

人工智能技术运用于教育的伦理风险及其规避

翁瑜铃

福建师范大学马克思主义学院，福建福州，350117；

摘要：在人工智能技术深度赋能教育的背景下，其作为引领新一轮科技革命的战略力量，展现出提升教学质量、优化资源配置等变革性潜能的同时，一系列显在与潜在的教育伦理风险逐渐暴露。为促进人工智能技术支撑教育向“善”、向“好”发展，有效防控人工智能技术教育应用的伦理风险，需从技术革新、人性回归等多维度构建规避策略，确保技术理性与价值理性相互促进，推动人机协同育人、德性良知重塑与教育优质均衡发展，实现人工智能技术在教育领域的可持续发展。

关键词：人工智能技术；伦理风险；规避策略

DOI：10.64216/3104-9702.25.08.021

习近平总书记深刻洞察人工智能的战略意义，指出：“人工智能是引领这一轮科技革命和产业变革的战略技术，具有溢出带动性很强的‘头雁’效应。”^[1]当前，我国正处于抢抓人工智能发展历史机遇、加快实现高水平科技自立自强的关键阶段，在教育领域，人工智能技术推动了教学模式创新和资源优化，正深刻影响教育发展格局，但随着技术工具理性与教育价值理性冲突的日益凸显，我们亟需审视人工智能技术运用于教育的伦理风险并构建完善的伦理规范。

1 伦理阴霾：人工智能教育应用的茧缚剖析

1.1 技限蔽途：风险丛生，人机协育受阻

算法霸权潜藏不良风险，师生深陷算法“泥潭”。人工智能以大数据驱动的算法和模型为核心，将“万物皆可计算”视为其基本逻辑。^[2]算法却凭借其隐蔽性与主导地位，构建了一种无形却强大的技术霸权，衍生出决策偏见、信息窄化等问题，不知不觉间将师生囿于算法规则之中，削弱其自主思考与行动能力，使其沦为算法操控的“囚徒”，对教育伦理构成威胁。其中人工智能对教育伦理的挑战主要体现在“算法黑箱”、“算法歧视”和“算法推荐”三个方面。

一是人工智能系统依赖预设算法和程序指令生成决策，但深度神经网络的复杂性使其运算逻辑难以解释，步骤推导透明度低，决策形成过程难以追溯，被称为“算法黑箱”现象。当教育场景全面嵌入此类算法体系时，精密的程序设计会重构学习行为的底层逻辑，学生在高度结构化的算法体系中学习路径被简化为数据模型中的概率分布，生活选择被压缩为系统生成的有限选项，逐渐丧失其自主性，沦为智能技术的附庸。

二是算法系统并不是完全客观的，开发人员在设计程序时带有主观选择。若算法被利益或人为目的操控时，只向特定群体推荐单一化的教育内容，忽视其个性化需

求，或在不同群体间形成差异化的价值观引导造成“算法歧视”。诚如海德格尔所述，“此时的技术不单是目的或手段，它在人们的使用过程中总是承载着各种‘偏见’，从而使人们受到技术的束缚和统治，与其认为人们在使用技术，不如说技术在利用人类，使其沦为技术的从属、工具或组成部分。”^[3]在此情形下，技术演变为监管学生、干预个人意志与行为的手段，导致教育主体在自主规划和决策上的能力和权力受限，进而沦为缺乏全面性的“单向度的人”，与教育伦理所追求的发展目标相背离。

三是“算法推荐”通过分析用户数据构建画像，只推送符合学生兴趣的信息，导致其被局限于与其原有观点一致的信息茧房中，难以接触到多元化的观点，使其难以形成全面、客观的认知体系，还可能强化其固有偏见，这实质消解着教育选择的多样性和学生批判性思维的培养空间，将影响教育伦理的公平性和实效性。^[4]

1.2 本真蒙尘：人性受扰，德善之光被遮蔽

首先，技术异化引发教育功利化，消解教师育人责任伦理。人工智能技术的发展为教育创新提供了新可能，但其在教育领域能否发挥正向功能与教师的价值判断与伦理抉择相挂钩。在教育实施过程中，智能技术的融入为教师提供了丰富多样的多媒体教学资源，包括智能课件的自动制作、互动式游戏的设计等，增强了课堂的趣味性，但若教师只图便利和一为迎合学生的娱乐需求，忽视对知识的深度挖掘和思维能力的培养，过度依赖人工智能技术进行泛娱乐化教学，仅把课堂视为单纯的感官刺激环境，此类教学实践就与“以文化育人”的教育理念相悖，使教师从“灵魂工程师”异化为“技术操作员”，削弱了教育应有的人文温度与道德引领功能，背离了教育应有的理性情感导向和职业道德规范，是教师教学失职的表征。

其次,虚拟环境暗藏道德危机,削弱学生道德人格培育实效。立德树人是教育的根本任务,习近平总书记强调“人无德不立,育人的根本在于立德”^[5]。在教育领域,人工智能技术的应用于情感维度上仍存在着较于教师的不足,其缺乏真实的情感体验与道德评估能力,因此难以模拟教师展现的言传身教的育人魅力及情感感召力,因此教师对学生道德品质的塑造是任何技术也不能代替的。教育人际关系是一种对等公正的心际关系,强调彼此互相关怀和尊重。^[6]人工智能技术的融入推动了教学互动方式的革新,由原先的“教师——学生”双向交流转变为“教师——机器——学生”的三方互动模式。人工智能技术以智能教学系统、虚拟教师等形态深度融入教育场景,这种虚拟教育空间虽突破了传统教育的时空限制,但因其脱离真实社会情境,导致道德教育实效性弱化。AI教师、智能拍照搜题及网络课程学习平台等功能的应用,颠覆了教师传统意义上的知识传播者角色。一方面,虽拓宽了学生获取多元化知识和信息的渠道,但若学生过度沉溺于技术营造的虚拟教学场景,而忽视现实中师生的情感互动,使情感交流被数据符号替代,会使其难以形成与教师的道德共鸣。长此以往,学生难以在教育教学中提高道德判断能力,其道德认知可能趋于抽象化,致使不能真正内化为现实社会所需的道德品质。另一方面,学生若只依靠虚拟环境学习,容易导致其道德认知的模糊,易将数字行为与现实伦理割裂,可能诱发网络暴力、信息滥用等失德行为,对学生道德人格塑造带来重大考验。

再者,技术迭代催生资本盛宴,侵蚀技术开发者职业操守。人工智能技术开发群体作为教育智能系统的开发者,其价值取向直接影响着技术应用的伦理边界。技术开发者包括系统架构师、算法工程师等专业群体,他们的技术选择与伦理判断深刻影响着教育人工智能的价值导向。当前,人工智能技术的快速迭代与资本驱动形成共生关系,部分技术开发者在商业利益诱惑下陷入功利化泥淖。一方面,数据资源成为重要资产,而由于数据产权界定模糊、问责机制缺失,技术伦理失范行为难以得到有效规制,部分开发者通过违规获取教育机构数据、售卖用户隐私信息谋取私利,形成了“数据滥用——利益输送”的灰色产业链。另一方面,企业绩效导向迫使技术人员将商业目标置于教育公益之上,而将职业道德底线一位放低。为实现利益目标,部分开发者做出缺德行为,在算法设计中嵌入过度商业化元素,甚至故意保留技术漏洞以维持用户粘性。此种“唯利是图”的做法,更会引发“蝴蝶效应”,导致教育人工智能系统偏离育人本质,埋下诱导消费、信息茧房等伦理风险。

2 破局之路:规避伦理风险的多维策略

习近平总书记强调,要“高度重视人工智能对教育的深刻影响,积极推动人工智能和教育的深度融合,促进教育变革创新。”^[7]为规避智能技术对教育伦理产生的某些消极影响,使智能技术真正成为教育创新的正向推动力,保障教育“善”的可持续性与稳定性,推动教育均衡发展,实现社会公平正义,迫切需要从“根基-中介-外在保障”等多元维度下寻找破局之道,确保技术赋能教育与伦理规约的协同发展。

2.1 复归本真:重塑人性,重铸教育德辉

人工智能时代个体教育“善”是教育伦理的实践外化,其核心在于教育主体伦理自觉向善向上发展。破除因人工智能技术深度介入教育领域而衍生的道德伦理风险和筑牢个体德性的根基就在于提高个体的内在驱动力,推动人性回归本真,提升其伦理自觉。

首先,要唤醒教育者的伦理自觉。教师作为学生成长的引路人应有其教育良心。在面对当今时代人工智能技术对教育所产生不良影响时,要恪守职业道德与教育初心,始终遵循教师职业伦理,在技术运用教育实践中进行思辨反思。《2024年提升全民数字素养与技能工作要点》中指出,要培育高水平的复合型数字人才,包括全面提升师生数字素养与技能履职能力。^[8]因此,在面对人工智能技术带来的教学自主性削弱时,教师应树立清晰的数字化、智能化思维,坚守教育初心,辩证认识技术工具的辅助性本质,提高驾驭智能技术的数字素养,拒绝低俗的泛娱乐化内容,使技术工具真正成为提高教育效果的有力支撑,而不是使自身丧失教育良心,沦为只追求便利、沉溺技术的“奴隶”。

其次,要重塑受教育者的道德人格。消除认知误区是强化信息筛选能力、培养学生道德人格的重要环节。一方面,教师应主动辅导学生掌握人工智能技术的核心理论及其发展动态,理性评估算法模型的局限性及潜在风险。另一方面,教师应坚持在马克思主义理论指导下充分利用案例教学与情景模拟等互动式教学方法,使学生深入剖析人工智能技术引发的教育伦理矛盾,运用辩证唯物主义方法论厘清智能技术生成内容的价值导向,提升其发现问题、分析问题、解决问题的能力,使其在使用人工智能技术时始终秉持道德原则,防止学术不端或违法行为的发生。再一方面,教师应深度联动学生社团与实践平台举办各类现实教育实践活动,并引入批判性思维能力培养机制,构筑一个开放且理性的交流环境,使其在现实的实践活动中深化其对虚拟环境下道德规范的认知,提高其道德判断与选择能力,强化其道德责任感,增强其在面对虚拟环境的具体情境中运用批判性思维对抗错误思想侵害的能力,使学生在现实的师生双向互动中培养更加健全的人格。

最后,在于提升技术开发群体的伦理境界。一方面,借助自主学习与集体教育等方式,培养其伦理道德意识,提高其职业道德水平和道德素养,使其坚守“利民”的伦理价值理念,在技术设计中自觉优化算法设计,提高算法透明度和可解释性,弥补技术漏洞,确保创造的新技术对教育起只发挥正面辅助作用。另一方面,通过系统培训与榜样教育等方式提高其“以人为本”的技术研发价值观和超越功利主义的技术伦理观,使其摒弃“资本至上”的错误价值观,实现其在面对资本诱惑时不向利益低头,主动保护用户隐私,优化数据收集程序,避免泄露风险的发生,实现智能系统设计始终服务于教育的本质需求而不是依靠教育数据赚取背德钱财。

2.2 革新技术: 弥补漏洞, 推动人机协进

人工智能在教育实践中的伦理冲突,本质源于数据采掘和算法设计的固有局限。因此,为规避人工智能技术异化和使学生沦为“单向度的人”等的风险,亟需在人工智能技术开发设计的源头,融入伦理价值考量,将教育伦理的基本原则内化为算法设计的基本要素。在教育领域技术本体的开发构建因基于前瞻性和预防性的伦理设计理念,坚持以保障师生权益为基本要素,确保算法设计的公正性与正义性,防止机器算法对个人权利的侵犯,避免其产生损害教育伦理的负面影响,在坚持公平正义的伦理原则下,突破技术桎梏,消散技术伦理迷雾。

2.3 加强保障: 完善机制, 支撑教育发展

教育作为一个复杂系统,其技术研发、运用及管理各个环节紧密相连,均具有重要作用,且教育公平直接关系到人的成长、社会公平与和谐稳定,因此,要实现人工智能时代教育的向好向善及均衡可持续性发展,亟需更为精确且细致的伦理规范予以引导和支撑,构建人工智能时代教育发展的规范新范式。故而,针对教育人工智能技术应用中所呈现的伦理失范现象,需正确利用外在保障机制,基于现有的人工智能伦理规范体系完善其规范指南,并围绕该指南,出台相应法律法规、责任清单以及问责机制,为研发者、使用者、受用者与管理者等相关主体提供正确行动导向,明晰技术与人的权力界限,规范各类主体行为。

其一,强化机制的规范约束作用。首先,针对教师群体,要建立教师伦理培育机制与责任落实机制,通过定期研讨智能教育伦理议题,深化教师对技术伦理的认知,同时建立师德考评体系,防止其教育初心的坍塌。其次,针对学生群体,要完善道德发展机制与行为监督机制,利用数字媒介普及网络法规知识,提高其道德判断力,并借助教师监督的动态机制作为保障。最后,针

对技术研发相关群体,构建价值导向的职责分工与问责机制,明确系统设计者、架构者、开发者与监测者的伦理责任,从源头上确保人工智能应用过程的透明度、安全性及合规性,并建立技术文档跟踪制度,确保各环节可追溯、可问责,使智能系统从设计到运行的全过程都遵循教育伦理规范,确保人工智能技术在透明、公平、安全和可控的框架内运行,实现技术真正为教育赋能,而非成为压制教育发展的无形牢笼。

其二,发挥政府的补偿性调控作用。长期以来,地域差异是导致教育不均衡的结构性原因,只有精准地采取措施,才能消弭资源分配的不均。政府可以设立专项补偿计划,用政策工具和经济杠杆作为支撑,精准地满足不同区域和学校的实际需求。或是建立教育资源需求的动态监测机制,通过大数据分析找出资源匮乏的地区,实施差异化的配置策略,让资源充分流向最需要的地方。

参考文献

- [1] 高文. 抢抓人工智能发展的历史性机遇——深刻领会习近平总书记关于人工智能的重要论述[N]. 人民日报, 2025-02-24 (01).
- [2] 贾向桐. 当代人工智能中计算主义面临的双重反驳——兼评认知计算主义发展的前景与问题[J]. 南京社会科学, 2019(1).
- [3] 马丁·海德格尔. 海德格尔选集(下)[M]. 孙周兴, 译. 上海: 上海三联书店, 1996.
- [4] Crawford K, Schultz J. AI Ethics: Algorithms and Social Inequality[J]. Yale Law Journal, 2021, 130(3).
- [5] 习近平在北京大学师生座谈会上的讲话[N]. 人民日报, 2018-05-03 (02).
- [6] 何菊玲. 教育正义: 对教育合法性的价值判断[J]. 教育研究, 2020, (11).
- [7] 中共中央党史和文献研究院. 习近平关于网络强国论述摘编[M]. 北京: 中央文献出版社, 2021.
- [8] 中央网信办等四部门印发《2024年提升全民数字素养与技能工作要点》[EB/OL]. (2024-02-21). https://www.cac.gov.cn/2024-02/21/c_1710183257270007.htm.

作者简介: 翁瑜铃(2002年-), 女, 汉, 福建福州, 硕士在读, 研究方向: 思想政治教育理论与实践。