

数字伦理视域下青少年网络暴力防治的素养教育路径创新

陈威, 陈锋

浙江安防职业技术学院, 浙江 温州 325000

DOI:10.61369/EIR.2025060023

摘要：本研究聚焦人工智能时代青少年网络暴力的复杂伦理困境，系统剖析算法助推语言暴力、数字身份审美霸凌、技术匿名性道德感降低及生成式 AI 滥用四大典型场景，揭示其背后的伦理混淆与责任模糊性特征。基于数字伦理导向，构建包含伦理认知、共情责任、批判性算法素养及积极行动四大模块的素养教育框架，并提出教育体系深化、技术治理重构、法律规制协同及社会支持网络培育四大创新实践路径，形成“四位一体”素养模块与多方协同的闭环治理生态系统，实现从技术认知到伦理自觉的范式转变，为青少年网络暴力防治提供理论支撑与实践方案。

关键词：数字伦理；青少年网络暴力；素养教育路径；技术治理重构

From the Perspective of Digital Ethics: Innovations in Literacy Education Paths for the Prevention and Control of Teenagers' Cyberbullying

Chen Wei, Chen Feng

Zhejiang College of Security Technology, Wenzhou, Zhejiang 325000

Abstract : Focusing on the complex ethical dilemmas of teenagers' cyberbullying in the era of artificial intelligence, this study systematically analyzes four typical scenarios: algorithm-facilitated verbal violence, aesthetic bullying based on digital identity, reduced moral sense due to technical anonymity, and abuse of generative AI. It reveals the underlying characteristics of ethical confusion and responsibility ambiguity. Guided by digital ethics, the study constructs a literacy education framework encompassing four modules: ethical cognition, empathic responsibility, critical algorithmic literacy, and positive action. Furthermore, it proposes four innovative practical paths: deepening the education system, reconstructing technical governance, coordinating legal regulation, and cultivating social support networks. This forms a closed-loop governance ecosystem integrating the "four-in-one" literacy modules and multi-stakeholder collaboration, realizing a paradigm shift from technical cognition to ethical awareness. The research provides theoretical support and practical solutions for the prevention and control of teenagers' cyberbullying.

Keywords : digital ethics; teenagers' cyberbullying; literacy education paths; reconstruction of technical governance

引言

随着人工智能技术深度嵌入青少年数字生活，网络暴力已突破传统形式，演变为技术特性与伦理困境交织的复杂问题。如2023年杭州互联网法院审理的 AI 换脸侵犯公民个人信息案^[1]，这些事件暴露出传统网络暴力防治体系在技术伦理挑战面前的失效：当算法推荐放大群体极化、匿名社交消解道德约束、AI 工具降低施暴门槛时，简单的“禁止骂人”式教育已完全无法应对。中国互联网络信息中心数据显示，截至2024年6月，我国10-19岁青少年网民规模约为1.5亿人^[2]。当前网络暴力形态正发生深刻转变——从直接的言语攻击转向隐蔽的符号暴力，从个体自发行演变为算法驱动的群体行为，从现实空间的矛盾延伸至虚实交融的数字身份领域，这种演变对防治体系提出了全新挑战。

现有研究多聚焦技术防控的工具优化或个体行为干预的策略探索，却鲜少从数字伦理的核心视角切入，剖析数字暴力生成的根源机制与结构性诱因，导致现有治理方案往往停留在“治标”层面，难以应对其复杂多变的演化态势。一方面，平台长期将“技术中立”作为免责盾牌，刻意淡化算法设计中暗藏的流量至上逻辑对冲突的放大效应。在资本逐利导向下，算法往往通过精准捕捉用户情绪偏好，

课题信息：2024年浙江省妇联、浙江省妇女研究会课题《新时代青少年网络素养教育提升对策的研究——聚焦网络十大领域》（编号202423）。

作者简介：陈威（1990—），男，汉族，浙江温州人，讲师，硕士研究生，研究方向：计算机应用技术、物联网应用技术，数字伦理与网络素养。

优先推送具有争议性、对抗性的内容以提升用户停留时长与互动频率。另一方面，青少年在数字环境中面临双重伦理认知困境：一是数字空间的虚拟性特征弱化了他们对行为后果的具象感知，使其难以充分理解数字行为背后的物质性影响；二是在技术工具理性的过度渗透下，部分青少年陷入“技术至上”的认知误区，将技术操作的有效性置于价值判断之上，认为只要借助技术手段实现自身诉求，便无需恪守道德准则，最终导致数字行为的伦理边界持续模糊。

基于上述现实困境，本研究以数字伦理理论为核心支撑，通过剖析四大典型场景伦理困境，构建数字伦理导向防治体系，旨在破解技术中立性责任模糊困局，培育青少年数字伦理自觉，实现教育、技术、法律、社会的协同治理，重建数字文明时代公共伦理秩序。

一、青少年网络暴力的典型场景与伦理困境分析

人工智能时代，青少年所遭遇及施加的网络暴力，已不再局限于简单的辱骂与攻击，而是深度嵌入数字技术的特性之中，呈现出复杂的伦理困境。下面针对四大典型场景及其背后的伦理困境进行深入剖析。

（一）算法助推下的语言暴力

在短视频、社交媒体等平台，算法基于用户的浏览轨迹、互动数据精准捕捉偏好，通过“点赞越多推送越密集”的正向反馈机制，极易为青少年构建封闭的“信息茧房”。当校园纠纷的片段化剪辑、服饰打扮的争议性评价、偶像偏好的对立言论等内容出现时，算法会持续向立场相近的青少年推送片面化、极端化信息，叠加数字空间的匿名性降低行为约束，他们在群体情绪发酵中迅速丧失独立判断能力，以“维护正义”“守护偶像”等正当化名义，对当事人发起大规模的语言围攻、人身攻击和人格贬损。平台算法看似中立，但其追求用户参与度的设计逻辑，实质上助长了冲突与对立。在此场景下，这种人机协同催生的暴力行为，既造成青少年道德责任认定模糊、平台技术责任追溯困难的治理困境，更暗藏目的正义与手段非正义的伦理混淆。传统素养教育若只强调“不骂人”的表层规范，若未能引导青少年解构虚假正义的包装，识别算法的流量陷阱，终究难以从根源上遏制此类暴力行为。

（二）数字身份建构中的审美霸凌

青少年在抖音、小红书等以图像和视频为主导的社交媒体上，通过滤镜美化、场景搭建、内容筛选等方式精心建构“数字身份”，以此获取关注、拓展社交圈。在这一过程中，他们不自觉地自我异化为被观看、被评价的客体，陷入自我物化的认知陷阱。这种倾向不仅让他们对他人的外貌、身材、服饰、消费水平评价极度敏感，一旦不符合群体推崇的标准便易遭受符号化打压，沦为暴力受害者；同时，为了融入圈层、获得群体认同，他们又可能主动迎合单一审美与价值标准，通过恶意P图对比、标签化嘲讽、圈层内排斥等方式，对“异类”实施符号暴力，转而成为施暴者。这种新型暴力摒弃了粗俗语言，却以“审美正义”“圈层共识”为伪装，形成隐蔽的霸权。青少年在从众心理的驱使下，不自觉地成为这种霸权的共谋与执行者，将“不同”等同于“不好”，彻底背离了数字空间应有的尊重多元、包容差异的伦理原则，加剧了网络社交的伦理失序。

（三）技术匿名性下的道德感降低与关系欺凌

在 Soul 等匿名社交软件、MOBA 类游戏的公屏聊天频道、各

类临时组建的兴趣或矛盾群组中，青少年借助虚拟身份的遮蔽，往往会产生短暂脱离现实身份约束的“去抑制效应”。在这里，他们惯常实施的人肉搜索、制作传播带有丑化元素的侮辱性表情包、在群组内发起针对性公开批判等行为，变得比现实中更为肆无忌惮。他们常常将校园里的同学矛盾、日常人际间的嫉妒情绪，通过线上平台进行放大转移，对目标对象展开持续性的精神打压、社交孤立，甚至煽动圈层内其他人参与围攻。由于身份的隐匿性，施暴者所感知到的道德谴责压力与法律惩戒威慑力显著降低，多数人将此类行为轻视为“开玩笑”“闹着玩”，完全未能意识到其行为可能给受害者带来长期焦虑、抑郁、社交恐惧等深刻心理创伤，这一现象深刻反映了青少年群体对数字行为背后真实物质性后果的伦理感知严重迟钝，也暴露了其数字伦理认知体系的明显缺失。

（四）生成式 AI 滥用下的新型污名化手段

随着人工智能生成内容（Artificial Intelligence Generated Content, AIGC）技术的普及，一种技术门槛更低、危害性更强的暴力形式开始出现。青少年利用 AI 换脸技术将同学的头像嫁接至不雅图片或视频中，或利用 AI 语音模仿技术伪造他人言论进行扩散。这类行为不仅造成直接的诽谤和侮辱，更动摇了人际信任的基石——真实性，对受害者造成严重的精神打击。当任何音视频都可能是伪造时，受害者极易陷入自证的困境，这不仅是个体间的暴力，更是对整个社群信任生态的攻击。

二、数字伦理导向的网络暴力防治素养教育框架构建

传统侧重于技术操作技巧传授与表层行为规训的素养教育，因缺乏对数字暴力背后算法逻辑、身份建构、伦理责任等核心问题的深度挖掘，已难以应对当前复杂的技术伦理挑战。为推动青少年网络暴力防治从外在被动的行为约束，转向内在主动的伦理自觉与价值坚守，本部分构建了数字伦理导向的网络暴力防治素养教育框架。该框架针对前文剖析的各类场景背后的特定伦理困境，精准设计对应素养模块，四个模块分别聚焦伦理认知、共情与责任、批判性算法素养、积极行动者等核心维度，相互关联且层层递进。通过进一步引导学生清晰认识到自身行为与算法、平台、社群等数字生态之间相互塑造的动态关系，扭转青少年将技术视为纯粹中立工具的普遍认知偏差，是探寻青少年网络素养提

升有效路径、夯实数字空间治理教育根基的关键所在。

（一）伦理认知模块

技术作为一种工具，其物理属性或许具有中性特征，但当它被设计转化为承载设计者意图、商业逻辑与文化观念的数字产品或平台后，其非中立性愈发凸显。伦理认知模块以此为核心切入点，聚焦破解青少年对“技术中立”的认知偏差。模块通过解析算法流量逻辑对冲突的放大机制、平台商业导向对暴力内容的隐性纵容等核心议题，结合算法助推语言暴力、匿名场景施暴等典型案例，采用案例研讨、算法透明化模拟实践等方式，引导青少年穿透技术工具的表层表象，认清数字产品背后的价值倾向与伦理风险。同时，模块强化数字行为物质性后果的认知培育，通过梳理数字暴力引发的心理创伤、法律责任等真实案例，帮助青少年建立“技术使用—伦理后果—责任归属”的关联认知，最终培养其对数字技术的伦理洞察力与批判性思维，为后续素养养成奠定认知基础^[9]。

（二）共情与责任模块

这一模块旨在重建青少年在数字空间中的道德情感，并加速青少年责任意识的觉醒。利用可视化展示方法展示一条网络暴力言论如何通过算法扩散，如某学生发布的侮辱性评论首先被算法识别并推送给更多的与该学生有着相似心理的用户，这类用户进一步点赞转发被更多的网友看到，最终触及受害者的社交圈。这种可视化展示方法能够让青少年直观地感受到“小恶”如何在数字空间中被算法不断放大，最终造成远超预期的系统性伤害。通过角色扮演活动，让学生亲身体验网络暴力的施害者与受害者角色，将抽象的网络暴力转化为可感知的情绪上海，使青少年从旁观者转变为共情者。

（三）批判性算法素养模块

首先，通过拆解算法逻辑的形式引导学生理解推荐算法、排序算法等核心技术逻辑；其次，通过具体实例让学生直观看到算法会优先推送同质化内容，从而使学生意识到算法并非客观中立，而是嵌入设计者价值观与商业逻辑的载体。这能够使青少年在复杂网络舆论场中保持独立思考，既避免沦为网络暴力的盲从者，又能成为理性发声者。

（四）积极行动者模块

积极行动者模块本质价值在于解构网络暴力的技术逻辑与认知误区，通过建设性行动重建数字公共空间的伦理秩序。通过系统地模拟冲突场景训练，可使得青少年从被动承受网络暴力的受害者转向构建理性对话空间的领导者。这种转变使青少年从“数字消费者”升级为“数字公民”，其行动不仅提升个体伦理素养，更通过“行动—反馈—优化”的闭环机制，推动平台算法调整、社区规范重构、制度政策完善，最终实现数字生态的整体优化。

三、数字伦理视域下青少年网络暴力防治对策

基于数字伦理导向的网络暴力防治素养教育框架，本章提出四大创新实践路径，旨在将“四位一体”的素养模块融入青少年的学习与生活场景，实现从理论到实践的跨越，构建一个多方协

同、闭环治理的教育生态系统。

（一）教育体系的伦理深化与实践路径

通过分层课程设计、教学方法创新与师资专业化建设形成防治网络：小学阶段以“尊重差异”主题绘本培育基础伦理意识；中学阶段开展批判性算法教育，拆解短视频推荐机制中的流量逻辑与偏见放大效应；大学阶段通过反网络暴力活动与数字公民项目强化行动能力，形成认知—实践闭环。教学方法创新实现技术逻辑可视化与伦理体验场景化，如开发放大算法效应的模拟软件，通过角色扮演建立“技术行为—心理影响—社会后果”认知联结。师范院校需将数字伦理纳入必修课，中小学配备专职教师构建三维教学体系，最终实现伦理自觉范式转变。

（二）技术治理的伦理化重构

技术治理需打破技术中立性的思维，构建主动防御体系。算法设计遵循“伦理优先”原则，推送多元观点打破信息茧房。通过设计既能保障临时社交场景的隐私安全感的匿名制社交方式，又通过技术预警强化“匿名不等于免责”的伦理认知，最终实现从技术架构到青少年用户行为的双重伦理约束。青少年用户需通过实名认证获得匿名权限，平台同步记录数字行为轨迹，确保极端行为可定位至真实主体，直观呈现“小恶”如何通过技术机制演变为系统性伤害。

（三）法律规制与平台责任的协同强化

法律规制与平台责任的协同强化是破解青少年网络暴力追责难题的核心路径，立法层面，修订《网络安全法》增设算法伦理审查条款，要求平台提交算法伦理影响评估报告，将算法放大暴力纳入连带责任；细化《民法典》“AI伪造侵权”条款，设立青少年数字权益保护专项法庭，降低维权门槛。平台层面，升级青少年模式，实施匿名社交实名认证过滤、AIGC内容24小时审核及12小时快速响应机制。通过明确算法设计者、平台运营者与使用者的责任划分，形成“事前审查—事中处置—事后追责”的全链条闭环，有效回应道德责任模糊性困境，实现技术与责任的良性互动，为青少年网络暴力防治提供制度化解决方案^[10]。

（四）社会支持网络的多元协同与伦理培育

社会支持网络作为防治青少年网络暴力的基础性生态系统，需通过家庭、学校与社区的多元协同构建。家庭层面通过系统指导家长掌握沟通技巧避免教育焦虑转化为网络形象过度评判，并参与科普活动增进家庭成员对数字环境的理性认知；学校作为关键社会化场域需构建安全和谐的校园环境，将网络暴力防治纳入德育体系核心内容，成立师生共治的《校园网络行为公约》；社区层面依托青少年活动中心，配备专业社工与心理咨询师形成分级响应机制，对受害者提供创伤修复、认知行为疗法等危机干预，对施暴者开展“同理心训练营”重建数字权益尊重意识。

本研究立足数字文明时代青少年成长的现实困境，通过对算法助推语言暴力、数字身份物化符号暴力、匿名场景肆意施暴等典型场景的深度解剖，结合数字伦理导向的素养教育框架系统性搭建，不仅精准揭示了青少年网络暴力背后算法逻辑、商业逐利与伦理失范交织的技术伦理本质，更针对性提出“教育深化、技术重构、法律协同与社会支持”四大维度的实践路径。在实践落

地层面，该路径进一步细化为分层分类的阶梯式教育课程、嵌入伦理审查与冲突降温机制的算法重构设计、涵盖事前预防、事中干预、事后追责的法律责任闭环体系，以及整合家庭、学校、平台、监管机构的多元社会支持网络，这一四维协同模式通过明确各主体权责边界、强化跨领域联动，有效破解了网络暴力治理中常见的“责任分散效应”，为实践操作提供了具象化、可落地的行动方案。

未来，将紧密追踪 AI 生成内容、元宇宙等新兴技术引发的网

络暴力新形态与新特征，动态迭代素养教育模块的核心内容与教学方法；同时加强跨学科协作，结合心理学、社会学、法学等多领域理论，完善多方协同机制的运行规则与保障措施，并充分考量不同地区、不同群体青少年的差异化需求，推动防治体系的精准适配。最终目标是构建起兼具理论深度、实践效力与动态适应性的青少年网络暴力防治长效体系，为青少年营造安全、健康、包容的数字成长环境，助力数字空间公共伦理秩序的持续完善。

参考文献

-
- [1] 贾雯静. AI应用须有界 [J]. 瞭望, 2024(34): 43-43.
 - [2] 中国互联网信息中心. 第54次《中国互联网络发展状况统计报告》[R/OL]. (2024-08-29). <https://www.cnnic.cn/n4/2024/0829/c88-11065.html>.
 - [3] 燕道成, 李进怡. 崇德向善: 数字强国时代青少年网络行为的伦理精神 [J]. 福建江夏学院学报, 2023, 13(2): 1-10.
 - [4] 李莉. 网络治理视域下网络服务提供者版权过滤义务 [D]. 中南财经政法大学, 2023.