

生成式人工智能背景下大学生学术诚信建设的困境与路径重构

王仰飞

北京联合大学马克思主义学院, 中国·北京 100101

摘要: 在生成式人工智能快速发展的背景下, 大学生学术诚信建设面临新的困境。一方面, AI 工具的普及为学术不端行为提供了便利, 例如自动化生成论文、篡改实验数据等, 增加了学术诚信监管的难度; 另一方面, 传统学术诚信规范尚未适应技术变革, 导致治理标准模糊、责任界定困难。为重构学术诚信建设路径, 需从多维度入手。需构建“预防-检测-惩戒”分层治理体系; 利用 AI 模拟案例, 提升学生诚信判断能力, 建立“AI-教师-学生”协同机制。需要政府、高校、师生多方协同, 完善制度规范, 共同维护学术诚信体系。

关键词: 生成式人工智能; 学术不端; 诚信教育

The Dilemma and Path Reconstruction of Academic Integrity Construction for College Students in the Context of Generative Artificial Intelligence

Wang Yangfei

School of Marxism, Beijing Union University, China Beijing100101

Abstract: Against the backdrop of the rapid development of generative artificial intelligence (AI), the construction of academic integrity among university students faces new challenges. On one hand, the widespread adoption of AI tools facilitates academic misconduct, such as automated essay generation and data manipulation, increasing the difficulty of academic integrity oversight. On the other hand, traditional academic integrity norms have yet to adapt to technological advancements, resulting in ambiguous governance standards and difficulties in accountability attribution. To reconstruct the path for academic integrity development, a multi-dimensional approach is necessary. A tiered governance system of “prevention-detection-penalty” should be established. Additionally, AI-simulated case studies can enhance students' ethical decision-making abilities, while a collaborative “AI-teacher-student” mechanism should be implemented. Government bodies, universities, educators, and students must work together to refine regulatory frameworks and collectively uphold academic integrity.

Keywords: Generative artificial intelligence; Academic misconduct; Integrity education

0 引言

近年来, 生成式人工智能的迅猛发展正在深刻重塑高等教育的学术生态。以 ChatGPT 等为代表的大语言模型, 不仅为学术研究提供了智能化辅助工具, 同时也对传统学术诚信规范带来了前所未有的挑战。一方面, AI 工具的普及为学术不端行为提供了便利; 另一方面, 传统学术诚信规范尚未适应技术变革, 导致治理标准模糊、责任界定困难。面对生成式人工智能对大学生学术诚信建设带来的挑战, 大学生学术诚信建设路径也需要重构。

1 生成式人工智能在大学生学术领域的应用场景

在当今数字化时代, 生成式人工智能已深度融入大学生的学术行为当中, 在课程作业、论文写作、学术研究辅

助等多个场景中发挥着重要作用。

一是在课程作业场景中, 生成式人工智能成为大学生获取思路和解决难题的得力助手。以语言类课程作业为例, 面对英语写作任务时, 学生常借助 ChatGPT 等工具进行语法检查和语句润色。当学生完成一篇英语作文初稿后, 将其输入 ChatGPT, 它能迅速识别出语法错误, 如主谓不一致、时态错误等, 并给出修改建议, 同时还能提供更丰富、更地道的表达方式, 提升作文的语言质量。在数理学科作业方面, 生成式人工智能同样大显身手。对于复杂的数学问题, 如高等数学中的微积分计算、线性代数中的矩阵运算等, 学生可以利用专门的数学计算生成式人工智能工具, 输入题目条件, 工具便能快速给出详细的解题步骤和答案, 帮助学生理解解题思路, 掌握知识点。

二是论文写作场景中,生成式人工智能从选题到完稿的各个环节都为大学生提供了有力支持。在选题阶段,学生可借助生成式人工智能工具分析学术前沿动态和研究热点。例如,通过输入相关学科领域的关键词,工具能够整合大量学术文献,梳理出当前该领域的热门研究方向和尚未解决的问题,为学生提供新颖且具有研究价值的选题建议。在文献综述部分,生成式人工智能可帮助学生快速筛选和整理文献。学生输入关键词后,工具能从海量的学术数据库中检索相关文献,并生成文献综述的大纲,对每篇文献的核心观点、研究方法和主要结论进行概括,极大地提高了文献综述的撰写效率。在论文内容创作阶段,生成式人工智能可以根据学生给定的主题和大纲,生成段落内容。

三是在学术研究辅助场景中,生成式人工智能也展现出独特的价值。在数据收集与分析环节,对于大规模的问卷调查数据,生成式人工智能可以运用自然语言处理技术对开放性问题的回答进行自动分类和主题提取。在实验设计阶段,生成式人工智能能够根据研究目的和已有研究成果,为学生提供实验设计的参考方案,包括实验变量的控制、实验步骤的安排等,帮助学生设计出更科学、更合理的实验方案。

据第三方机构“麦可思”对3000多名高校师生的问卷调查揭示,近六成的高校师生日常频繁使用生成式AI,其中,大学生群体在使用时,有近三成将其主要用于撰写论文或完成作业。这一现象凸显了人工智能时代学术诚信建设的紧迫性和复杂性。

2 生成式人工智能带来的大学生学术诚信挑战

2.1 模糊性剽窃问题

随着生成式人工智能在教育领域的广泛应用,学术诚信面临新的挑战。以ChatGPT为代表的大语言模型基于海量文本数据进行训练,通过算法重组生成新的内容,这一过程往往缺乏规范的引用标识。当学生使用AI生成学术内容时,模型可能将多篇文献的观点和表述进行无标注拼接,导致学生在不知情的情况下构成抄袭。这种由技术特性引发的剽窃行为,因其隐蔽性和非故意性,对传统学术诚信体系提出了严峻考验。“它模糊了原本清晰的知识侵权界定,使学术道德失范从显性抄袭转向隐性共谋,导致传统学术道德规范被遮蔽。”

生成式AI的出现使得剽窃行为的界定变得异常困难。传统查重系统主要检测文本相似度,而AI生成内容经过算法重组后,与源文献的相似度可能显著降低,从而规避检测。更棘手的是,当AI生成内容与现有文献存在相似时,很难判定这是模型训练的自然结果还是学生的故意抄袭。学生可能辩称仅将AI作为辅助工具,但“合理使用”与

“学术不端”的界限尚未明确。这种模糊性不仅增加了教师评判的难度,也为学术诚信教育带来了新的课题。

高校中关于AI代写论文的实际案例也揭示了这一问题的复杂性。学生使用ChatGPT生成的论文虽经修改,但仍包含大量未标注的学术观点,最终被认定为模糊性剽窃。这些案例表明,AI辅助写作可能产生看似原创实则侵权的学术成果,这种新型剽窃形式不仅破坏学术公正性,也可能误导学生对学术规范的理解。随着AI技术的普及,如何建立适应技术发展的学术诚信规范,已成为亟待解决的重要议题。

2.2 责任主体争议

生成式人工智能在学术创作中的广泛应用,引发了关于责任主体认定的新争议。传统学术署名机制要求作者对研究成果承担完全责任,但当AI深度参与研究过程时,这一机制面临挑战。当前学术界对AI是否应被视为“合著者”存在分歧:一方面,AI不具备主观能动性和道德判断能力,无法承担学术责任;另一方面,AI在数据分析、内容生成等方面的实质性贡献又难以忽视。部分学术机构建议在致谢部分注明AI辅助,而非将其列为合著者。“AIGC不能履行文章作者的角色,也不能被列为文章作者。”这种折中做法虽然暂时缓解了争议,但尚未形成统一规范,导致学生在实际应用中缺乏明确指引。

教师群体在判定AI辅助与学术不端的界限时同样面临困境。由于缺乏统一的判断标准,教师往往需要依靠主观经验来评估学生对AI的使用程度。这种评估面临双重困难:其一,AI生成内容与学生原创内容在语言风格上日趋接近,难以区分;其二,不同教师对AI辅助的容忍度存在显著差异,导致评判标准不一。更复杂的是,学生可能对AI生成内容进行二次修改,使其既能通过查重检测,又保留AI的核心观点。这种“半人工”作品进一步模糊了合理使用与学术不端的界限。

现有的检测技术尚不足以有效应对这一挑战。虽然部分查重系统开始整合AI内容识别功能,但其准确率仍有待提升。技术局限性导致教师往往陷入两难:过度怀疑可能损害师生信任,过度宽容又可能纵容学术不端。这种判断困境不仅影响具体案例的处理,更对学术评价体系的公正性构成潜在威胁。建立适应AI时代的学术规范,需要教育机构、技术开发者和学术共同体共同参与,制定兼顾技术创新与学术诚信的新标准。

2.3 能力退化风险

生成式人工智能在高等教育中的深度应用正在引发一系列值得警惕的能力退化风险。“若师生在文科教育教学中一味依赖AIGC快速的内容生成和数据分析能力,而不进行自我思考,无异于知识的流水线生产,会使师生的知识思维能力日渐堕化。”

批判性思维作为学术研究的核心素养,其培养过程正受到 AI 工具广泛使用的显著影响。传统学术训练要求学生通过广泛阅读、深入分析和反复质疑来建构知识体系,而过度依赖 AI 的学生往往满足于接受算法生成的现成答案,缺乏对问题本质的探究。这种思维惰性导致学生在面对开放性问题时表现出明显的分析能力不足,其学术产出往往停留在对 AI 生成内容的表层整合,难以形成具有原创性的学术观点。更值得关注的是,这种思维模式的转变可能具有长期累积效应,最终影响学生整体的认知发展轨迹。

研究能力的系统性退化是另一个突出表现。规范的学术研究要求研究者掌握从方法选择、数据采集到文献综述的完整技能链,而 AI 工具的便捷性正在消解这一训练过程。在研究方法层面,学生倾向于直接采纳 AI 推荐的研究设计,缺乏对方法适用性的审慎考量;在数据处理环节,算法暗箱操作取代了学生亲手处理数据的实践机会,导致其数据分析能力停留在工具使用层面;在文献综述方面, AI 生成的综合性文本虽然覆盖面广,却难以替代研究者对文献脉络的深度把握。这种能力缺陷在实证研究中表现得尤为明显,学生往往无法解释 AI 生成的数据分析结果,更遑论对其有效性进行评估。这种研究能力的空心化趋势,将严重影响未来学术人才的质量。

从长远来看,这种能力退化可能产生深远的负面影响。在学术发展维度,当学生进入研究生阶段后,其薄弱的独立研究能力将难以应对更高层次的学术要求;在职业发展层面,批判性思维和研究能力的缺失将限制其在知识密集型行业的发展空间。更值得警惕的是,这种能力退化可能形成代际传递效应:当缺乏扎实学术训练的学生未来成为教育者时,其教学方式可能进一步强化对 AI 工具的依赖,形成恶性循环。面对这一挑战,高等教育机构需要重新审视 AI 时代的培养方案,在利用技术便利的同时,坚守学术能力培养的核心目标,确保技术创新与学术传承的平衡发展。

3 现行学术诚信规范体系在生成式人工智能背景下的不足

3.1 传统学术诚信规范的滞后性

生成式人工智能的快速发展对传统学术诚信规范体系提出了全新挑战。现行学术诚信规范主要针对人类作者间的学术行为,尚未对 AI 生成内容的使用和归属作出明确规定。这种制度空白导致学生在使用 AI 辅助学术创作时缺乏明确指引:一方面,规范未界定 AI 生成内容的合理使用范围,使学生难以把握使用边界;另一方面,对 AI 生成内容的知识产权归属、学术贡献认定等关键问题缺乏判定标准。学术诚信规范完全未提及 AI 生成内容,导致师生在论文写作和评价过程中陷入困惑。这种制度滞后性不仅增加了学术不端风险,也影响了学术评价的公正性。“当高校研究者

卷入涉及 AIGC 的学术道德失范指控处境时,责任认定面临以下争议:归责主体应采用人类单一主体模式还是人机联合模式?联合归责时的责任比例如何划分?对这些关键问题的处理在现行学术规范体系中仍处于无法可依的尴尬境地。”

传统学术诚信规范在技术适应性方面面临严峻考验。其核心理念强调学术的独立性和原创性,但 AI 的介入使学术创作转变为“人机协作”的新模式,原有理念亟待更新。在具体规则层面,传统反抄袭机制主要依赖文本相似度检测,而 AI 生成内容通过算法重组可轻易规避查重系统。更棘手的是责任界定问题:当 AI 生成内容出现学术不端时,现行规范无法明确划分开发者、使用者之间的责任。某研究案例中,学生使用 AI 生成的错误数据导致研究结论失实,但由于缺乏责任认定标准,最终陷入开发者与使用者相互推诿的困境。这些不适应现象严重削弱了学术规范的约束力,亟须建立兼顾技术创新与学术诚信的新规范体系。

3.2 高校管理政策的缺失

当前高校学术不端检测体系面临生成式人工智能带来的严峻挑战。传统查重系统主要针对人类作者间的文本复制行为,通过比对已有文献数据库来判断抄袭,如知网查重系统通过连续文本相似度检测来识别剽窃行为。然而,这类系统对 AI 生成文本的识别存在明显局限:首先, AI 通过算法重组多源信息生成内容,其文本相似度往往低于检测阈值;其次, AI 文本特有的模式化语言风格和逻辑结构差异,超出了传统系统基于字面比对的检测能力。例如, ChatGPT 生成的论文段落虽借鉴多篇文献观点,但经过算法重构后,其重复率可能完全符合学术规范要求,导致现有检测工具失效。这种技术滞后性使得 AI 辅助的学术不端行为得以规避传统监管,严重削弱了学术诚信保障机制的效力。

高校在处置 AI 相关学术不端行为时,面临惩戒标准体系的双重困境。一方面,现行惩戒规范主要针对传统抄袭行为,难以适用于 AI 生成内容的新型学术不端。各高校在处理尺度上存在显著差异:有的将未声明使用 AI 视为轻微违规,有的则认定为严重学术失信,这种标准不统一损害了学术纪律的严肃性。另一方面,责任主体认定存在模糊地带。当学生使用 AI 生成核心观点却未标注时,既难以界定其主观故意程度,也无法明确 AI 开发者、教育机构等各方的连带责任。学生因使用 AI 生成论文主体内容被警告处分,将引发关于“技术使用”与“学术不端”界限的激烈争议。这种标准缺失不仅导致处理结果缺乏公信力,更反映出既有学术规范体系在应对技术革新时的系统性不足,亟须建立与 AI 时代相适应的新型学术治理框架。

3.3 教育与引导的不足

当前高校 AI 伦理教育体系存在显著不足,难以应对

生成式人工智能带来的学术诚信挑战。传统学术道德课程仍聚焦于抄袭、剽窃等常规学术不端行为,对AI伦理议题涉及甚少。课程内容缺乏对AI生成内容知识产权归属、使用边界等关键问题的系统阐述,导致学生在实际应用中缺乏明确指引。更突出的是,教学方法仍以单向讲授为主,未能充分利用AI技术本身开展互动式、情景化的伦理教育,使学生难以深入理解AI应用中的伦理困境。这种教育内容的滞后性与教学方式的单一性,导致学生在使用AI工具时既缺乏必要的伦理认知,也欠缺独立判断能力。

高校在引导学生正确使用AI方面同样存在系统性缺失。在认知层面,宣传教育活动往往片面强调AI的效率优势,而忽视对其潜在学术风险的警示,使学生对技术应用形成偏颇认知。在实践层面,既缺乏针对不同学科特点的AI使用指南,也缺少对教师相关能力的系统培训,导致师生在面对具体问题时无所适从。更关键的是,相关教育措施多为临时性、碎片化安排,未能构建贯穿人才培养全过程的AI伦理教育体系。这种引导不足使得学生既可能因过度依赖AI而弱化研究能力,又可能因规范缺失而陷入无意识的学术不端,亟须建立系统化、常态化的AI应用指导机制。

4 生成式人工智能背景下大学生学术诚信建设的路径重构

4.1 构建分层治理框架

生成式人工智能的快速发展对大学生学术诚信建设提出了全新挑战,亟需构建“预防-检测-惩戒”三位一体的分层治理体系。

在预防层面,应将AI伦理教育系统融入学术道德课程体系,通过模块化设计将AI生成内容的知识产权归属、使用规范等核心议题纳入教学内容。创新采用案例研讨、情景模拟等互动教学方式,结合跨学科真实案例,培养学生的伦理判断能力。同时加强师资培训,建立学术诚信教师与人工智能专家的协同机制,确保教学内容的前沿性和实践性。这种预防性教育能帮助学生建立正确的技术使用认知,从源头上减少学术不端行为的发生。

技术检测是学术诚信建设的关键环节,需要开发基于多模态分析的新型检测系统。通过文本特征分析识别AI生成内容在语言风格、逻辑结构等方面的特征差异,结合行为日志分析学生的学习行为模式,构建多维度的学术不端识别模型。系统应整合自然语言处理、机器学习等前沿技术,持续优化算法以适应快速迭代的AI生成技术。同时建立动态更新的样本数据库,提升系统对新型学术不端行为的识别能力。这种技术赋能的检测机制能够有效弥补传统查重系统的不足,为学术诚信建设提供可靠的技术支撑。

在惩戒层面,应建立梯度化、规范化的处罚机制。根据学术不端行为的性质、程度及影响,制定从警告教育到

开除学籍的分级处罚标准。对轻微违规行为以教育矫正为主,通过诚信教育引导规范学术行为;对严重学术造假行为则实施严厉处罚,维护学术尊严。完善申诉复核程序,确保处罚的公正性和透明度。同时建立学术诚信档案,将学生学术行为与评奖评优等挂钩,形成长效约束机制。这种刚柔并济的惩戒体系既能有效震慑学术不端,又能保障学生的合法权益,促进学术生态的健康发展。

4.2 基于AI技术的学术诚信教育

生成式人工智能的快速发展为高等教育带来了全新的学术诚信挑战,亟需构建系统化的治理体系。基于AI技术运用教育学生“进行自主学习、个性化学习和探究学习,以提升知识构建能力和价值认知水平,养成良好的人文素养和职业道德,从而成长为新时代全面发展的社会主义创新型人才。”

在教育层面,应充分利用AI技术本身开展创新性伦理教育。通过AI模拟学术不端案例,可以构建高度仿真的虚拟情境,让学生通过角色扮演深入体验学术诚信困境。这种沉浸式教学方法能够针对不同专业特点设计个性化案例,如为理工科学生模拟实验数据造假场景,为文科学生设计文献抄袭情境。在模拟过程中,AI系统可根据学生的伦理决策实时生成反馈,形成互动式学习体验。相比传统说教式教育,这种方法更能激发学生的参与热情,帮助其深入理解学术规范的本质,培养独立判断能力。同时,这种教育方式还能有效训练学生的批判性思维、团队协作等核心素养,实现伦理教育与能力培养的有机统一。

在技术应用层面,需要构建“AI-教师-学生”三位一体的协同学习模式。这一模式的关键在于明确各方的角色定位:AI作为智能助手,提供个性化学习支持与实时反馈;教师发挥引导作用,帮助学生把握人机协作的边界;学生作为学习主体,主动参与知识建构。具体实施中,AI可辅助完成文献检索、数据分析等基础性工作,教师则聚焦于启发学生思考、培养创新能力。为保障模式有效运行,学校需建立配套支持体系,包括技术平台建设、师生AI素养培训等。这种协同模式不仅能提升学习效率,更能让学生在实践过程中形成正确的技术使用观,避免过度依赖AI。通过定期组织成果展示与互评活动,可以进一步强化学生的学术规范意识,实现技术赋能与伦理约束的平衡发展。

4.3 加强政策引导与多方协同

生成式人工智能的快速发展对高等教育领域的学术诚信建设提出了全新挑战,需要政府、高校、教师和学生多方协同参与,构建系统化的治理体系。

政府层面应当发挥主导作用,通过制定具有前瞻性的政策法规来规范AI在教育领域的应用。首先需要明确界定AI工具在学术活动中的合法使用范围,要求学生在使用

AI 辅助完成作业或论文时必须进行明确标注, 并对生成内容进行批判性思考和合理引用。同时应建立严格的数据安全监管机制, 规范教育数据的收集、存储和使用流程, 实施最小必要原则和加密保护措施, 确保学生隐私不受侵犯。此外, 政府还应推动建立 AI 生成内容检测标准、学术不端判定标准等规范体系, 为高校提供统一的治理依据。这些政策法规的制定需要教育部门、科技部门和行业专家共同参与, 既要促进技术创新, 又要维护学术公正。

高校作为学术诚信建设的主体, 需要建立“制度 - 教育 - 监督”三位一体的治理机制。在制度层面, 要制定详细的 AI 使用规范, 明确学术成果中 AI 生成内容的比例限制和标注要求。在教育层面, 应将 AI 伦理纳入学术道德课程体系, 通过案例教学、情景模拟等方式培养学生的学术诚信意识。在监督层面, 要建立学术诚信委员会, 完善论文抽查和学术不端处理流程。教师作为具体执行者, 需要提升 AI 素养和伦理判断能力, 在教学过程中引导学生合理使用 AI 工具, 培养学生的独立思考和创新能力。学生作为学术活动的主体, 应当树立正确的学术价值观, 既要善用 AI 提高学习效率, 又要保持学术独立性。通过建立高校、教师、学生之间的常态化沟通机制, 形成多方协同的治理合力, 才能有效应对生成式 AI 带来的学术诚信挑战, 维护健康的学术生态。

参考文献:

[1] 央视网[新闻直播间] 记者观察“大学生论文“AI 味儿”渐浓 怎么破?” 2025 年 2 月 26 日.<https://tv.cctv.cn/2025/02/26/VIDEAETkgDK4T9ptbZhOylBP250226.shtml>.

[2] 杜静, 冯凯瑞, 王东. 生成式人工智能背景下高校学术道德失范的内涵重塑与治理策略[J]. 苏州大学学报(教育科学版), 网络首发时间 2025 年 4 月 27 日.

[3] 中国科学技术信息研究所:《学术出版中的 AIGC 使用边界指南 2.0》. 2024 年 9 月 26 日.

[4] 金华, 王汉英. 生成式人工智能视域下高校文科教学的伴生性风险与防范[J]. 黑龙江高教研究, 2025 年第 5 期, 总第 373 期.

[5] 杜静, 冯凯瑞, 王东. 生成式人工智能背景下高校学术道德失范的内涵重塑与治理策略[J]. 苏州大学学报(教育科学版), 网络首发时间 2025 年 4 月 27 日.

[6] 房宏君, 蔡红, 汪昕宇. 数智化教育背景下知识图谱赋能课程思政教学研究与设计[J]. 北京联合大学学报第 39 卷第 2 期总 144 期, 2025 年 3 月.

[7] 高奇琦. 知识革命还是知识堕化——ChatGPT 与知识生成秩序[J]. 中国社会科学评价, 2023(2):8-13.

[8] 周洪宇, 李宇阳. ChatGPT 对教育生态的冲击及应对策略[J]. 新疆师范大学学报(哲学社会科学版), 2023(4):102-112.

[9] 周洪宇, 常顺利. 生成式人工智能嵌入高等教育的未来图景、潜在风险及其治理[J]. 现代教育管理, 2023(11):1-12.

[10] 李菲, 吕欣. AIGC 在传媒艺术教育中的应用场景与生产机制研究[J]. 传媒, 2024(19):12-15.

[11] 王佑镁, 王旦, 梁炜怡, 等. ChatGPT 教育应用的伦理风险与规避进路[J]. 开放教育研究, 2023(2):26-35.

[12] 吴南中, 陈咸彰, 冯永. 从“失序”到“有序”: 生成式人工智能教育应用的转向及其生成机制[J]. 远程教育杂志, 2023(6):42-51.

基金项目: 本文为 2023 年度北京市社会科学基金重点项目(项目编号: 23LLMLB090)“习近平总书记关于党的青年工作重要论述的问题指向、践行进路和重大意义研究”的阶段成果。

作者简介: 王仰飞(1978-), 男, 安徽六安人, 北京联合大学马克思主义学院讲师, 博士, 研究方向: 大学生思想政治教育。