

生成式人工智能赋能高校心理健康服务的兴起、挑战与规范化发展路径

冀艳云, 刘佳音, 陈佳宜, 甘晓芳*

内蒙古医科大学卫生管理学院, 内蒙古 呼和浩特

收稿日期: 2026年9月5日; 录用日期: 2026年10月12日; 发布日期: 2026年10月21日

摘要

高校学生心理健康问题日益严峻, 传统心理健康服务体系在专业资源、服务可及性、学生病耻感等方面面临多重困境。生成式人工智能(Generative Artificial Intelligence, GenAI)技术的迅猛发展为破解上述难题提供了新的技术路径。现有研究表明, 人工智能聊天机器人在心理健康科普、情绪支持、自我调节和部分数字化干预任务中具有扩大服务可及性、降低使用门槛和增强个性化交互的潜力, 但其实际应用仍面临专业边界模糊、数据隐私、风险识别、情感依赖及责任界定难等问题。本文基于国内外相关研究, 对生成式人工智能赋能高校心理健康服务的兴起、技术演进和应用现状进行梳理, 并进一步讨论提示工程与指令微调、检索增强生成(RAG)、情感计算和多角色协作架构等关键技术机制。在此基础上从功能分层、专业知识约束、安全评估、数据治理、人机协同和校园生态融合以及中国高校文化适应等方面提出规范化发展路径。研究认为, 高校生成式人工智能心理健康服务的定位不应替代专业人员, 而是在以人为中心和专业人员监督的前提下, 承担低风险、高频次、即时性的心理支持与情绪疏解任务, 并通过有效转介机制与高校现有心理健康服务体系形成互补, 从而实现技术效率与心理服务专业性的平衡。

关键词

生成式人工智能, 大语言模型, 高校心理健康, 人机协同, 人工智能伦理

The Rise, Challenges, and Path to Standardized Development of Generative AI-Empowered Mental Health Services in Higher Education

Yanyun Ji, Jiayin Liu, Jiayi Chen, Xiaofang Gan*

School of Health Management, Inner Mongolia Medical University, Hohhot Inner Mongolia

*通讯作者。

文章引用: 冀艳云, 刘佳音, 陈佳宜, 甘晓芳(2026). 生成式人工智能赋能高校心理健康服务的兴起、挑战与规范化发展路径. *心理学进展*, 16(10), 223-231. DOI: 10.12677/ap.2026.1610487

Abstract

Mental health issues among university students are becoming increasingly severe, and traditional mental health service systems face multiple challenges regarding professional resources, service accessibility, and student stigma. The rapid development of Generative Artificial Intelligence (GenAI) offers a new technological pathway to address these difficulties. Existing research indicates that AI chat bots have the potential to expand service accessibility, lower barriers to use, and enhance personalized interaction in areas such as mental health education, emotional support, self-regulation, and certain digital intervention tasks; however, their practical application faces issues such as blurred professional boundaries, data privacy concerns, risk identification challenges, emotional dependency, and difficulties in defining liability. Drawing on relevant domestic and international research, this paper reviews the emergence, technological evolution, and current application status of GenAI-enabled mental health services in universities. It further discusses key technical mechanisms, including prompt engineering and instruction fine-tuning, Retrieval-Augmented Generation (RAG), affective computing, and multi-role collaborative architectures. Based on this analysis, the paper proposes a pathway for standardized development, covering aspects such as functional stratification, professional knowledge constraints, safety assessment, data governance, human-AI collaboration, integration into the campus ecosystem, and cultural adaptation within Chinese universities. The study concludes that GenAI-based mental health services in universities should not replace professionals; instead, operating under a human-centric approach and professional supervision, they should undertake low-risk, high-frequency, and immediate tasks related to psychological support and emotional relief. By establishing effective referral mechanisms, these services can complement existing university mental health systems, thereby achieving a balance between technological efficiency and the professionalism of mental health services.

Keywords

Generative Artificial Intelligence, Large Language Models, Mental Health in Colleges and Universities, Human-Machine Collaboration, Artificial Intelligence 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

心理健康是高校人才培养体系中不可忽视的一环，也是学生个体成长发展的重要支撑。近年，大学生在学业、人际关系、就业和社会适应等方面面临着多重压力，高校心理健康工作也逐渐由应对传统的心理问题转向涵盖健康教育、预防监测、心理咨询和危机干预等的综合服务体系。2023年，教育部等十七部门联合发布了《全面加强和改进新时代学生心理健康工作专项行动计划（2023—2025年）》¹，提出要坚持健康第一的教育理念，切实把心理健康工作摆在更加突出位置，通过构建“健康教育、监测预警、咨询服务、干预处置”四位一体的工作体系，不断提升学生心理健康工作的科学性、针对性和有效性，最终实现学生心理健康素养的全面提升。然而，高校心理服务仅仅依靠有限数量的班主任、辅导员和心

¹https://www.gov.cn/zhengce/zhengceku/202305/content_6857361.htm

理咨询人员,很难覆盖所有学生在不同时间、不同情境下产生的大量心理支持需求。隐私泄露的担忧、被贴上标签的恐惧、时间安排上的冲突等因素,也让一部分学生在遇到心理困扰时选择沉默。如何在守住专业边界的前提下,为数字化的心理服务打开一个安全有效的入口是当下关键的问题。2025年4月,《教育部等九部门关于加快推进教育数字化的意见》²明确提出要推动心理健康等领域的大模型垂直应用(吴砥等, 2026)。紧接着,2026年出台的“十五五”规划进一步强调要健全社会心理服务体系和危机干预机制³。几份政策文件的发布为生成式人工智能(英文全称 GenAI)赋能高校心理健康服务指明了方向。以 ChatGPT、DeepSeek 等为代表的 GenAI 技术正在解决传统高校心理健康服务模式面临的多重挑战。随着大语言模型(英文全称 LLMs)的成熟,这类系统开始具备更强的语义理解、多轮交互和开放式文本生成能力。它们不仅能够回答心理健康问题,还能够结合历史对话,借助知识库、工具调用等方式完成更加复杂的交互任务。然而,LLMs 虽表现出一定潜力,但大模型的运用依然存在信息不一致、隐私风险、可解释性不足和伦理规范缺失等问题,现阶段仍不能将其视为专业心理健康服务的替代品(Guo et al., 2024)。因此,对高校情境下的人工智能,需要进一步回答三个问题:生成式人工智能在高校心理健康服务体系中能承担哪些任务?哪些任务不应由人工智能独立承担?如何将技术能力转化为安全、可靠和可持续的心理健康服务。本文围绕这些问题,对国内外研究进行梳理,并尝试从技术能力、服务边界和制度治理三个维度提出一条可行的规范化发展路径。

2. GenAI 赋能高校心理健康服务的兴起

2.1. 从国家战略到教育实践的政策驱动

GenAI 赋能高校心理健康服务的兴起,得益于国家政策的推动。2023 年学生心理健康专项行动计划提出汇聚心理科学、人工智能等资源开展研究;2025 年教育数字化政策进一步提出推动心理健康领域大模型应用;2026 年“十五五”规划则强调健全社会心理服务和危机干预机制。上述政策为 GenAI 在高校心理健康领域的应用提供了制度基础。这些政策制度的部署不但为技术创新与应用落地提供了制度保障,也明确指出 GenAI 要从国家战略高度落实到教育实践中去。

2.2. 高校心理健康服务现实困境的需求牵引

政策驱动的背后是复杂的现实需求。传统高校心理健康服务在以下方面存在明显不足:高校心理咨询中心普遍面临专业师资短缺、服务能力有限的困境;传统的面对面咨询模式受限于时间和空间条件,无法覆盖学生在深夜或假期等时段的心理需求;大量学生因担心被另眼看待而回避主动寻求心理帮助,导致心理问题被延迟发现与干预;早期数字心理健康系统主要依赖规则式聊天机器人,通过预先设计的问答流程进行交互,当用户叙述复杂或同时呈现多种情绪时,系统会难以理解导致情绪支持不及时。这些困境共同构成了 GenAI 进入高校心理健康服务领域的现实基础。

3. GenAI 介入心理健康服务的技术演进

3.1. 从规则式对话系统到开放式生成

人工智能心理服务并非随着 LLMs 的出现开始发展。以 Woebot 等早期对话系统为代表的研究表明,通过对话形式提供心理健康相关内容具有一定可行性,但它在开放式交互理解方面依然受到限制(Guñu et al., 2021; Yeh et al., 2025)。LLMs 改变了这一技术基础,它会通过大量交互训练获得自然语言生成和建模的能力。在此基础上,提示工程与指令微调成为通用模型向心理健康领域适配的重要技术。提示工程通

²https://www.gov.cn/zhengce/zhengceku/202504/content_7019045.htm

³https://www.gov.cn/zhengce/zhengceku/202607/content_7075214.htm

过系统提示、角色设定和任务链的方式,规定模型交互步骤和功能边界;而指令微调是用整理好的问答、对话或任务指令进一步训练模型,以达到遵循特殊任务的要求。因此,技术演进正由扩大规模逐步转向通过提示工程、指令微调的方式提高模型在特定心理服务场景中的可控性与一致性。

3.2. 从通用语言能力向领域知识增强发展

通用 LLMs 虽然具有良好的语言表达能力,但表达流畅并不意味着专业可靠。如果完全依赖通用模型,容易产生事实错误、概念混淆和缺乏依据的建议。检索增强生成(Retrieval-Augmented Generation, RAG)为上述情况提供了一种可控的技术路径。它首先对提出的问题进行语义表示,然后在专业知识库中检索相关材料,最后将检索到的专业内容与当前问题共同输入 LLMs 完成生成,形成了问题识别-知识检索-上下文增强-回答生成的链条。有研究表明,RAG 已被较多 LLMs 心理咨询系统用于引入专业心理健康知识并支持上下文生成(Cho et al., 2026)。但另一项研究发现,优化模型后,它的规范性、情感理解以及回答一致性都能得到改善,但专业能力提升有限(黄峰等, 2025)。由此来看,较稳妥的系统不应该把知识生成与风险决策混在一起,而是应努力实现 AI 系统与人工服务的有效衔接。

3.3. GenAI 在心理健康领域的突破发展

GenAI 之所以能够进入心理健康服务领域,情感计算是重要的技术路径之一。它会从文本中的情绪词、语义强度、上下文表达变化中提取情感信息,将其分解为离散情绪类别、效价-唤醒度等连续维度,最后根据结果调整回应语气和服务路径。除此之外,多模态情感计算还可以结合语音和行为等信号进一步作出判断。相关综述指出,当前情感计算正由传统任务分类向 LLMs 驱动的情感理解与情感生成发展,但模型仍受到训练数据偏差、文化差异以及多模态信息一致性等因素影响(Schlicher et al., 2025)。因此,在高校中,情感计算更适合作为辅助,而不能把模型识别出的情绪类别直接等同于心理状态的直接诊断。

3.4. 从单一对话模型向多角色协作架构发展

随着对话任务复杂度的提升,依赖 LLMs 同时完成多项系统任务容易造成冲突,多角色协作架构因此成为被关注的重要技术方向。该架构将心理服务流程拆分为很多功能相对独立的智能体或模块,协作完成整个对话任务。与单模型相比,这种架构的核心是利用任务分工和交叉检验的步骤来降低单一模型同时承担多个任务所带来的风险。2026 年关于 LLMs 心理咨询与 GenAI 的研究已经将代理式系统架构列为该领域的重要发展方向,但同时指出外部验证、真实环境验证和伦理安全报告仍然不足(Cho et al., 2026)。因此,多角色协作目前更适合作为提高系统可控性的工程框架,它的实际心理健康效能需要进一步验证。

4. GenAI 赋能高校心理健康服务的应用现状

4.1. 心理健康知识教育与信息支持

心理健康知识教育是 GenAI 风险相对较低的应用领域之一。高校学生可能在压力、焦虑、睡眠、情绪变化及人际关系等方面需求巨大,传统心理健康网站和宣传材料需要依靠学生主动查找,信息不透明的情况下学生的心理需求无法得到及时干预,而 GenAI 可以打破限制,降低信息获取门槛,为学生提供较为即时的心理健康信息支持和低强度心理支持。与普通搜索不同,GenAI 可以根据用户的提问方式调整解释深度,将专业概念转化为更加易于理解的语言。已有研究开始探索人工智能在大学生心理健康评估中的应用(袁小雅, 刘仪辉, 2025)。然而,智能心理评估的信效度边界还需要进一步验证。

4.2. 日常情绪支持与低强度心理调适

对于考试压力、人际冲突、就业压力或短期负面情绪等情境,GenAI 可以通过语言回应和结构化建

议提供支持。它的核心价值在于为学生提供即时的自我调适入口。2025 年针对大学生人工智能聊天机器人的系统研究中包含了 9 项子研究及 1082 名参与者,其中多数研究报告了焦虑、抑郁或心理健康等指标的改善,但同时存在异质性较高、部分研究样本量较小、退出率较高等局限(Nyakhar & Wang, 2025)。一项针对青少年和青年群体的系统综述纳入了 31 项随机对照试验,结果表明人工智能聊天机器人对心理困扰有不同程度的改善效应(Feng et al., 2025)。这些结果表明, GenAI 不仅能够提供基础的情感陪伴,还可以实现更具针对性的心理干预,但其稳定性还需要验证。

4.3. 情绪识别与危机预警

情绪识别与危机预警是 GenAI 在高校心理健康服务中具有独特价值的功能维度。通过与用户的多轮互动进行分析并识别潜在的风险。部分高校开始探索将 AI 危机评估功能纳入心理服务体系,比如中山大学“逸心”模型提供的“4P 危机评估”功能等⁴。然而,研究也指出, GenAI 聊天机器人在风险监测与评估方面的表现都不理想。多数 GenAI 聊天机器人在治疗方法的适切性(仅 31%达标)和风险监测能力方面表现较差(Sobowale et al., 2025)。这一发现凸显了当前技术在危机情境下的局限性。

4.4. 当前应用实证及其局限性

关于人工智能心理健康服务的研究快速增加,但现有研究使用的 AI 系统技术路线差异较大,不能将规则系统、传统机器学习模型与生成式 LLMs 的效果直接等同。已有随机对照研究显示,优化后的 LLMs 可在短期内改善部分心理指标,但效果维持和专业能力提升仍有限(黄峰等, 2025)。因此,高校在引入 GenAI 心理健康服务时必须区分四个层次:系统是否能生成合理回答;用户是否愿意使用;系统是否能够长期稳定运行;能否真正改善心理健康结局,最后一点也是 GenAI 的效能所在。现有研究容易将前三者直接解释为第四者,这是需要纠正的一点。

5. GenAI 心理健康服务的主要风险

5.1. 语言能力与专业能力错位

GenAI 最容易造成误解的地方,是语言表达能力非常接近人类的交流方式。然而, LLMs 本质上依据输入上下文和训练所得参数进行文本生成,而不是依照心理咨询师或医疗专业人员的专业判断。因此,高校 AI 心理健康应用在功能设计方面应当始终明确基础心理教育、一般支持与专业咨询、诊断治疗三者之间的边界。

5.2. 模型依赖与信息可靠性风险

World Health Organization (WHO, 2024)发布的文件指出,大模型可能生成错误、不完整、有偏见但语言形式非常具有说服力的信息,这种特点在健康相关场景有很大的风险。如果模型在信息不足时随意下结论,就可能将普通生活困扰错误解释为心理疾病,也可能对真正需要专业介入的问题作出简单的回应。因此,模型设计应避免确定性诊断回复,不能为了追求表面上的专业性而生成未经验证的结论。

5.3. 数据隐私与信息安全风险

心理健康对话属于高度敏感的隐私信息,其内容可能涉及家庭、个人经历等其他私人信息。GenAI 大多都依赖云端模型接口进行数据处理,因此数据治理的风险远高于普通智能问答。现有研究已将数据留存、第三方服务调用及敏感信息暴露等列为生成式心理健康聊天机器人的重要隐私与网络安全问题

⁴<https://psycorner.sysu.edu.cn/mhecc/trends/d78dd1e0f26e49d8a13f33289459486f.htm#>

(Sawesi et al., 2026)。另外,当 AI 心理支持系统给出不当建议或未能识别危机信号时,责任由谁承担?现有的监管框架对这一问题的回答并不明晰。

5.4. 情感依赖与人机边界模糊

LLMs 能够持续表达理解,增强与用户的共情能力,但这种高度拟人化的交互方式,也可能使学生在长期使用中逐渐形成对 AI 助手的情感依赖,把本应投向真实人际关系的信任与依恋转移到机器身上,进而模糊人机之间的本质边界。因此,人机交互的设计应该既有温度也有边界,避免某种意义上的情感绑定。真正有效的心理支持应当指向用户的现实适应与自主成长,因此更不能用延长用户的互动时间作为心理健康服务质量的主要评价指标。

5.5. 应用落地障碍与体系融合困境

当前,尽管学生对 AI 心理支持工具普遍持开放态度,但信任的建立需要时间。最需要心理支持的学生往往也是最不愿主动求助的群体。如何设计有效的推广与引导机制,触及这些真正有需求的潜在用户,是 AI 心理支持系统面临的重要挑战。需要注意的一点是,它不能取代人类咨询师的专业判断,所以在技术效率和人文关怀之间找到平衡点也是需要持续探索的课题。

6. 生成式人工智能赋能高校心理健康服务的规范化发展路径

6.1. 由技术驱动转向需求与边界驱动

生成式人工智能最有价值的应用场景往往不是处理最复杂、最疑难的心理干预问题,而是兼顾传统的心理健康服务体系难以覆盖的大规模低风险需求,比如心理健康知识解释、日常压力管理、即时的情绪疏导等。只有明确任务边界、厘清哪些环节适合由 AI 承担、哪些必须交由人类专业人员处理,再据此选择相应的技术方案,才能防止为了用技术而用技术。

6.2. 建立专业知识约束与多层安全机制

高校心理健康 GenAI 应将提示工程与指令微调、RAG、风险识别和输出审核等机制嵌入多层安全护栏(Safety Guardrails)。输入层可利用情感与风险识别模块将对话场景分类;知识层通过 RAG 调用经过专业审核的知识库;生成层利用系统提示、任务指令和白名单功能约束模型的回答范围;输出层由独立的审核功能进行安全检查,如事实一致性、过度确定性表达或不恰当的人机关系表述;达到风险阈值后,则停止自主处理并进入人工服务。已有研究显示,心理健康 LLMs 系统已开始组合提示工程、指令微调、RAG 和代理式架构,但安全机制和真实环境验证仍缺乏统一标准(Cho et al., 2026; Olisaeloka et al., 2026)。因此,高校在上线前还应通过多轮测试,评估系统运行的各个环节。

6.3. 加强数据全生命周期治理

对高校 GenAI 心理健康服务而言,数据治理应遵循目的明确和最小必要原则。心理对话中涉及医疗健康等敏感个人信息的,属于《中华人民共和国个人信息保护法》规定的敏感个人信息,应依法采取严格保护措施,在适用情形下取得单独同意,并明确告知处理目的、方式和保存期限。对于敏感个人信息处理、自动化风险分级和第三方数据处理等情形,还应开展个人信息保护影响评估,并按照岗位职责设置访问权限(徐天然,王慧,2024)。因此,高校应建立覆盖数据采集、存储、调用和销毁的全生命周期管理制度,不得将学生完整心理对话用于未经授权的模型训练或其他用途。即使采用第三方大模型,高校也不能将治理责任转移给技术供应商(Holmes & Miao, 2023)。从《生成式人工智能服务管理暂行办法》的视角看,大模型时代生成式人工智能治理应逐步探索既有制度框架的优化方案(秦瑞标,2024)。基于此,

“GenAI4MH”框架提出的数据安全、信息公正、用户安全和伦理治理四个维度,可为高校进一步完善治理机制提供参考(Wang et al., 2025)。

6.4. 从人工智能替代转向人机协同

未来高校心理健康服务的方向不是比较 AI 和心理咨询师谁更好,而是优化两者分工。AI 可以扩大心理服务的广度,专业人员可以保证服务的深度;AI 提供数字入口,专业人员承担专业责任。这一方向也符合现有研究的结论。尽管 LLMs 心理咨询系统表现出巨大的潜力,但真实环境下测评验证数据不足、伦理机制报告不完整等问题说明当前阶段依旧需要专业人员进行把控和监督。在这一领域中,南京财经大学的“AI 辅助 + 人工守护”双轨服务体系⁵,中山大学“逸心”模型的“师 - 生 - 机”三方协同模式⁶都是这个方向的积极探索。

6.5. 将 AI 纳入高校现有心理健康服务生态

真正成熟的高校心理健康 AI 应该与现有心理健康教育、预约咨询和学生工作形成连接。教育部等十七部门提出,高校应完善“学校 - 院系 - 班级 - 宿舍/个人”四级预警网络,并畅通预防、转介、干预和就医通道⁷。在这一框架下,GenAI 更适合承担前端入口和流程连接的功能,既能在日常阶段提供心理知识、压力管理和情绪疏导的自助支持,也能在心理普查后为低风险学生提供解释,当对话出现风险信号时,由系统触发转为人工处理,由经过授权的心理专业人员进行后续判断。辅导员可告知学生可用的心理资源、提供日常关怀并承担组织协调的角色,但不宜通过后台无限制查看学生完整心理对话。因此,高校心理健康 AI 最终应该关注三个问题:学生是否获得了准确的心理健康信息;需要进一步疏导的学生是否更容易获取专业资源;人工智能的加入是否提高了高校整体心理健康服务体系的可及性与安全性。

6.6. 中国高校情境下的 GenAI 治理

在中国高校情境中,学生对心理问题的表达会受到病耻感、社会评价以及隐私暴露顾虑的影响。研究显示,匿名性和可及性能够降低大学生数字心理资源的使用门槛,但污名以及对工具可信度的担忧仍会影响他们主动求助(Ma et al., 2026)。与面对面咨询相比,GenAI 匿名化、即时性和低压力的特点,可能降低部分学生表达心理困扰的门槛。所以,GenAI 更适合作为正式心理服务之前的低门槛入口,而不是专业咨询的替代渠道。然而,这种特点也具有双重效应,如果学生担心对话内容被学校、平台或第三方获取,其优势也可能被削弱。因此,针对 GenAI 的本土化设计而言,应该同时处理语言、关系和制度三个层面。语言层面,知识库和对话模板应该覆盖中国高校常见的学业竞争、就业压力、宿舍关系、家庭期待、恋爱与社会适应等情境,避免直接生成专业的诊断术语和表达;关系层面,明确 AI 的非人类身份,采用支持但不过界的表达,并在识别到重要风险时选择人工帮助;制度层面,要把 AI 接入学校已有的心理健康系统,按照岗位职责和最小必要原则明确数据访问权限。另外,还需要区分普通心理健康教育和个体风险处置,实现系统与人工的双向转换。

7. 讨论

GenAI 正在推动高校心理健康服务提供了更加即时、交互和情境化的数字支持。LLMs 在自然语言理解、多轮对话以及低强度情绪支持等方面已经表现出一定的潜力,部分人工智能聊天机器人也能够从对话文本中识别焦虑、抑郁等心理状态相关的情绪特征或风险信号。然而,这并不意味着我们可以将技术

⁵https://jyt.jiangsu.gov.cn/art/2025/11/12/art_57813_11673733.html

⁶同脚注 4。

⁷同脚注 1。

可行性、用户接受度直接等同于心理干预的有效性。

从高校应用场景看,生成式人工智能合理的定位应该是传统心理健康服务体系的数字补充。其优势主要体现在高频、低风险和及时,面对复杂的情况,可以迅速转入由专业人员主导的服务体系。因此,高校心理健康 AI 应遵循风险越高, AI 自主处理权限越低,人工专业介入程度越高的基本原则,通过风险分层和专业介入相结合的方式,实现技术效率与服务安全之间的平衡。未来研究也应从 AI 能否提供心理支持进一步转向 AI 适合提供什么样的支持、支持到什么程度以及何时必须退出并转交人工服务,并通过真实高校场景验证它的安全性、可用性和实际服务价值。

同时,本文也存在一定局限。(1) 本文属于叙述性综述,并未按照系统评价标准完成穷尽式文献检索和分析;(2) GenAI 迭代速度极快,不同模型之间差异较大,现有的结论不能照搬照抄、直接迁移至未来模型;(3) 高校心理健康服务的效能还没有充分论证,国内研究结果依然需要通过我国高校真实场景进一步探索。

8. 结论

GenAI 的发展为高校心理健康服务数字化转型提供了新的技术基础。依托 LLMs 的自然语言理解、多轮交互、内容生成和知识调用能力,人工智能已经能够在心理健康知识教育、日常情绪支持、自我调适等领域承担一定的辅助功能,并在提高服务即时性、可及性和个性化程度方面表现出应用潜力。然而,研究同时表明,GenAI 的技术成熟度并不等同于心理健康服务的专业成熟度,它在专业判断、风险识别、信息可靠性、数据隐私、人机关系及责任承担等方面存在明显的局限性,真实高校场景下的长期效能证据也尚不充分。因此,高校 GenAI 心理健康服务应该坚持以人为中心、风险分层、功能有界、安全优先和人机协同的基本原则。对于心理健康教育和一般性情绪支持任务,可以发挥即时性优势;对于复杂心理问题和高风险情境,应该及时转入专业人员主导的服务体系,使技术优势真正嵌入高校已有的心理健康服务体系。

未来高校心理健康人工智能发展的关键,是建立更加专业的知识体系、嵌入更加安全的技术架构、完善数据治理制度、制定效果评价标准和划分人机责任边界。只有实现从技术可用向专业可靠、从人工智能替代向人机协同的转变,GenAI 才能真正成为高校心理健康服务体系的有效补充,实现技术效率与医学人文价值之间的平衡。

基金项目

内蒙古自治区研究生教育教学改革项目“新医科背景下医学研究生人工智能素养分层培育模式构建与实践”(JG2025029C);内蒙古医科大学高等教育教学改革研究与实践项目(NYJXGGSJ2024032)。

参考文献

- 黄峰,丁慧敏,李思嘉,等(2025). 基于大语言模型的自助式 AI 心理咨询系统构建及其效果评估. *心理学报*, 57(11), 2022-2042.
- 秦瑞标(2024). 大模型时代生成式人工智能治理的逻辑更新与回应路径——以《生成式人工智能服务管理暂行办法》为视角. *时代法学*, 22(3), 28-41.
- 吴砥,王雪,尉小荣(2026). 全面推进智能化:时代方位、现实图景与实践路向——《教育部等九部门关于加快推进教育数字化的意见》智能化内容分析与思考. *中国教育信息化*, 32(2), 32-40.
- 徐天然,王慧(2024). 《中华人民共和国个人信息保护法》在高校学生工作实践中的应用. *高校辅导员学刊*, 16(1), 57-62+98.
- 袁小雅,刘仪辉(2025). 人工智能在大学生心理健康评估与干预中的应用探究. *心理学进展*, 15(1), 182-187.
- Cho, H. N., Wang, J., Hu, D., & Zheng, K. (2026). Large Language Model—Based Chatbots and Agentic AI for Mental Health

- Counseling: Systematic Review of Methodologies, Evaluation Frameworks, and Ethical Safeguards. *JMIR AI*, 5, e80348. <https://doi.org/10.2196/80348>
- Feng, X., Tian, L., Ho, G. W. K., Yorke, J., & Hui, V. (2025). The Effectiveness of AI Chatbots in Alleviating Mental Distress and Promoting Health Behaviors among Adolescents and Young Adults: Systematic Review and Meta-Analysis. *Journal of Medical Internet Research*, 27, e79850-e79850. <https://doi.org/10.2196/79850>
- Guo, Z., Lai, A., Thygesen, J. H., Farrington, J., Keen, T., & Li, K. (2024). Large Language Models for Mental Health Applications: Systematic Review. *JMIR Mental Health*, 11, e57400. <https://doi.org/10.2196/57400>
- Guțu, S. M., Cosmou, A., Cojocaru, D., Turturescu, T., Popoviciu, C. M., and Giosan, C. (2021). Bot to the Rescue? Effects of a Fully Automated Conversational Agent on Anxiety and Depression: A Randomized Controlled Trial. *Annals of Depression and Anxiety*, 8, Article No. 1107. <https://doi.org/10.26420/anndepressanxiety.2021.1107>
- Holmes, W., & Miao, F. (2023). *Guidance for Generative AI in Education and Research*. UNESCO Publishing.
- Ma, J., Tan, J. S., Shi, H., & Bao, Y. (2026). The Role of Digital Mental Health Resources in Encouraging Help-Seeking: A Qualitative Study of University Students' Experiences. *Cambridge Prisms: Global Mental Health*, 13, e116. <https://doi.org/10.1017/gmh.2026.10216>
- Nyakhar, S., & Wang, H. (2025). Effectiveness of Artificial Intelligence Chatbots on Mental Health & Well-Being in College Students: A Rapid Systematic Review. *Frontiers in Psychiatry*, 16, Article ID: 1621768. <https://doi.org/10.3389/fpsyt.2025.1621768>
- Olisaeloka, L., Richardson, C. G., Wang, A. Y., Munthali, R. J., & Vigo, D. V. (2026). Safety Mechanisms and Risk Mitigation in Generative AI Mental Health Chatbots: A Systematic Scoping Review. *Healthcare*, 14, Article 1395. <https://doi.org/10.3390/healthcare14101395>
- Sawesi, S., Sabbineni, H., Shagamreddy, R., & Rashrash, B. (2026). Cybersecurity and Privacy Risks of Generative AI Mental-Health Chatbots: A Systematic Review and Regulatory Framework. *Journal of Multidisciplinary Healthcare*, 19, 1-21. <https://doi.org/10.2147/jmdh.s581251>
- Schlicher, M., Li, Y., Murthy, S. M. K., Sun, Q., & Schuller, B. W. (2025). Emotionally Adaptive Support: A Narrative Review of Affective Computing for Mental Health. *Frontiers in Digital Health*, 7, Article ID: 1657031. <https://doi.org/10.3389/fdgth.2025.1657031>
- Sobowale, K., Humphrey, D. K., & Zhao, S. Y. (2025). Evaluating Generative AI Psychotherapy Chatbots Used by Youth: Cross-Sectional Study. *JMIR Mental Health*, 12, e79838-e79838. <https://doi.org/10.2196/79838>
- Wang, X., Zhou, Y., & Zhou, G. (2025). The Application and Ethical Implication of Generative AI in Mental Health: Systematic Review. *JMIR Mental Health*, 12, e70610. <https://doi.org/10.2196/70610>
- World Health Organization (2024). *WHO Releases AI Ethics and Governance Guidance for Large Multi-Modal Models*. World Health Organization.
- Yeh, P. L., Kuo, W. C., Tseng, B. L., & Sung, Y. H. (2025). Does the AI-Driven Chatbot Work? Effectiveness of the Woebot App in Reducing Anxiety and Depression in Group Counseling Courses and Student Acceptance of Technological Aids. *Current Psychology*, 44, 8133-8145. <https://doi.org/10.1007/s12144-025-07359-0>