限制。

第二,数字技术自主发展权。该权利旨在保障企业在数字环境中的经营自主性,避免因算法不透明与技术依赖而受到不当约束。一是算法知情权,该权利要求技术提供者在合理范围内说明算法决策的基本逻辑与主要影响因素,以防止企业在平台规则与算法机制的双重作用下处于不利地位。算法透明应采取分层实现路径:在用户层面,通过简明说明提升对算法运行与风险的基本认知;在监管层面,由监管机构或授权第三方在保密条件下对算法模型及相关数据进行审查;在社会层面,通过发布算法影响评估等方式实现结果性透明,在保护商业秘密的前提下强化外部监督。二是技术选择与适配权,其保障企业根据自身发展需要选择和使用数字工具的自由。平台不得通过捆绑服务、技术限制或不兼容安排,变相压缩企业选择空间。企业有权基于自身能力与发展规划决定技术应用路径,而不应被动接受平台设定的技术升级或运行规则。

第三,数字规则制定参与权,该权利旨在保障企业在平台治理与算法规则形成过程中的程序性参与地位。当前,算法治理与平台规则的制定权高度集中于技术资本手中,企业(尤其是中小微企业)的利益诉求往往被排除在决策程序之外。据此,有必要在制度层面引入协商与参与机制,在涉及算法排序、流量分配等关键规则制定时,保障企业代表的参与权,使其合理诉求能够进入规则形成过程。同时,在国家标准与行业规范制定中,应当赋予企业平等参与的机会,防止通过技术标准与专利安排形成不当竞争优势。此外,参与权的实现不应仅停留于形式层面,还应通过影响评估与反馈机制,确保相关主体的意见能够在规则设计中得到实质性回应,从而提升规则制定的公平性与包容性。

4.3 数字弱势群体权益的保障机制

数字弱势群体权益的实现，有赖于相应保障机制的支撑。在生成式 AI 背景下，单一治理路径难以有效应对相关问题，有必要通过制度设计构建多方协同的保障机制，以提升权益保护的實際效果。

首先，技术研发者与平台服务提供者应通过责任内化助推数字弱势群体权益实现。生成式 AI 的技术架构不仅涉及技术选择，也在一定程度上影响利益分配与价值取向。因此，技术主体不宜以技术中立为由回避责任，而应在技术生命周期各阶段嵌入相应的伦理与合规要求。从域外实践看，欧盟《人工智能法案》第 12 条第 1 款明确要求高风险人工智能应具备运行时自动记录事件的能力，第 2 款则强调 AI 开发者将 AI 产品投入市场之后应当承担产品追踪记录义务^[41]。基于此，技术研发者可在模型开发阶段引入风险评估与测试机制，并履行相应的记录与报告义务，以提升系统的可追溯性与安全性。同时，平台服务提供者应在运营规则中纳入对数字弱势群体的保护要求，在用户协议与数据处理规则中以明确方式提示相关权利义务，并优化产品设计以降低使用门槛，例如在用户协议之中对于涉及个人数据的相关条款使用显著标识。此外，无论是技术研发者还是平台服务提供者均应当承认生成式 AI 公共物品的属性，允许第三方机构对算法模型进行伦理审查。例如，当数字弱势群体对算法决策存有异议时，可引入在保密义务约束下的专业审查机制，由具备资质的独立机构开展算法合规与伦理评估，在保护商业秘密的前提下实现外部监督^[10]。

其次，司法机关通过倾向性个案裁判和出台相应司法解释提供价值指引作用。“在裁判规则未明的案件中，法官通常通过解释扩大规则的适用范围，将既有规则拓展到新兴技术领域^[21]。”可见，司法的功能不只是定分止争，当出现新型疑难案件或典型案件时，发挥判决的价值指引作用是司法推动数字弱势群体权益实现的重要方式之一。面对生成式 AI 引发的侵权和不等保护，借助典型案例的标杆效应对数字弱势群体权益的平等保护以及倾斜性裁判，能通过示范效应提升规则适用的统一性。然而，典型案例的适用范围具有一定局限，其规范效力仍依赖个案情境。因此，有必要通过司法解释将个案中形成的裁判规则加以抽象与整合，转化为具有普遍适用性的规范标准，将新兴法益的权利属性和伦理价值诉求真正变为可予以司法救济的法治诉求^[22]。

最后，国家应通过立法与制度供给，为数字弱势群体权益提供基础性保障。作为公共利益的维护主体，国家有必要在法律层面明确相关权益的地位，并将以人为本的价值原则具体化为可操作的规范。虽然《人工智能法（专家建议稿）》明确规定了数字弱势群体权益应当受到法律保护，但在配套机制与实施路径方面仍有完善空间。因此，应进一步健全相关制度体系，例如完善监管机制，探索设立专门协调机构，以提升治理的系统性与专业性。此外，为避免相关权益停留于形式层面，可通过赋予社会组织一定的程序参与权，推动其参与监督与权益保护过程，从而形成多元主体协同的治理格局。相应地，还应完善投诉与举报机制，畅通权利救济渠道，以强化制度运行的实际效果。

5 结语

生成式 AI 正深度嵌入社会运行，对个体认知方式与权利实现结构产生持续影响。在此过程中，数字弱势群体的权益问题已超出传统权利救济框架，成为衡量数字社会包容性的重要指标。然而，当生成式 AI 以“普惠”之名分配数字资源时，技术的非对称性扩散却加剧了数字鸿沟的纵向深度。上述问题表明，生成式 AI 的发展不仅涉及技术进步，还关系到权利结构与价值秩序的调整。因此，有必要在技术治理中重新确立以人为本的基本立场，将个体权利与人格尊严作为制度设计的重要依据。应通过以人为本原则对技术应用设置必要边界，防止技术机制对主体地位产生不当影响；通过倾斜保护弥合能力与资源差异带来的不平等；通过多元主体协同，明确技术研发者、平台主体、社会组织与国家在权益保护中的责任分工。在此基础上，方能防止数字不平等进一步转化为权利实现的不平等，进而在生成式 AI 背景下保障数字弱势群体的基本权利与发

展空间。

参考文献:

- [1] 习近平.论党的宣传思想工作[M].北京:中央文献出版社:2020:202.
- [2] 马长山.迈向数字社会的法律[M].北京:法律出版社:2021:42.
- [3] 马长山.数字公民的身份确认及权利保障[J].法学研究,2023,(4):21-39.
- [4] 王恒.权利正当性后形而上学重建的困境——评哈贝马斯《在事实与规范之间》[J].法制与社会发展,2006,(3):145-155.
- [5] 中共中央.中共中央关于进一步全面深化改革 推进中国式现代化的决定[Z].2024-7-21.
- [6] 中华人民共和国国务院.中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要[Z].2021-3-13.
- [7] 中国互联网络信息中心.第 55 次中国互联网络发展状况统计报告[EB/OL].[2026-03-10].
- [8] 龙卫球.科技法迭代视角下的 AI 立法[J].法商研究,2020,(1):57-72.
- [9] 龙柯宇.生成式 AI 应用失范的法律规制研究——以 ChatGPT 和社交机器人为视角[J].东方法学,2023,(4):44-55.
- [10] 刘乃梁,叶照琪.数字平台成瘾性技术的风险规制[J].电子政务,2025,(3):1-15.
- [11] 刘艳红.公共空间运用大规模监控的法理逻辑及限度——基于个人信息有序共享之视角[J].法学论坛,2020,(2):5-16.
- [12] 李汉卿.协同治理理论探析[J].理论月刊,2014,(1):138-142.
- [13] 李成.AI 歧视的法律治理[J].中国法学,2021,(2):127-147.
- [14] 李钧鹏,王东猛.社会关系中的主体叙述[J].江汉论坛,2024,(2):127-135.
- [15] 杨嵘均.“技术索权”视角下信息弱势群体公共服务供给的偏狭性及其治理[J].中国地质大学学报(社会科学版),2018,(6):123-130.
- [16] 宋保振.“数字弱势群体”权利及其法治化保障[J].法律科学,2020,(6):53-64.
- [17] 宋保振.论“数字弱势群体”及其权益体系[J].山东师范大学学报(社会科学版),2025,(1):107-120.
- [18] 宋维志.重回价值治理:韦伯技术治理理论研究[C]//周尚军.法律和政治科学:第 1 辑.北京:社会科学文献出版社,2019:169-218.
- [19] 张叶东.论绿色元宇宙的法理构造[J].东方法学,2024,(6):148-160.
- [20] 阿马蒂亚·森.以自由看待发展[M].于真,等译.北京:中国人民大学出版社:2002:62-63.
- [21] 陈阳.互联网新兴权利的司法证成——以法官解释权为视角[J].学习与探索,2018,(9):77-84.
- [22] 陈金钊,宋保振.新型人格权的塑造及其法律方法救济[J].北京行政学院学报,2015,(3):32-42.
- [23] 陈建平,刘浩龙.数字弱势群体权利的双重面向及其概念重释[J].人权法学,2024,(5):25-41.
- [24] 陈桂生,史珍妮.数字赋权与数字平权:迈向共同富裕的数字治理[J].学习论坛,2022,(7):66-75.
- [25] 季卫东.主权的嬗变——数字化“魔兽世界”与法律秩序创新[J].交大法学,2023,(5):5-17.
- [26] 周尚军.数字社会对权力机制的重新构造[J].华东政法大学学报,2021,(5):17-26.
- [27] 单勇.跨越“数字鸿沟”:技术治理的非均衡性社会参与应对[J].中国特色社会主义研究,2019,(5):68-75.
- [28] 胡玉鸿.新时代民生保障法治中的“弱有所扶”原则[J].法学家,2022,(5):28-43.
- [29] 星野英一.私法中的人[M].王闯译.北京:中国法制出版社:2004:70-71.
- [30] 侯学宾,郑智航.新兴权利研究的理论提升与未来关注[J].求是学刊,2018,(3):89-99.
- [31] 俞吾金.如何理解康德关于“人是目的”的观念[J].哲学动态,2011,(5):25-28.
- [32] 梅因 H.古代法[M].郭亮译.北京:法律出版社:2019:112.
- [33] 龚向和.数字人权的概念证立、本原考察及其宪法基础[J].华东政法大学学报,2023,(3):6-21.

- [34] 龚向和.数字人权的概念证立、本原考察及其宪法基础[J].华东政法大学学报,2023,(3):6-21;
丁晓东.论“数字人权”的新型权利特征[J].法律科学,2022,(6):52-66;解志勇,吴梦玉.发展与共享:数字社会的权利鸿沟及其平权治理[J].法学评论,2025,(1):13-25.
- [35] 崔靖梓.算法歧视挑战下平等权保护的危机与应对[J].法律科学,2019,(3):29-42;李成.AI歧视的法律治理[J].中国法学,2021,(2):127-147;叶明,冉隆宇.“数字弱势群体”权益的反垄断法保障:功能、限度与路径[J].新疆社会科学,2023,(3):103-114.
- [36] 康德.历史理性批判文集[M].何兆武译.北京:商务印书馆:1997:22.
- [37] 康德.判断力批判[M].邓晓芒译.北京:人民出版社:2002:287.
- [38] 博登海默 E.法理学:法律哲学与法律方法[M].邓正来译.北京:中国政法大学出版社:2017:265.
- [39] 福柯 M.规训与惩罚[M].刘北成,等译.北京:生活·读书·新知三联书店:1999:14.
- [40] Blau J.Theories of the welfare state[J].Social Service Review,1989:26-38.
- [41] European Union.Artificial intelligence act[EB/OL].[2026-03-06].<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32024R1689>.
- [42] Harari Y N.Homo deus:a brief history of tomorrow[M].New York:Harper:2017:356-357.
- [43] Rubin C T.Eclipse of man:human extinction and the meaning of progress[M].New York:Encounter Books:2014:1-200.
- [44] Smith M R,Marx L.Does technology drive history?The dilemma of technological determinism[M].Massachusetts:MIT Press:1994:2-3.

The Jurisprudential Justification and Institutional Construction of the Protection of Rights and Interests of Digitally Disadvantaged Groups from the Perspective of Generative Artificial Intelligence

Feng Chunyang¹, Sheng Zhengyang^{*2}

1 School of Law, Chongqing University, 1104502983@qq.com

2 Faculty of Law, Macau University of Science and Technology, yyszeng@163.com

Abstract: The iterative advancement of generative artificial intelligence technology is not only quietly reshaping the logic of social operation, but also intensifying the marginalization of digitally disadvantaged groups. Protecting the rights and interests of digitally disadvantaged groups is both a practical necessity for bridging the digital divide and an inherent requirement of social civilization transformation, as well as an essential step toward achieving common prosperity. The protection of such rights and interests should adhere to the principles of people-centered governance, preferential protection, and multi-actor collaborative governance. The people-centered principle aims to curb the encroachment of technological rationality upon human subjectivity; preferential protection seeks to rectify the unequal distribution of digital dividends; and multi-actor collaboration addresses

*Corresponding Author

governance dilemmas arising from the dispersion of algorithmic power. Furthermore, to ensure the effective realization of the protection system for digitally disadvantaged groups, efforts should begin at the level of top-level institutional design, including the scientific identification of relevant subjects, the systematic construction of rights content, and the establishment of diversified guarantee mechanisms. Through these approaches, the institutional framework for protecting the rights and interests of digitally disadvantaged groups can be enriched, thereby promoting a virtuous interaction between the high-quality development of generative artificial intelligence technologies and the high-level protection of digitally disadvantaged groups within the rule-of-law framework.

Keywords: Generative Artificial Intelligence, Digitally Disadvantaged Groups, Protection of Rights and Interests, Institutional Construction