

AI赋能教学下的学生监控：伦理困境与教育价值的平衡研究

张莹, 张楠

宁夏大学教师教育学院, 宁夏 银川

收稿日期: 2026年6月22日; 录用日期: 2026年7月20日; 发布日期: 2026年7月27日

摘要

随着人工智能技术的迅猛发展, AI在教学中的广泛应用正在重塑传统教育模式, 尤其是其在学生监控领域的使用日益频繁。尽管AI监控技术在提升教学效率、精准管理学生行为方面展现出强大能力, 但其引发的伦理争议也不容忽视。本文以“AI赋能教学下的学生监控”为研究对象, 梳理其现实应用背景, 分析当前面临的伦理困境, 探讨其与教育价值之间的张力, 最终提出实现二者平衡的伦理规范与实践机制, 以为未来教育智能化发展提供可持续的治理路径。

关键词

人工智能, 教学监控, 教育伦理, 学生隐私, 技术治理

Student Monitoring in AI-Enabled Teaching: A Study on Balance between Ethical Dilemmas and Educational Values

Ying Zhang, Nan Zhang

College of Teacher Education, Ningxia University, Yinchuan Ningxia

Received: June 22, 2026; accepted: July 20, 2026; published: July 27, 2026

Abstract

With the rapid development of artificial intelligence technology, the widespread application of AI in teaching is reshaping the traditional education model, especially its increasingly frequent use in the field of student monitoring. Although AI monitoring technology has demonstrated great ability to

improve teaching efficiency and accurately manage student behavior, the ethical controversy caused by AI monitoring technology cannot be ignored. This paper takes “student monitoring under AI-empowered teaching” as the research object, sorts out its practical application background, analyzes the current ethical dilemma, discusses the tension between it and educational value, and finally proposes an ethical norm and practice mechanism to achieve a balance between the two, in order to provide a sustainable governance path for the future development of intelligent education.

Keywords

Artificial Intelligence, Teaching Monitoring, Educational Ethics, Student Privacy, Technology Governance

Copyright © 2026 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. 研究缘起与问题提出：AI 赋能教学的时代背景与伦理挑战

随着人工智能(Artificial Intelligence, AI)技术的快速发展, AI 在医疗、交通、安防、金融等多个领域已得到广泛应用。教育作为国家现代化的基础系统之一,近年来也逐步成为 AI 技术重点渗透的对象。从智能评测、个性化学习推送、作业自动批改,到行为分析、考试防作弊、人脸识别签到等, AI 正以前所未有的方式重构教学结构和课堂生态。

尤其在“双减”政策推进、教育信息化的背景下, AI 在教学过程管理中的运用呈现爆发式增长。例如,部分中小学已试点使用“智能课堂行为识别系统”,通过图像识别技术自动识别学生是否走神、打瞌睡、玩手机;又如,线上平台广泛嵌入 AI 学习分析系统,根据学生学习数据智能推送题目、标注薄弱点。可以说, AI 技术正从辅助教学工具向“教育治理主体”角色转变[1]。

然而,技术的快速应用也带来了深刻的伦理冲击。一方面, AI 赋能提高了教育管理的精度与响应力,极大缓解了教师工作负担、提升了教学生产率;另一方面, AI 监控学生所引发的隐私侵犯、知情同意缺失、算法偏见、教育关系异化等问题,已经超出了单一技术问题范畴,进入伦理与教育哲学的核心地带。AI 在教学中的运用呈现出越来越强的“监控性”特征,其实质是将学生身体、行为、心理等各类信息纳入一种“可计算”“可识别”“可预警”的系统逻辑中[2]。这种系统化、量化的控制背后,是对学生成长过程的全景记录与行为约束。正是在这一背景下,教育中使用 AI 监控学生的行为不再是中性行为,而逐渐带有对学生自主性、隐私权甚至人格权的潜在侵犯。AI 是否应当具有“识别懒惰”“判断不专注”的权力?学生在被技术观察时是否具有反驳与选择的权利?系统反馈是否等同于教学判断?这些问题引发了深刻的伦理冲突。

这些问题表明,在“技术合理性”之外,还需对 AI 监控的“伦理合法性”作出系统性反思。特别是在中小学阶段,学生尚处于人格与判断力发展期,对技术不具备足够的辨识能力,更需要教育伦理作为制度护栏。

2. AI 监控技术在教学中的应用与主要伦理困境

人工智能监控技术在教学场域中的应用正以前所未有的速度推进。特别是在基础教育领域,借助 AI 进行学生行为分析、课堂动态记录、学习轨迹识别等方式,已成为当前教育数字化转型的重要路径之一。其背后的逻辑是通过数据驱动优化教学流程,提升管理效率,实现对学生学习行为的可视化与精准化干

预。表面上,这种“智慧化”教学方式有效缓解了传统教学中教师人力不足、课堂反馈滞后等难题;然而,随着AI监控的深入渗透,关于其伦理合法性与教育正当性的质疑也日益增多,集中体现为隐私侵犯、知情同意缺失、算法歧视以及教育关系异化等四个方面的伦理困境[3]。

2.1. AI 教学监控技术三维类型学建构与分类型专属伦理风险

为避免笼统化批判,本文从监控目的、技术手段、应用场景三个互嵌维度建立AI学生监控系统类型学,对不同类别工具的运行逻辑、覆盖范围、风险落点进行细分,精准拆解每一类技术独有的伦理隐患,实现靶向性伦理审视。

依据系统核心服务目标,可划分为纪律管理型、学习分析型、校园安全保障型三类监控系统:

1. 纪律管理型 AI 监控

代表工具:课堂行为识别摄像头、智能头环、考场AI防作弊系统、课堂走神打分系统。核心目的:标准化约束学生课堂行为,降低教师课堂管理成本,统一课堂纪律评判标准。专属伦理风险:一是行为标准单一化,以算法预设的“标准坐姿、专注表情”定义合规行为,压抑学生差异化学习姿态;二是惩罚前置化,系统实时标记违规行为并同步教师、家长,形成持续性行为规训,催生学生表演式学习;三是生物特征强制采集,依靠面部表情、头部动作判定纪律,无差别抓取未成年人面部生物数据,隐私泄露风险突出。

2. 学习分析型 AI 监控

代表工具:线上学习行为追踪平台、学情画像系统、答题路径AI分析工具、课后作业智能监测系统。核心目的:采集学习数据定位知识薄弱点,生成个性化学习方案,实现规模化因材施教。专属伦理风险:一是数据终身化留存,持续积累学生答题速度、错题类型、学习时长等学习轨迹,形成可长期追踪的学生数字画像;二是标签固化效应,算法依据学习数据划分“学困生”“慢思考学生”,易形成自我实现预言;三是数据商业化外流,平台将海量学情数据用于算法迭代、商业培训,学生与家长缺乏知情与否决渠道。

3. 安全保障型 AI 监控

代表工具:校园人脸识别门禁、校内异常行为预警系统、校园视频AI追踪、校外接送人脸核验系统。核心目的:识别陌生人、追踪学生活动轨迹、预警校园欺凌与学生出走风险,保障人身安全。专属伦理风险:一是空间全景无死角监控,学生课间、午休、往返校门等非教学私密活动全程被记录,模糊公共空间与私人活动边界;二是风险标签泛化,将奔跑、独处、低声交谈等正常行为标记为“异常”,扩大监控预警范围;三是轨迹数据跨机构流转,人脸、出入时间、活动路线同步至学校、教育局、安保企业,数据流转缺乏透明监管。

2.2. 基于技术手段的分类及对应伦理风险

依据数据采集与分析技术路径,分为图像识别类监控、数字数据追踪类监控:

1. 图像识别类 AI 监控

代表技术:人脸识别、人体动作捕捉、微表情识别、课堂全景视频解析。核心采集对象:学生面部特征、肢体动作、神态表情、实时影像。专属伦理风险:生物敏感信息不可逆采集,人脸、微表情属于高度敏感个人信息,一旦泄露无法重置;识别算法存在肤色、口音、外貌偏见,对少数民族、深肤色学生、外貌特殊学生产生误判;影像存储周期过长,学生日常画面长期存档,存在泄露、外传隐患。

2. 数据追踪类 AI 监控

代表技术:鼠标轨迹捕捉、屏幕停留时长统计、登录行为记录、答题时序追踪。核心采集对象:线上

操作行为、学习交互数字数据。专属伦理风险：无感知静默采集，学生无法察觉自身操作行为被实时记录；数据颗粒度过细，将短暂停顿、反复翻页等正常思考行为判定为学习低效；线上线下数据联动拼接，整合校内课堂影像与居家线上学习记录，构建完整学生数字全景档案。

2.3. AI 技术的伦理性探讨

(1) 侵犯学生隐私

AI 监控技术在教学中被广泛应用于多个层面，构建了一个覆盖教学全过程的“数据感知环境”。在实体课堂中，越来越多的学校配备了高清摄像头与行为识别系统，用以监测学生的专注状态、课堂举止、互动频率等。这些系统借助图像识别、动作捕捉与面部特征提取技术，实时记录并分析学生是否低头走神、是否频繁交头接耳，甚至可以判断其情绪状态。在在线教学场景中，AI 通过分析鼠标点击、屏幕停留时间、视线轨迹等方式，实现对学生学习参与度的跟踪与预测，生成个性化学习画像。此外，为了应对考试作弊、校内安全等问题，部分学校还引入人脸识别门禁系统、行为异常预警机制等手段，对学生的进出记录、活动轨迹进行全面监控。这些技术的应用虽然提升了管理效率，但其核心特征——“全景式监控”与“算法介入判断”——却带来了严重的伦理风险。

在传统教学关系中，教师对学生的观察更多依赖于专业经验与人际互动，具有一定的模糊性与间接性。而 AI 系统则将学生的生物信息(如面部特征、行为动作、表情变化)和行为轨迹(如浏览记录、操作频次)系统化、数据化、结构化，使得学生在教学空间中几乎处于一种“持续暴露”的状态。尤其是在中小学阶段，学生尚未具备足够的隐私意识与自我保护能力，其个人数据一旦被系统自动采集、存储并可能被第三方机构分析使用，将面临极高的滥用风险。例如，有部分教育平台会将学生行为数据用于商业模式训练，而学生与家长却对此毫不知情，这种技术与权力的极度不对称显然违背了基本的隐私伦理。

(2) 知情同意缺失

AI 监控在教学中的大规模应用，不可避免地触及“知情同意”这一伦理与法律中的核心概念。所谓“知情同意”，本质上是指数据主体在充分了解数据采集、处理、使用、存储与共享的全部内容后，自主、自愿地作出授权决策，并拥有随时撤销该授权的自由。这一原则广泛存在于医学、社会调查、数据隐私等领域，是维护个体人格权利的基本机制。然而在教育系统中，尤其是面向未成年人的教学场景中，AI 系统的快速部署往往绕过了“知情-理解-授权-拒绝”这一伦理闭环，转而以一种隐蔽的、默认的、不对称的方式完成数据收集与处理，形成了结构性困境。

从系统部署过程来看，AI 监控的实施大多由学校、教育局或第三方平台推动，学生和家长被排除在决策程序之外。平台方和校方在系统上线前多以“提升教学质量”“优化学生管理”为由，进行形式化的信息披露，通常仅通过家校群、书面通知或开学说明会进行简略介绍，缺乏对系统运行原理、采集数据类型、使用目的、算法风险等关键内容的详尽说明。而即便有书面材料，其专业术语繁复、信息不透明，也极大抬高了家长和学生的理解门槛，使得所谓“同意”更像是一种被默认前提的合约安排，而非真正基于理解的授权。

系统在技术设计上普遍采用“默认启用”机制，这意味着学生在课堂上一旦进入监控范围，即自动成为数据收集对象，无需额外同意。这种“系统内嵌”的默认逻辑实质上剥夺了学生选择参与的权利，在行为上构成了“被动数据贡献者”。更严重的是，部分 AI 监控工具还延伸至学生在校外的学习行为追踪，如在线作业平台对登录时间、答题路径、学习时长的持续采集，使得学生几乎处于一种“无处不在的算法监视”状态，完全失去了对数据命运的控制权。

现行教育环境中普遍缺乏“知情撤回”的制度保障。即使部分家长和学生事后意识到数据采集存在问题，也难以有效退出或追踪数据流向。学校往往并未设置“数据退出机制”或“数据删除通道”，家

长只能通过非正式渠道提出质疑, 而平台也多以“系统设计限制”或“教学管理需要”为由拒绝撤销请求。这种机制性不对称进一步固化了教育场域中的技术支配结构, 使学生与家长在技术运行中处于弱势地位, 无法真正行使其个人信息控制权。

尤其值得警惕的是, 这种知情同意的缺失对弱势学生群体具有更显著的不利影响。在农村、边远地区或信息素养较低家庭中, 学生家长难以理解系统背后的技术逻辑, 也无能与学校或企业协商。相较而言, 拥有法律常识和技术背景的城市家庭更可能通过主动沟通维护孩子的数据权利, 这种“数字知情鸿沟”使 AI 监控在无形中放大了原有的教育差距, 形成结构性歧视。例如, 一些学校对行为“异常”的学生设立重点行为追踪档案, 而这一过程往往未经家长允许, 也不告知学生是否已被“重点关注”, 这种隐形标签不仅侵犯隐私, 也影响了学生的发展预期。

因此, 知情同意的缺失不仅暴露出 AI 系统“以效能压伦理”的发展路径, 也揭示出当前教育治理体系在技术伦理建设上的滞后性。要实现真正有温度、有尊严的教育技术使用, 必须从制度上确立“真实知情、自由授权、可撤销同意”的操作机制, 并赋予学生及其监护人实质性的决策权与追责权, 将技术纳入教育伦理治理的基本框架。

(3) 算法歧视

在 AI 赋能教学的语境中, 算法不仅承担着信息处理与学习预测的功能, 更在一定程度上介入了教育判断与资源配置。然而, AI 系统本质上并非“价值中立”的技术工具, 其运行逻辑依赖于对历史数据的学习与建模, 这意味着算法极有可能继承甚至放大原有的数据偏见。尤其在学生监控系统中, 算法通过对学习行为、课堂举动、答题效率等表征数据的分析, 将学生划分为不同类别、风险等级或行为特征标签。这种“自动分类-标签化”的过程在提升系统效率的同时, 也埋下了潜在的教育歧视与不公平隐患。

算法偏见的产生源于模型训练数据的不平衡性与隐含价值倾向。AI 系统往往依赖大量已有教育数据进行监督式学习, 这些数据可能来源于某一特定地区、学校、群体, 难以全面代表学生多样性。比如, 在一个模型中, 如果“高效率”学生的数据主要来自城市中学、男性群体或某一学习风格, 那么该模型就可能在无形中把“城市”“男性”“速答型”视为优质学习特征, 从而对农村学生、女生、慢思考型学生形成“不合范式”的隐性偏见。更有甚者, 一些早期采集的数据本身就带有社会性别刻板印象、地域歧视等问题, 而算法无法区分这些“结构性不公”的背景, 只会将其作为“客观规律”进行再现与放大。

AI 监控系统的自动分类机制在教育实践中往往以“标签”形式呈现, 如“高专注学生”“干扰型行为”“注意力波动群体”“潜在学习困难对象”等。虽然这些分类意在帮助教师了解学生差异, 但标签化一旦固化为系统属性, 便可能成为教师判断学生的前置框架, 进而引发“算法决定论”的认知偏误。例如, 教师在教学互动中可能因为系统提示某位学生“经常走神”, 而下意识减少与其眼神交流、提问频率与个性化关注, 从而形成一种“技术诱导下的差别对待”。这不仅违背了教育公平原则, 也削弱了教育者的专业判断力与情感介入能力。

算法标签还会对学生自身产生深远影响。一旦某位学生被反复归入“低效学习者”或“注意力不足者”的类别, 系统将持续反馈这类评价结果给教师、家长甚至学生本人, 在校园生态中构建起一种隐形的“技术画像”。学生在感知这些标签后, 极易出现“自我实现预言”现象, 即在被持续暗示“不如别人”的情况下降低自我期待、逃避学习挑战, 甚至将系统标签内化为身份认同。这种由技术定义而非教育理解所带来的概念, 不仅违背了教育促进个体成长的价值宗旨, 更在心理层面压抑了学生的主动性与成长潜能。

更为复杂的是, 一些算法偏见具有深层的结构性与不可察觉性, 使得教育管理者甚至教师自身都难以意识到其负面影响。例如, 行为识别系统的摄像头若在肤色感应方面存在失衡, 可能导致对深肤色学

生“专注度”识别精度下降,形成误判;或者系统对语音识别中口音差异不加区分,可能将部分方言学生识别为“不回应”或“干扰行为”,而这种行为往往难以追责,也难以被系统自动纠正。一旦这些误差嵌入教师评价体系、学生学业跟踪报告与家校沟通机制,就会形成长期影响,最终演变为结构化的教育不平等。

(4) 教育关系异化

AI 监控对传统教育关系造成了深刻异化,破坏了教育本应具有的人文性与信任机制。教育不仅仅是一种管理行为,更是一种关照人的生命成长的实践。在传统师生关系中,教师通过观察、对话、感受来理解学生行为背后的动机与情感,从而作出恰当的引导与教育判断。然而在 AI 系统主导下,教师越来越依赖系统生成的数据与报告来判断学生是否“表现正常”,这在无形中将师生关系转化为“技术-管理者与被管理者”关系。当学生意识到自己时时刻刻被系统记录与分析时,容易产生强烈的被监视感与心理负担,部分学生甚至会出现焦虑、抵触、顺从表演等非自然行为,严重影响其真实自我表达与主体性发展。此外,教学场景的情感因素被技术干预稀释,教学过程日益趋于“程序化”,教师专业判断空间被压缩,教育的创造性与人文关怀被技术逻辑所吞噬。

AI 监控技术在教学中的大规模应用虽具备管理与效率优势,但其带来的伦理困境不容忽视。从学生权利的被动化,到教师角色的弱化,再到教育关系的异化,这些问题都表明 AI 监控在教育系统中的应用亟需伦理审视与制度规范。在没有建立起明确边界、伦理框架与权力监督机制之前,任何“技术优化”都可能沦为“教育异化”的隐性推手。因此,唯有正视并回应这些伦理困境,方能推动 AI 与教育价值实现真正的协同共生[4]。

2.4. AI 监控的适用性探讨

现有研究多聚焦 AI 监控的伦理缺陷,容易形成单向度批判视角,忽略技术在特定教育场景下不可替代的正向功能。本节辩证分析 AI 监控的合理适用场景与滥用红线,客观界定技术的价值边界,构建“约束下使用、规范中赋能”的中立分析立场。

(1) 面向特殊学习需求学生的精准支持

针对自闭症、注意力缺陷多动障碍(ADHD)、学习障碍等特殊学生群体,学习分析型 AI 监控具有独特的干预价值。AI 可持续捕捉学生课堂情绪波动、学习耐受时长、互动反馈模式,量化记录学生何时出现情绪崩溃、注意力涣散,为特教教师提供连续、客观的行为数据,弥补人工观察碎片化、主观性强的短板。依托数据可定制个性化课堂节奏、适配性学习任务,降低特殊学生的学习挫败感,实现无障碍融合教育。该场景下监控以“干预帮扶”为核心目的,而非纪律管控,具备天然伦理正当性。

(2) 校园人身安全的底线保障

安全保障型 AI 监控是校园风险防控的关键工具:其一,人脸识别门禁可快速甄别外来无关人员,阻断校园暴力、校外人员闯入风险;其二,AI 行为识别可实时识别追逐推搡、孤立欺凌、学生独处哭泣等异常场景,第一时间推送预警给安保与班主任,弥补人力巡查存在的时间盲区;其三,上下学人脸核验系统可精准匹配学生接送人,降低未成年人走失、错接风险。对于寄宿制中小学、低龄小学,AI 监控是守护学生人身安全的低成本、高效率手段,具备公共安全层面的合理性。

(3) 大规模普及个性化均衡教育

在师资薄弱、城乡教育资源不均衡的区域,学习追踪类 AI 监控能够破解大班教学“因材施教难”的现实困境。AI 自动采集全体学生答题正确率、知识漏洞、学习时长,批量生成分层学习方案,无需教师投入海量时间单独分析每位学生学情;线上教学 AI 监控可同步记录偏远地区学生课堂参与度,让异地支教教师实时掌握远端学生学习状态,缩小城乡课堂反馈差距,推动大规模教育公平落地。

(4) 标准化考试的公平维护

考场 AI 防作弊监控依托图像识别、视线追踪技术, 统一甄别偷看、传递纸条、携带电子设备等作弊行为, 消除人工监考中主观偏袒、识别疏漏问题, 建立统一、客观的考试评判标准, 维护升学、学业测评的公平秩序。

3. 教育价值视域下的 AI 监控正当性探讨

(一) 教育价值的本质张力: 技术理性与教育人文的冲突

人工智能技术的迅速嵌入教育场域, 使得技术理性日益主导教学过程的组织逻辑与管理方式。然而, 教育作为一项典型的人文事业, 其核心价值并不只是传递知识或管控行为, 而是致力于激发个体的内在潜能、促进人的全面发展与自由成长。AI 监控技术的广泛使用固然可以提高教学效率、优化课堂管理, 但其背后所包含的“算法决定论”“数据主导论”等技术主义逻辑, 正在悄然改写教育价值的传统结构。在这种情境下, 亟需从教育本体论与伦理视角出发, 探讨 AI 监控在教学中的正当性边界及其与教育价值之间的深层冲突。

首先, 教育的根本目的不仅仅是行为管理和知识传授, 更在于塑造人的独立人格、批判思维与社会责任。从夸美纽斯、卢梭到杜威等经典教育思想家均强调, 教育不应沦为单向度的输入输出过程, 而应是一种唤醒个体生命意义的过程。这种意义上的教育具有高度的开放性、复杂性与生成性, 而 AI 监控系统的设计逻辑则倾向于以“规范 - 反馈 - 纠偏”为核心模式, 对学生进行行为识别、状态分析与结果预判。其背后隐藏着一种隐性的工具理性倾向, 即以“效率最大化”“可量化”为最高准则。这种技术理性与教育所追求的价值理性之间构成了本质性冲突。

AI 监控系统在教育实践中的合法性, 必须以“教育正义”与“伦理合法性”为基础, 不能仅以技术是否“有效”“准确”“先进”作为评判标准。例如, 在一些课堂中部署的“智能头环”“行为打分系统”, 虽然在提升课堂纪律和行为规范方面具有明显效果, 但其核心机制是以技术手段强化外在控制, 并未真正尊重学生的成长规律与心理需求。教育应是基于理解、信任与引导的关系性活动, 而非依赖机器对“偏差”的即时识别与矫正。过度依赖 AI 技术进行学生评价与管理, 实际上会削弱教师的教育判断力与情感感知力, 进而导致教学关系的去人化, 学生也可能逐渐将自身行为调整为“被系统接受的状态”, 丧失内在动机与自我调节能力[5]。

(二) 伦理正当性与公平性前提下的合法性判断

从教育伦理的角度出发, AI 监控学生的行为必须遵循“最小干预”与“目的正当”原则。教育伦理强调“尊重学生主体性”, 意味着教育工作者必须承认学生在认知、情感与价值形成上的独立地位, 不得将其简化为“被管理的对象”或“被优化的数据模型”。然而, AI 监控系统的设计与实施通常围绕“行为改进”“学习提效”等目标展开, 忽视了学生在成长过程中的多样性与差异性。系统将学生纳入“理想行为轨道”的过程中, 往往以统计学意义上的“平均值”“标准行为”作为判断依据, 对偏离轨道者进行提醒甚至惩戒, 这种机制事实上形成了一种“教育中的技术性歧视”, 不利于保护个体差异, 反而助长了教育的标准化倾向。

此外, 教育评价本应是多维度、动态化、过程性的, 而 AI 监控系统则倾向于提供单一、即时与碎片化的数据反馈, 其算法评价的逻辑很难覆盖学生复杂的学习动因与心理状态。举例来说, 一个被系统判断为“注意力不集中”的学生, 可能只是因短时情绪波动或思维方式差异而产生“行为异常”; 但系统并不具备人类教师的情感理解与发展性视角, 很可能将这一行为固化为“消极学习”的标志。这种算法评价对学生形成潜在负面影响, 不仅可能扭曲教师的教育判断, 也会改变学生对自我能力与价值的认知, 从而妨碍其自尊、自信的建立[6]。

更重要的是, AI 监控系统在教育场景中的合法性不仅是伦理问题, 也涉及社会公平与教育正义的问题。大量研究表明, 教育技术的部署往往与学校财力、地区资源分布、家长数字素养等因素紧密相关。这意味着在不同地区、不同层级的学校中, AI 监控的实施水平与监管能力存在显著差距。一些重点学校或发达地区可能拥有成熟的监管机制和家长反馈渠道, 而边远或资源薄弱地区则更容易出现“技术滥用”“无人问责”的现象。这种数字鸿沟不仅加剧了教育不平等, 也让技术治理沦为教育权力失衡的新工具。因此, 在推动 AI 赋能教学的过程中, 必须充分考虑制度保障的可行性与风险分担机制, 确保技术应用不会进一步放大教育公平的裂缝。

在此背景下, 判断 AI 监控系统在教学中是否具有正当性, 应从以下几个维度进行系统考量: 其一, 目的正当性, 即系统是否服务于学生发展、是否以教育价值为中心, 而非单纯服务于效率或控制; 其二, 方式的适当性, 即系统是否采取最小干预原则, 是否在尊重学生人格的前提下设计技术介入路径; 其三, 过程的公开透明性, 即学生、家长与教师是否有知情权与参与权, 系统是否具有申诉、纠错与解释机制; 其四, 数据使用的可控性, 即学生数据的收集、存储、处理是否遵循法律法规与伦理规范, 是否具备数据最小化与权责明晰原则。

AI 监控技术在教学场景中的正当性并非天然存在, 而是建立在教育伦理、人文价值与社会责任三重基础之上的。忽视对 AI 系统的价值审视, 只会导致教育从“人性关怀”退化为“技术治理”, 从“激发潜能”沦为“行为操控”。要实现 AI 与教育的真正融合, 必须坚持以教育本质为根基, 以学生主体性为核心, 以伦理框架为保障, 构建起面向未来的、有温度的教育技术应用体系。

4. 平衡路径建构: 面向未来的伦理规范与实践机制

AI 技术在教学场域中的快速应用已成为不可逆转的趋势, 其在课堂管理、学习分析和学生行为识别中的强大能力, 使其越来越多地介入教育决策过程。然而, 正如前文所述, AI 监控系统一方面提高了教学效率, 另一方面却引发了一系列伦理争议与教育价值的张力问题。要真正实现“AI 赋能”而非“AI 控制”, 关键在于构建一套以教育本质为导向的平衡机制, 确保技术进步不以牺牲学生权益与教育正义为代价。因此, 必须从教育制度、教师角色、学生主体与治理机制四个维度, 协同推进伦理规范与实践路径的建构, 推动技术与人文的深度融合。

(一) 构建教育制度层面的伦理规范与治理边界

AI 监控系统的广泛部署必须以系统性的制度约束为前提, 而不能依赖学校或企业的自我约束。首先, 应在国家层面制定关于人工智能在教育场景中使用的专门法规或指导性规范, 明确 AI 技术在数据采集、行为分析、学习评价等方面的红线。例如, 可借鉴欧盟 GDPR (通用数据保护条例) 中“数据最小化”“目的限制”“知情同意”等原则, 对学生数据的采集范围、使用目的、保留时间进行强制规范, 从源头上避免学生信息被过度监控或滥用。

其次, 应设立由教育行政部门牵头的“教育技术伦理评估机制”, 对 AI 系统进行事前、事中和事后的伦理审核。任何一套 AI 系统进入学校之前, 必须通过伦理影响评估程序, 对其算法模型、决策逻辑、数据安全、学生权利保护等方面进行全面检视, 并接受社会监督。此外, 学校应建立“技术使用备案与公示制度”, 将所有正在使用的 AI 系统名称、功能、数据处理方式等信息向师生和家长公开, 确保使用过程的透明与可追责。

同时, 要引导教育技术企业强化“教育伦理责任”。技术开发者在设计 AI 系统时, 不能仅以“商业效率”“用户体验”为主导, 而应设立“教育伦理设计岗”或咨询机制, 确保系统架构嵌入教育价值观。例如, 系统设置中应避免对学生进行默认标签化, 应增加教师自主判断权限, 并在界面中设置学生知情说明、数据撤回申请等功能, 实现技术工具的人本化转型。

(二) 重构教师角色与 AI 协同的教育边界

教师作为教育活动的核心主体,在 AI 监控系统中的作用不应被削弱为“执行者”或“系统辅助员”,而应被赋予更加积极的“技术解释者”“伦理把关人”“教育协调者”角色。要实现这一角色转型,需从以下几方面着力。

首先,强化教师的“AI 素养”与“伦理判断能力”。当前教师在面对 AI 系统时常存在“技术依赖”“算法迷信”等倾向,缺乏对技术逻辑的深度理解与反思能力。为此,教师职前职后培训中应设置 AI 教育伦理课程,帮助教师掌握 AI 技术的基本运行原理、算法偏差风险、数据伦理边界等知识内容,增强其在面对 AI 判断时的质疑与干预能力。例如,当 AI 系统判断某学生“注意力不集中”时,教师应能够基于自身观察与学生互动,综合判断该数据是否准确,而非盲目接受系统建议。

其次,应明确教师对 AI 系统拥有“最终教育判断权”,避免教学行为被技术逻辑主导。可通过制度设计赋予教师一定的数据解释权与“技术介入限制权”,如当教师认为系统数据有误或使用不当时,拥有暂停系统运行、申请重新评估的权利。同时,教师在使用 AI 系统过程中应承担对学生的“数据责任教育”任务,帮助学生理解自身数据被采集与使用的逻辑,建立学生对技术世界的反思性认知。

最后,要推动教师形成“人机协同”的教学能力,在数据支持下提升教育个性化水平。例如,教师可将 AI 生成的学习报告作为参考,在此基础上开展个别化谈话、差异化教学设计,而非仅依据系统排序进行管理决策。只有实现技术与教师的有机协同,才能真正发挥 AI 系统在支持教育中的价值,而不是替代教育判断。

(三) 强化学生主体性意识与数据权利保护机制

学生作为 AI 监控系统的直接影响者,应成为技术治理中不可或缺的主体力量。然而,现实中学生常常被排除在 AI 系统部署的决策链之外,对自身被如何记录、如何评价、如何处置几乎毫无知情与干预权。因此,必须在制度与实践层面全面强化学生的数据权利保护与主体性意识。

其一,要建立学生“数据知情-授权-撤销”三位一体的权利机制。任何涉及学生个人信息的 AI 系统在部署前,必须向学生及家长提供完整、易懂的知情说明,内容包括数据采集内容、使用目的、保管期限、第三方访问权限等。同时,应明确规定学生及监护人有权拒绝授权、不受不利影响,有权随时撤销同意并要求删除个人数据。这不仅是对《未成年人保护法》《个人信息保护法》的具体落实,更是对教育伦理中“尊重主体性”的重要体现。

其二,要将“数据素养”纳入学生核心素养体系,引导学生形成“数字公民”意识。教育部门与学校可在课程中增设“AI 与社会”“个人数据与权利”等主题活动,通过情境模拟、案例分析、角色扮演等方式,帮助学生理解算法如何影响其评价、选择与命运。同时,鼓励学生参与学校 AI 系统的评估与反馈机制,形成“有反思的技术使用者”身份。

其三,要关注弱势群体在 AI 系统中的特殊风险。例如,对行为较为外向或学习风格非传统的学生,AI 系统可能因识别精度不足而形成“异常标记”,造成其在评价与资源分配中的边缘化。为此,应建立系统纠错机制与教师干预路径,保障所有学生在 AI 系统下的发展平等权利。

(四) 构建可持续发展的教育 AI 治理机制

从长远看,AI 赋能教育的治理体系不能仅依赖某一层级或单一主体的努力,而需建构一个多元协同、持续优化的治理机制,实现技术发展与教育伦理的动态平衡。

首先,建立“教育 AI 伦理联盟”平台,推动教育行政机构、技术开发商、教师组织、家长协会、高校研究机构等多方共同参与教育 AI 系统的伦理议题讨论与标准制定。通过跨领域对话,确保伦理规范既有理论高度,又具现实可行性,推动从“技术导向”转向“治理导向”。

其次,推动国家层面的“教育 AI 伦理实验区”建设,选取部分地区或学校作为试点,对 AI 系统在

真实教育环境中的应用进行伦理评估与制度模拟,探索数据最小化、算法透明、责任归属、师生协同等核心问题的实践路径。实验结果可反哺国家政策与行业规范,形成“制度-实践-反馈”闭环体系。

最后,设立“教育 AI 审计机制”,定期对正在使用的 AI 教育系统进行功能审查与伦理测评,包括算法更新、误判率统计、学生反馈分析、教学关系影响评估等。将伦理指标纳入教育信息化考核体系中,使学校不仅关注技术覆盖率,也重视技术应用中的伦理质量,推动技术治理向纵深发展。

AI 监控技术在教育中的伦理问题并非单靠某一制度或行为可以化解,它需要教育制度、教师行为、学生认知与技术机制的系统性重构。在 AI 全面赋能教育的时代,只有通过多维度的治理逻辑,才能实现技术效能与人文价值的真正融合,为未来教育构建一条有尊严、有温度、有规范的智能发展之路。

参考文献

- [1] 连燕纯,刘盾.数字平台促课后服务质量质齐升[N].中国教师报,2024-06-12(004).
- [2] 陈婷,张淑,王静.数字化赋能高质量课后服务课程建设的机理、困境与路径[J].教育与教学研究,2024,38(12): 1-11.
- [3] 吴南中,唐之茜,朱喜祥.“人工智能+教育”行动的理论超越与实践旨趣[J].教学与管理,2026(20): 1-6.
- [4] 方东玲,赵一婷,唐烨伟,等.智能驱动与范式跃迁:人工智能赋能学生评价转型研究[J].中国电化教育,2026(6): 78-86.
- [5] 王娟.生成式 AI 时代数据伦理教学的“破”与“立”[J].贵州师范学院学报,2025,41(2): 38-44.
- [6] 刘丽书,陈旭远,胡洪强.AI 时代教师高质量发展的伦理缺位及补位分析[J].现代中小学教育,2025,41(5): 58-62.