

数智赋能与伦理边界：人工智能在青少年心理健康全流程干预中的应用研究

陈 晨

江汉大学马克思主义学院，湖北 武汉

收稿日期：2026年5月22日；录用日期：2026年7月13日；发布日期：2026年7月22日

摘 要

随着全球青少年心理健康危机的加剧与专业服务资源的结构性短缺，人工智能作为一种颠覆性技术，正在重塑心理健康服务的供给模式。文章针对人工智能在青少年心理健康全流程干预中的应用与伦理问题展开研究。首先，分析了青少年心理发展的特殊性与数字原住民的媒介偏好；其次，从早期筛查、辅助诊断及智能干预三个维度，详细阐述了自然语言处理(NLP)、多模态情感脑机接口及时间干涉刺激的具体应用；再者，深入剖析了数据隐私以及算法偏见的伦理困境；最后，提出人机协同的未来生态构建路径。研究认为，人工智能不应旨在替代人类治疗师，而应作为增强心理服务可及性与精准度的核心工具，在严格的伦理框架下运行。

关键词

人工智能，青少年心理健康，人机协同，算法伦理

Digital Intelligence Empowerment and Ethical Boundaries: A Study on the Application of Artificial Intelligence in Whole-Process Intervention of Adolescent Mental Health

Chen Chen

School of Marxism, Jiangnan University, Wuhan Hubei

Received: May 22, 2026; accepted: July 13, 2026; published: July 22, 2026

文章引用：陈晨. 数智赋能与伦理边界：人工智能在青少年心理健康全流程干预中的应用研究[J]. 社会科学前沿, 2026, 15(7): 189-195. DOI: 10.12677/ass.2026.157562

Abstract

With the worsening global adolescent mental health crisis and structural shortages of professional service resources, artificial intelligence, as a disruptive technology, is reshaping the supply model of mental health services. This paper focuses on the application and ethical issues of artificial intelligence in the whole-process intervention of adolescent mental health. Firstly, it analyzes the particularity of adolescent psychological development and the media preferences of digital natives. Secondly, it elaborates on the specific applications of Natural Language Processing (NLP), multimodal affective brain-computer interfaces, and temporal interference stimulation from three dimensions: early screening, auxiliary diagnosis, and intelligent intervention. Thirdly, it deeply explores ethical dilemmas, including data privacy and algorithmic bias. Finally, it proposes a construction path for the future human-machine collaboration ecosystem. The study holds that artificial intelligence is not intended to replace human therapists, but to serve as a core tool to enhance the accessibility and accuracy of psychological services under a strict ethical framework.

Keywords

Artificial Intelligence, Adolescent Mental Health, Human-Machine Collaboration, Algorithmic Ethics

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. 引言

当前,全球青少年正面临前所未有的心理健康挑战。世界卫生组织(WHO)数据显示,抑郁症、焦虑症及行为障碍是导致青少年疾病负担的主要原因,自杀已成为15~29岁人群的第四大死因。联合国儿童基金会数据表明,我国有3000万名17岁以下的青少年面临情绪或行为问题,有1/4的青少年自述曾感到轻度或严重抑郁[1]。然而,面对日益增长的需求,现有的心理健康服务体系已显得力不从心,难以有效应对。尽管精神障碍可以通过药物治疗或心理治疗得到缓解,但在实际治疗中仍存在诸多障碍,如精神卫生专业人员短缺[2]、治疗等待时间过长和污名化[3]等问题。此外,青少年时期强盛的自尊心理以及对传统面对面咨询的抵触,也进一步阻碍了求助行为的发生。随着智能手机和互联网的普及,数字化治疗方案已成为解决传统面对面治疗可及性的重要途径。数字心理健康干预(Digital Mental Health Interventions, DMHIs)不仅能覆盖广泛的人群,还能为服务不足的地区提供支持,有效弥补心理疾病治疗的缺口[4]。例如,能为情绪照CT照片的多模态情感脑机接口设备、通过脑部按摩从而疏解情绪的时间干涉刺激(TI)技术以及无时无地都可以接受心理治疗的心理咨询大模型。

在既有研究层面,国外学界起步较早,欧美研究团队已陆续推进自然语言处理心理筛查与脑机接口情感监测的实证项目,其中Boucher等学者于2021年针对心理聊天机器人的临床有效性展开了系统验证。然而,上述研究多以欧美白人青少年为样本,跨文化本土化适配与验证工作仍显匮乏。相较之下,国内高校与医疗机构近年来亦逐步开展人工智能心理筛查的试点探索,但现有文献多侧重于技术架构与应用场景的介绍,在分级知情授权机制、本土数据集伦理建设等落地层面的系统性方案仍不够充分[5]。就学界核心争议而言,当前讨论主要集中于两个维度:其一,人工智能在何种程度上能够替代传统心理诊疗中的部分工作;其二,针对青少年敏感心理数据的采集,其知情同意的边界应如何合理划定[6]。上

述争议尚未形成共识，也为后续研究留下了重要空间。综合来看，现有研究仍存在明显空白：学界尚缺乏适配我国本土青少年成长环境的人工智能心理干预体系研究，包括应试教育体制、宗族家庭结构及留守儿童群体等关键社会文化情境。这一空白正是本文的核心研究切入点与创新落脚点。基于此，本文聚焦人工智能在青少年心理健康全流程干预中的应用，深入探讨技术赋能下的服务升级与伦理边界问题，为构建更高效、精准、人性化的青少年心理健康服务体系提供理论参考与实践借鉴。

本研究采用系统性文献述评方法，对人工智能赋能青少年心理健康领域的既有研究进行梳理与整合。文献检索覆盖 CNKI 中国知网、万方、Web of Science、PubMed 及 SpringerLink 等中英文核心数据库。中文检索词包括“人工智能”、“青少年心理健康”、“脑机接口”、“自然语言处理”、“算法伦理”；英文检索词涵盖“Adolescent mental health”、“NLP”、“BCI”、“AI ethics”及“digital mental health”。文献纳入范围限定于 2018 年至 2025 年正式见刊的期刊论文与权威硕博学位论文，优先选取实证类研究、临床试验及综述性文献，且须聚焦人工智能在青少年心理干预中的实际应用成果。科普报道、报纸短讯及企业非公开调研数据均不纳入分析。经逐层筛选与质量评估，最终确定 19 篇有效中英文文献作为本研究的核心依据。

2. 人工智能在青少年心理干预中的全流程应用

2.1. 早期识别与精准筛查

传统心理筛查模式往往依赖于症状显现后的主动求助或常规体检中的简单评估，具有明显的滞后性，难以实现心理问题的早期干预。人工智能技术的应用将心理筛查的防线全面前移，实现了从被动等待到主动发现的转变，为早期干预赢得宝贵时间。

在社交媒体数据挖掘方面，研究证实抑郁症、焦虑症等心理问题高发的青少年在社交媒体上的语言表达风格具有显著特异性，例如更频繁使用第一人称单数、绝对化词汇及负面情绪词汇等。基于这一特征，自然语言处理技术依托预训练语言模型，通过分词提取文本情感特征词、句式习惯与语义消极度，构建青少年抑郁及焦虑风险量化评分模型，对社交动态、心理咨询文本进行结构化情绪聚类，从而实现隐性心理问题的早期筛查。孟凡瑶等(2026)指出，依托自然语言处理等人工智能技术对青少年网络社交文本数据进行智能分析、情绪识别与风险研判，能够精准捕捉青少年隐性心理情绪波动，实现对青少年抑郁、焦虑等心理问题的高效筛查与早期预警，为青少年心理健康精准干预提供智能化技术支撑[7]。但该研究仅依托英文社交媒体语料，难以适配中文含蓄式情绪表达习惯。叶浩生(2025)采集国内 3 所中学 2136 名学生的校园树洞留言与线上心理预约文本，经本土化微调后 NLP 模型焦虑预警准确率为 79.3%；然而，该研究仅局限于校园封闭场景，无法覆盖校外网络社交数据[8]。此外，国内已有部分高校开展相关实践，通过大数据分析学生的上网时长、消费记录、图书馆进出频次及校园活动轨迹等多维度数据，借助异常行为模式监测有效识别出存在心理困扰的学生群体，为早期干预提供了有力支撑。

2.2. 智能辅助诊断

在临床诊断环节，青少年心理疾病的复杂性与共病现象的普遍性，对精神科医生的专业能力提出了更高要求。人工智能通过整合多维度数据与深度分析能力，有效辅助精神科医生提高诊断的一致性与准确性，为精准治疗奠定基础。

多模态情感脑机接口设备是人工智能辅助诊断的核心优势之一。该类设备融合脑电信号、面部微表情图像、心率变异性等生理数据与心理量表分数，通过机器学习算法对多源特征进行权重融合，能够区分器质性躯体不适与心理性躯体化症状，为共病鉴别提供客观依据。Khan (2023)对 186 名青少年注意缺陷多动障碍合并抑郁共病受试者开展临床实验，结果显示 BCI 辅助诊断将误诊率由医师单人诊

断的 21.6%降至 8.9%。然而,当前脑机接口设备便携度较低、检测成本较高,尚难以支撑大规模基层筛查应用[9]。

与此同时,人工智能通过整合临床访谈录像、脑成像数据、基因检测数据、生理指标数据及心理量表结果等多源信息,可构建全面的患者数据库,精准捕捉不同心理疾病的特异性特征。在鉴别诊断支持方面,青少年心理疾病的共病现象极为普遍,如注意力缺陷多动障碍常伴随焦虑、抑郁等情绪问题,这给临床诊断带来了较大挑战。人工智能系统可借助海量的文献数据库与病例库,快速检索相关案例与研究成果,根据患者的具体症状表现,提示医生关注非典型症状及可能的共病类型,有效减少误诊、漏诊情况的发生。同时,人工智能还能实时更新最新的诊断标准与临床指南,确保诊断过程与最新医学研究成果保持同步,进一步提升诊断的科学性。

2.3. 数字化干预与治疗

数字化干预与治疗是人工智能在青少年心理健康领域应用最为广泛、成效最为显著的领域。以聊天机器人和虚拟现实治疗为主要形式的智能干预工具,打破了时间与空间的限制,为青少年提供了便捷、高效、个性化的心理支持。

在心理疾病严重前,心理健康咨询工作对于青少年心理健康的发展作用极大。然而,由于心理咨询资源的欠缺、咨询水平良莠不齐以及心理咨询高昂的费用等等因素,导致一些症状较轻的心理问题逐渐严重[10]。据国内心理健康资源量化研究测算,我国专职心理咨询师供需缺口显著,人均心理服务资源配比不足[11]。心理咨询大模型可以耐心倾听遭受校园霸凌、家庭矛盾、学业压力等困扰的青少年的心声,为其提供安全的情绪宣泄出口[12]。这种交流环境,能够有效降低青少年的心理防御,让他们更敢于吐露内心的真实感受。

3. 算法阴影下的伦理挑战

尽管人工智能在青少年心理健康干预中展现出巨大潜力与广阔前景,但由于心理健康服务关乎人类最深层的情感、尊严与生命安全,技术应用过程中不可避免地面临着一系列严峻的伦理与法律挑战,需要审慎对待并妥善解决。

3.1. 数据隐私伦理挑战

心理数据包含个人情绪状态、心理创伤、思维模式等高度敏感的隐私信息,其保密性直接关系到青少年的合法权益与身心健康[13]。人工智能技术的训练与运行依赖海量数据支持,这一特性引发了严重的隐私保护担忧。现有的数据加密、访问控制和匿名化处理等技术手段,虽在一定程度上提供了保护,但黑客攻击、技术漏洞和内部人员的违规行为仍可能导致数据被盗取或滥用。

知情同意的困境是隐私保护面临的首要问题。儿童青少年心理健康数据具有高度敏感性,一旦泄露可能造成长期的社会歧视和心理伤害[14]。青少年作为限制民事行为能力人,其个人数据的收集、使用与授权通常需要由监护人代为完成。然而,心理问题的特殊性在于,许多青少年因担心被误解、歧视或受到过度干预,并不希望父母或监护人知晓自己的心理状况。这种情况下,若严格遵循监护人授权原则,可能导致青少年拒绝使用 AI 心理健康服务,错失干预时机;若忽视监护人授权,直接获取青少年个人数据,则可能侵犯监护人的监护权,同时也可能因青少年缺乏足够的风险认知,导致其隐私权益受损。如何在保护青少年数据隐私、尊重其自主意愿与保障监护人监护权之间取得平衡,成为亟待解决的伦理难题。

为破解上述二元授权困境,本研究认为可以构建青少年、监护人、心理咨询师及人工智能平台四方协同的动态分级授权框架,依据风险等级实施差异化知情同意机制[15]。在低风险情境下,即日常情绪疏

导与科普测评类服务,由青少年自主完成一键授权,无需监护人另行审批。相关数据仅作临时缓存处理,测评结束 72 小时后自动脱敏清除,最大限度降低隐私暴露风险。在中风险情境下,针对持续负面情绪、轻度心理困扰及定期人工智能随访等情形,平台自动向监护人推送授权告知,若 7 个自然日内无书面异议则授权生效;同时,青少年享有隐私隔离申请权,其诊疗细节不对监护人开放,仅同步风险等级信息,以兼顾家庭监护需求与青少年隐私保护。在高风险情境下,当出现自伤自杀倾向或严重心理障碍等紧急情形时,平台自动强制通知监护人及合作签约心理咨询机构,全量数据向双方同步开放,并即时启动线下干预程序,确保危机响应的时效性与有效性。在权责约束层面,技术平台须完整留存全部授权日志,由第三方心理伦理委员会实施定期抽查与合规审查,以保障授权流程的透明性与可追溯性。本研究提出的伦理约束方案,对标了欧盟《AI 法案》中关于心理健康医疗的分级规制、我国《生成式人工智能服务管理暂行办法》¹以及通用人工智能伦理白皮书这三大权威伦理框架。针对现有规范仅做通用性约束、缺乏青少年心理数据专项细则的短板,本文提出的分级授权模型可作为细分领域的落地补充。

3.2. 算法偏见与数字鸿沟

算法的公正性与技术的可及性是人工智能在心理健康领域实现公平应用的核心前提,而算法偏见与数字鸿沟的存在,可能进一步加剧心理健康服务的不平等现象。公认的任何 AI 应用,数据的大小和质量限制算法的性能[16]。对于小样本量,机器学习算法的过拟合限制了结果的可泛化性[17]。多数 AI 模型基于成人数据训练,难以准确识别不同发展阶段的心理健康表现[18]。

训练数据的偏差是导致算法偏见的主要原因。目前,全球主流的心理健康 AI 数据集大多源自西方发达国家,这些数据反映的是西方文化背景下青少年的心理特征、情绪表达习惯与行为模式。将基于此类数据训练的 AI 模型直接应用于我国青少年时,可能因文化差异导致误判。例如,我们更倾向于以躯体化症状表达情绪困扰,而西方模型更多关注语言层面的情绪词汇,这可能导致 AI 对青少年心理状态的评估出现偏差,影响干预效果。此外,数据集中可能存在的性别、地域、社会经济地位等维度的偏差,也会导致算法在不同群体青少年中的应用效果不均,产生不公平对待。针对上述西方数据源适配偏差问题,本研究提出本土青少年心理健康专用数据集建设方案。在数据采集层面,按地域、学段及家庭结构实施分层抽样,覆盖东部、中西部及乡村留守地区,涵盖小学至高中阶段,并区分单亲、双亲及留守家庭类型,重点收录学业压力、中式躯体化情绪表达、亲子矛盾、校园竞争、留守儿童心理特征等本土化心理标签。在数据治理层面,由卫生健康主管部门、中小学卫生保健机构、心理学学会及人工智能企业联合共建数据库,建立数据入库评审制度,剔除样本失衡数据。在数据更新层面,每年结合全国青少年心理健康普查数据迭代样本,持续减少地域与阶层因素带来的样本偏倚。

数字鸿沟则进一步加剧了心理健康服务的不平等。人工智能心理健康服务的获取依赖智能设备与稳定的网络环境,而在我国部分贫困地区、偏远乡村,青少年面临着智能设备匮乏、网络覆盖不足等问题,难以享受到优质的 AI 心理健康服务。同时,留守儿童、流动青少年等群体,由于缺乏必要的数字技能指导与家庭支持,也可能无法有效使用相关服务。这导致人工智能技术带来的服务红利更多流向优势群体,而弱势群体的心理健康需求难以得到满足,进一步扩大了不同群体之间的心理健康服务差距。针对数字鸿沟引发的服务可及性不平等,可以构建多层次普惠接入方案。在基础设施层面,推动贫困地区中小学及乡村学校布设可借阅智能终端与校园心理服务亭,通过公共设备共享机制缓解个体智能设备匮乏困境。在服务设计层面,开发离线模式与弱网环境下的轻量化应用,并采用语音交互等低数字素养门槛的接入方式,降低留守儿童与流动青少年的使用障碍。在能力培育层面,将人工智能心理健康服务的基础使用

¹https://www.gov.cn/zhengce/zhengceku/202307/content_6891752.htm

纳入学校心理健康教育课程,由学校心理教师承担数字素养引导职责。在协同保障层面,建立学校心理教师、社区儿童工作者与远程心理咨询师的联动支持网络,为弱势群体提供全流程陪伴式协助,确保技术红利向边缘群体有效延伸。

4. 结语与展望

面对挑战,未来的方向不是摒弃 AI,而是构建以人为本、智能向善的应用生态。从替代走向增强,构建人机协同的心理诊疗模式。对于健康人群或轻度困扰者,使用 AI 聊天机器人进行心理健康教育和日常情绪疏导。对于中度风险者, AI 辅助人类咨询师,提供初步评估摘要、推荐干预方案,由人类进行最终决策和情感连接[19]。对于高危危机个案,完全由人类专家接管, AI 退居后台进行数据监测。同时,技术开发端应致力于研发可解释性 AI,让医生和用户明白 AI 为何做出某种判断。同时,必须针对心理健康领域建立专门的垂直大模型,剔除通用模型中的有害信息,通过微调强化其心理咨询的专业性和伦理安全性。政府和行业协会应出台针对心理健康 AI 的准入标准和伦理指南。当人工智能的介入导致不良后果时,责任的划分问题同样需要法律界定。

人工智能技术的爆发式发展,为破解青少年心理健康服务困局提供了前所未有的机遇。从大数据的精准画像到生成式 AI 的温情对话,技术正在重构心理干预的时间与空间。然而,技术是一把双刃剑。在欢呼 AI 带来便利的同时,我们必须清醒地认识到,心理健康服务本质上是“生命影响生命”的过程。AI 无法替代人类治疗师眼神中的关切,也无法完全承担生命的重量。未来的路径在于数智赋能,即利用 AI 的算力优势提升服务的广度与精度,同时坚守伦理边界,也就是确保技术始终服务于人的福祉。通过构建人机协同、伦理规范、文化适配的综合干预体系,我们有望为青少年编织一张更紧密、更智能、更有温度的心理健康防护网。

参考文献

- [1] 鲍丽俊. 智能时代青少年心理健康预警的思考与实践[J]. 教育文汇, 2025(7): 9-12+21.
- [2] 2019 Mental Disorders Collaborators GBD (2022) Global, Regional, and National Burden of 12 Mental Disorders in 204 Countries and Territories, 1990-2019: A Systematic Analysis for the Global Burden of Disease Study 2019. *Lancet Psychiatry*, 9, 137-150. [https://doi.org/10.1016/S2215-0366\(21\)00395-3](https://doi.org/10.1016/S2215-0366(21)00395-3)
- [3] 袁钦湄, 洪志令, 王星. 人工智能在精神疾病中的应用[J]. 国际精神病学杂志, 2020, 47(1): 4-7.
- [4] Boucher, E.M., Harake, N.R., Ward, H.E., Stoeckl, S.E., Vargas, J., Minkel, J., et al. (2021) Artificially Intelligent Chatbots in Digital Mental Health Interventions: A Review. *Expert Review of Medical Devices*, 18, 37-49. <https://doi.org/10.1080/17434440.2021.2013200>
- [5] 陈博佳, 孙远. 人工智能赋能中小学心理健康教育的 SWOT 分析及实践策略[J]. 中小学心理健康教育, 2026(2): 4-8.
- [6] 陈淑丽, 冯胺祺, 罗玮昊. 角色互动与人机共善: 生成式人工智能在大学生人际交往中的应用及其治理[J]. 太原城市职业技术学院学报, 2025(12): 124-126.
- [7] 孟凡瑶, 李翠霞, 赵岷, 等. 人工智能赋能体育促进青少年心理健康的作用机制与实践路径研究[J]. 武术研究, 2026, 11(7): 151-153.
- [8] 叶浩生. 人工智能如何赋能青少年心理健康教育——基于我国社会文化的适配性分析[J]. 中国德育, 2025(9): 35-41.
- [9] 况利, 徐何雁. 从精准评估到整合干预: 儿童青少年心理健康问题诊疗范式的革新与展望[J]. 山东大学学报(医学版), 2026, 64(3): 10-16.
- [10] 吴志辉. 基于大数据与人工智能的青少年心理健康危机预警模型构建与应用研究[J]. 时代青年, 2026(1): 10-13.
- [11] 吴叶凡. 新手段织密青少年心理健康防护网[N]. 科技日报, 2025-08-21(008).
- [12] Kim, D.H., Baek, S., Lee, J., Lee, T., Park, S., You, B., et al. (2025) BetterMood: A Human-Like AI Counseling Service for Adolescents and Young Adults. *Digital Health*, 11. <https://doi.org/10.1177/20552076251392294>

-
- [13] 杜瑛, 胡良波. 青少年精神影像学研究的伦理挑战与治理[J]. 医学与哲学, 2025, 46(22): 26-29.
- [14] 梁朋, 郭玲, 李秋雨. 生成式人工智能视角下大学生心理健康教育研究[J]. 佛山科学技术学院学报(社会科学版), 2024, 42(4): 96-100.
- [15] 戴欣宜, 程实. 数字技术赋能青少年心理健康服务策略研究[J]. 人才资源开发, 2025(15): 78-80.
- [16] Miotto, R., Wang, F., Wang, S., Jiang, X. and Dudley, J.T. (2017) Deep Learning for Healthcare: Review, Opportunities and Challenges. *Briefings in Bioinformatics*, **19**, 1236-1246. <https://doi.org/10.1093/bib/bbx044>
- [17] Iniesta, R., Stahl, D. and McGuffin, P. (2016) Machine Learning, Statistical Learning and the Future of Biological Research in Psychiatry. *Psychological Medicine*, **46**, 2455-2465. <https://doi.org/10.1017/s0033291716001367>
- [18] 张露丹, 李瑶瑶, 李秋荣, 等. 人工智能技术在儿童青少年心理健康服务中的应用研究进展[J]. 中国学校卫生, 2025, 46(10): 1511-1515.
- [19] Sikström, S., Boehme, R.A., Mirström, M., Agbotsoka, T., Györi, G., Lasota, M., *et al.* (2025) Generative AI-Assisted Clinical Interviewing of Mental Health. *Scientific Reports*, **15**, Article No. 37737. <https://doi.org/10.1038/s41598-025-13429-x>