Advances in Social Sciences 社会科学前沿, 2026, 15(11), 59-69
Published Online November 2026 in Hans. https://www.hanspub.org/journal/ass
https://doi.org/10.12677/ass.2026.1511899

Hans汉斯

生成式人工智能时代高校思政课教学的挑战与创新

高雨欣[1]，田　璐[2]

[1]文华学院马克思主义学院，湖北　武汉
[2]华中师范大学人工智能教育学部，湖北　武汉

收稿日期：2026年9月11日；录用日期：2026年10月29日；发布日期：2026年11月9日

摘　要

生成式人工智能正在改变高校思政课教学的内容供给、互动方式与评价模式，同时也带来新的教学挑战，表现为教师主导地位弱化、教学内容质量失范、育人效能弥散以及伦理规范缺位。在此基础上，本文提出技术应用应坚持四项基本前提：教育性前提，即育人目标与技术应用的统一；主体性前提，即教师主导与学生能动的统一；发展性前提，即全面发展与素养提升的统一；规范性前提，即伦理约束与制度保障的统一。进而从创新路径角度提出，应以教师主导为核心构建人机协作教学新形态，以内容审核为重点建立技术过滤与教师把关机制，以三维融合为路径优化认知诊断与沉浸体验模式，以伦理制度为保障完善数据保护与算法透明体系。四者相互衔接，共同确保生成式人工智能在服务育人目标的前提下合理应用，为高校思政课教学创新提供参考。

关键词

生成式人工智能，高校思政课，教学挑战，教学创新，技术伦理

Challenges and Innovations in the Teaching of Ideological and Political Theory Courses in Colleges and Universities in the Era of Generative Artificial Intelligence

Yuxin Gao[1], Lu Tian[2]

[1]Institute of Marxism, Wenhua College, Wuhan Hubei
[2]Faculty of Artificial Intelligence in Education, Central China Normal University, Wuhan Hubei

Received: September 11, 2026; accepted: October 29, 2026; published: November 9, 2026

文章引用: 高雨欣, 田璐. 生成式人工智能时代高校思政课教学的挑战与创新[J]. 社会科学前沿, 2026, 15(11): 59-69.
DOI: 10.12677/ass.2026.1511899

Abstract

Generative artificial intelligence is transforming the content supply, interaction methods, and evaluation models of the teaching of ideological and political theory courses in colleges and universities, while also bringing new teaching challenges, manifested as the weakening of teachers' leading role, anomie in the quality of teaching content, dissipation of educational effectiveness, and absence of ethical norms. On this basis, this paper proposes that the application of technology should adhere to four basic premises: the educational premise, namely the unity of educational goals and technological application; the subjectivity premise, namely the unity of teacher leadership and student agency; the developmental premise, namely the unity of comprehensive development and competency enhancement; and the normative premise, namely the unity of ethical constraints and institutional safeguards. Furthermore, from the perspective of innovation paths, it proposes that a new form of human–machine collaborative teaching should be built with teacher leadership at its core; a mechanism of technological filtering and teacher gatekeeping should be established with content review as the focus; cognitive diagnosis and immersive experience models should be optimized through three-dimensional integration; and systems for data protection and algorithmic transparency should be improved with ethical regulations as safeguards. These four aspects are interconnected, jointly ensuring the reasonable application of generative artificial intelligence on the premise of serving educational goals, and providing a reference for innovation in the teaching of ideological and political theory courses in colleges and universities.

Keywords

Generative Artificial Intelligence, Ideological and Political Education in Colleges and Universities, Teaching Challenges, Teaching Innovation, Ethics of Technology

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

人类社会正加速迈向以数字化、智能化为特征的生成式人工智能时代。生成式人工智能凭借较强的内容生成能力与广泛的适用性，正在推动教育领域的转型重构。高校思政课教学作为落实育人目标的重要环节，也面临着新的技术环境与教学条件。生成式人工智能可以辅助知识供给、案例生成、学习反馈等环节，提升教学的交互性与针对性，但其在主体关系、内容质量、教学效能与伦理规范等方面也带来了新的挑战。如何在这一背景下推动思政课教学的合理创新，成为需要认真研究的问题。

技术条件的变化，并不必然改变思政课教学的根本目标。思政课教学承担着价值引导、思想交流与人格培育的职责，其育人目标不因技术手段的更新而改变。生成式人工智能改变的是教学的工具与环境，而非教学的目的与方向。因此，在讨论技术应用时，需一方面，正视技术给教学带来的现实挑战，明确教学的基本前提；另一方面，积极探索技术融合的教学路径，使技术应用服务于育人目标的实现。基本前提为技术应用确立方向与边界，创新路径则将前提要求转化为可操作的教学实践。

基于这一思路，本文首先分析生成式人工智能时代高校思政课教学面临的多维挑战，继而阐明技术应用应当坚持的基本前提，最后提出技术融合创新的具体路径，以期为思政课教学的合理变革提供参考。

2. 生成式人工智能时代高校思政课教学的多维挑战

生成式人工智能在高校思政课教学中的应用不断扩展，既改变教学效率与交互方式，也带来多重挑战。在主体关系上，AI 参与知识供给与互动反馈，可能削弱教师主导地位；在内容生产上，算法生成存在信息幻觉与价值偏差，影响内容质量；在教学实践上，沉浸体验可能遮蔽价值内涵，导致育人效能分散；在应用伦理上，数据采集与算法控制缺乏规范，伦理风险上升。四者相互关联，共同影响教学目标的实现。

2.1. 人机协同过程中教师教学主导地位趋于弱化

不能单纯将生成式人工智能视为人身体器官的延伸，而是要深刻认识到“各种专用、通用人工智能具有越来越高的类人智能，可以在生产和服务领域独立、自主地完成劳动任务”[1]。生成式人工智能正以类人化智能深度嵌入高校思政课教学流程。它在知识检索、实时答疑、案例生成等环节表现出较强的认知模拟能力，能够部分替代教师的知识供给与基础互动职能。这使得其在教学中的角色从被动辅助工具转变为具有一定自主性的行动者。传统“教师–学生”的权威传递结构由此被重塑为“教师-AI-学生”三元协同系统。在这一系统中，生成式 AI 既是内容共创者，又是学习互动的中介，其输出过程兼具创造性与适应性的类人。教育者的职能重心由此开始从信息传递转向价值塑造与思想引导。这一变革虽未赋予机器完全的主体地位，但已实质性推动教学权威从单一主导转向教师、生成式 AI 与学生多元主体的动态互动。

生成式人工智能对教学主体关系的重塑，可能造成教师主导作用弱化、学生认知自主受限、师生价值与情感联结削弱等现实困境。正如马克思所警示的那样，科学技术有时也“表现为异己的敌对的和统治的权力”[2]，生成式 AI“也可能造成教育者主导作用弱化、受教育者认知自主受限、主客体价值与情感联结削弱等现实困境”[3]。这些困境具体表现为三个方面。其一，生成式 AI 凭借知识覆盖的显著优势及对教学流程的算法化定义，削弱了教师的知识权威。当教学过程受制于智能算法对知识供给与认知规则的预设框架，教师容易从价值引导者降格为机械的知识传递环节。其二，“AI 智慧学伴”的即时应答功能，使知识获取简化为单向的“刺激–反应”模式。学生沉浸于技术营造的认知舒适区，批判性思维与独立思考能力因缺乏深度探究而逐渐弱化。这种“算法捷径”消解了思想淬炼过程，与思政课“启智润心”的目标相背离。其三，当生成式 AI 以标准化的信息输出替代教师的情感表达与价值传递时，课堂的“温度”被量化为师生互动的“频率”。师生在技术依赖中可能逐渐丧失主体自觉：教师因主导地位弱化而削减育人使命感，学生因认知惰性而轻视教师的价值引导。

2.2. 算法生成过程中教学内容质量面临失范风险

“知识生产是思想政治教育学科创新发展的动力之源和立身之本，是推动学科知识更新迭代、构建学科自主知识体系、实现学科高质量发展的基础和关键”[4]。生成式人工智能正在改变高校思政课教学内容的生产方式。传统内容主要由教育行政部门和教材编写机构主导，采用线性编辑与周期性修订模式，更新速度较慢，难以及时反映理论发展与社会变化。生成式 AI 凭借实时响应与智能整合能力，可在一定程度上突破这一局限。它可以追踪公开权威文献与政策文本，将重要理论表述转化为结构化教学方案，缩短内容更新周期。同时，生成式 AI 赋予师生更多创作自主性：教师可生成本土化教学案例，学生可借助轻量化工具制作政策图解、思辨短评等学习成果。经过审核的用户生成内容汇入校本资源库后，可与专业教材形成互补。由此，思政课教学内容从线性编纂转向动态更新的开放系统，内容供给的效率和多样性得到提升。

但算法生成也带来内容质量失范风险。首先，风险源于模型本体的认知局限。“大语言模型在开发

与运用过程中展现出了数据偏见、内容幻觉和受众本位等弊端”[5]。当生成式 AI 对重要理论概念进行解读时，可能只是对网络文本进行机械拼贴——表面完整，实则缺乏理论深度。思政课教学要求以彻底的理论说服人，而算法生成的内容若停留在数据相关性的计算层面，就容易削弱理论阐释的严谨性。其次，风险来自训练语料的价值偏差。若监管不足，互联网语料库中的极端观点、流量导向内容与片面叙事可能被模型吸收，并借由算法输出渗入教学资源。再次，风险还表现为内容审核与更新机制不健全。生成式 AI 可以快速产出大量素材，但若缺乏教师把关与专业审核，错误信息、片面观点和低质内容就可能进入课堂。思政课教学内容不仅要求知识准确，更要求价值导向正确。若算法生成内容未经严格筛选便用于教学，就会影响教学内容的思想性与规范性。因此，在利用生成式 AI 提升内容生产效率的同时，必须正视其质量失范风险，建立相应的审核与治理机制。

2.3. 沉浸体验过程中教学育人效能出现弥散倾向

生成式人工智能正在推动高校思政课教学从经验型教学转向精准教学。移动终端与云计算构建的学习入口，使课程答疑、知识测验等环节延伸到生活场景，历史文化场馆的 AR 导览也使“生活即课堂”成为可能。借助多媒体融合技术生成的互动历史场景，抽象理论被转化为可感知的立体叙事，有助于唤起学生兴趣与情感认同。AI 助教能根据对话内容聚焦学生困惑，生成分层教学建议，分布式协作系统支持多人同步构建思维导图，教师从知识讲授者转向教学引导者。教学交互性与针对性因此提升。

然而，当沉浸体验成为教学的重要形式，育人效能的实现也面临新的不确定性。技术呈现越是丰富，价值内涵越容易被视觉冲击所遮蔽。虚拟场景的感官体验如果强于理论阐释，技术手段就容易从辅助工具变成教学中心。学生可能记住画面与情节，却未理解理论背后的价值逻辑。思政课强调以理服人、以价值引导人，若沉浸体验停留在感官层面，教学的思想性就会被削弱。随着体验不断深入，知识结构的碎片化问题也会显现。个性化推荐容易形成信息茧房，AI 问答则可能把系统理论拆解为孤立知识点。学生习惯于即时获取答案，深度学习让位于碎片化认知，系统性的知识图谱被分散的信息片段替代，独立思考与批判分析的空间反而收窄。

更值得关注的是，师生之间的价值与情感联结可能因此弱化。情感计算模拟的共情无法替代教师基于自身经历与价值认同的引导。当学生习惯于向 AI 求解思想困惑，师生对话的信任基础就会受到影响；教师若过度依赖技术，也可能减少对学生的思想关注与价值引导。沉浸体验虽能提升教学吸引力，却可能使教学育人效能出现弥散。技术应用必须始终围绕育人目标，避免以体验替代思考、以形式替代内容、以技术替代教师的价值引导。沉浸体验的价值不在于技术本身的新颖，而在于其是否服务于学生的认知深化与价值认同。若缺乏教师引导与制度约束，技术越复杂，育人效能越容易分散。

2.4. 数据应用过程中教学伦理规范存在缺位问题

生成式人工智能在思政课教学中的深入应用，使教学数据的采集、存储与使用范围不断扩大，由此带来的伦理风险逐渐显现。深度学习模型以“精准教学”为名，大量收集学生的学习行为、思想动态、社会参与等信息。这些数据一旦上传云端，便可能脱离教学场景，成为商业机构分析青年群体的数据来源。学生对数据采集的范围与用途往往缺乏清晰认知，知情同意流于形式，数据授权机制不健全。当教学数据被过度采集和不当使用时，学生隐私与教学信任都会受到损害。

数据应用中的算法控制更为隐蔽。基于用户画像的“认知投喂”机制，根据学生偏好推送内容，容易形成信息茧房，使教育公平原则在算法逻辑中受到削弱。若训练语料中英文内容占比较高，主流价值内容可能在参数权重分配中被边缘化，成为算法输出的次要内容。算法的不透明使教师和学生难以判断内容生成的依据，也难以对输出结果进行有效干预。技术本身的不完善尚可通过改进加以弥补，但伦理

失范与监管缺位带来的问题更为复杂。当生成式 AI 伪造历史场景影像在校园传播，当学生利用开源模型批量生成课程论文冲击学术诚信，当 AI 助教输出错误或片面内容时，追责链条往往在开发方、校方与技术提供商之间循环推诿，难以落地。现行制度多以宽泛条款应对专业性极强的技术挑战，缺乏针对教学场景的细化的伦理规范。

思政课教学承担着价值引导与人格培育的职责，数据应用必须服务于育人目标，而不能使教学关系被数据关系所替代。若缺乏明确的伦理约束与制度保障，技术应用越深入，教学伦理风险就越突出。因此，在推进生成式人工智能与思政课教学融合的过程中，必须同步建立数据保护、算法透明与责任追溯机制，使技术应用始终运行在伦理规范的轨道上。

3. 生成式人工智能时代高校思政课教学的基本前提

在分析生成式人工智能给高校思政课教学带来的多维挑战之后，需要明确应对这些挑战的基本前提。技术应用不能仅从效率或形式出发，而应首先回答教学的根本方向问题。高校思政课教学的基本前提包括四个方面：教育性前提，即育人目标与技术应用的统一；主体性前提，即教师主导与学生能动的统一；发展性前提，即全面发展与素养提升的统一；规范性前提，即伦理约束与制度保障的统一。这四个前提相互关联，共同构成技术应用的基本框架。

3.1. 教育性前提：育人目标与技术应用的统一

在生成式人工智能深刻影响教学方式的背景下，高校思政课教学必须首先明确其教育性前提。思政课的本质规定性在于育人，而非知识传递本身。它承担着帮助学生形成正确价值判断、提升思想素养、促进全面发展的任务。这一任务不因技术条件的变化而改变。生成式人工智能可以改变教学内容的呈现方式、教学互动的组织形式以及教学评价的数据来源，但不能改变思政课教学的育人目标。教学手段的更新，应当服务于育人目标的实现，而不是替代或削弱这一目标。明确这一前提，是讨论技术与教学融合问题的逻辑起点。

生成式 AI 在带来内容生产便利的同时，也因其数据依赖、算法黑箱等技术特性，以及对人类价值体系理解的固有局限，可能生成偏离育人导向的内容。算法并不理解价值，它只能根据训练数据与参数权重输出文本。当生成内容涉及思想观念、价值判断和社会议题时，就可能出现片面、偏颇甚至错误的信息。这便凸显了以育人目标统摄技术应用的必要性。育人目标为评估、运用和规范生成式 AI 提供了基本标尺。技术应用是否有利于学生的认知深化，是否有助于价值引导，是否服务于学生的全面发展，应当成为判断其合理性的核心依据。凡不符合育人目标的技术应用，即便效率再高、形式再新，也应加以调整或限制。

坚持教育性前提，要求在教学实践中始终把育人放在首位。教师需要从育人目标出发，判断哪些环节适合引入生成式 AI，哪些环节必须由教师主导。内容生成、答疑辅导、数据分析等环节可以借助技术提高效率，但价值引导、思想交流、人格示范等环节仍须由教师承担。技术应用的范围与方式，应当依据育人目标加以界定。同时，教育性前提也要求对技术应用进行持续反思。教学效果的衡量不能只看学生获取信息的数量或课堂互动的频率，更要看学生是否形成了正确的价值判断和独立思考能力。唯有如此，生成式人工智能才能在思政课教学中发挥积极作用，而不至于偏离育人方向。教育性前提的确立，为后续讨论主体性、发展性与规范性前提奠定了基础。

3.2. 主体性前提：教师主导与学生能动的统一

高校思政课教学的基本前提之一，是坚持教师主导与学生能动的统一。生成式人工智能的引入，不

能改变教师与学生在教学中的主体地位，而应服务于二者主体性的更好发挥。明确这一前提，是判断技术应用是否合理的重要依据。教师主导是思政课教学不可替代的基本要求。教师不仅承担知识讲授任务，更承担价值引导、思想交流与人格示范的职责。生成式 AI 可以辅助知识检索、答疑反馈、案例分析等环节，但无法替代教师对教学目标的把握、对教学过程的调控和对学生思想动态的判断。若教师过度依赖技术，将教学决策权让渡给算法，就可能从教学主导者变为技术执行者，教学的方向性与思想性也会受到影响。因此，教师必须始终掌握教学目标设定、教学内容审核和教学效果评价的主导权。技术工具的使用，应当以增强而非削弱教师的主导作用为前提。

学生能动是思政课教学取得实效的关键。思政课的目标不是让学生被动接收信息，而是引导学生形成独立思考与价值判断能力。生成式 AI 的即时应答功能，可能使学生习惯于快速获取答案，减少深度思考与主动探究。若学生长期依赖算法推荐与智能问答，就容易停留在认知舒适区，批判性思维与自主学习能力难以得到锻炼。因此，在技术应用中必须保障学生的学习主体地位，引导学生从被动接受转向主动参与。教师应设计需要辨析、讨论与反思的学习任务，使学生在人机互动中保持独立思考，而不是让算法替代学生的认知过程。

教师主导与学生能动并非相互对立，而是相互支撑。教师主导为学生的能动学习提供方向与保障，学生能动则是教师主导效果的具体体现。生成式人工智能的应用，应当有助于强化这一关系，而非削弱任何一方。技术可以承担重复性、机械性工作，为教师腾出更多时间关注学生的思想需求；也可以为学生提供个性化资源，支持其自主探究。但无论技术如何发展，教师的价值引导作用和学生的主动建构过程都不能被替代。坚持主体性前提，就是要在技术应用中始终维护教师与学生的双主体地位，使二者在协同互动中共同实现思政课的教学目标。

3.3. 发展性前提：全面发展与素养提升的统一

发展性前提意味着，技术应用不能只追求知识传递效率或课堂形式新颖，而应服务于学生思想素养、认知能力与人格品质的整体提升。判断技术应用是否合理，关键看它是否有利于学生的长远发展，是否有助于实现全面发展与素养提升的统一。面对算法推荐可能导致的认知窄化和价值迷失，思政课教学需要强化学生的价值主体性建构。算法根据偏好推送信息，容易使学生长期接触单一观点，削弱其对多元信息的辨别能力。思政课应着力培育学生独立的价值判断力，使其在复杂信息环境中保持清醒认识，形成对社会责任、公共价值和自身发展的稳定理解。技术应用不能替代价值内化的过程。教师应引导学生穿透技术表象，理解信息背后的价值导向与立场逻辑，而不是把算法输出当作现成结论。只有学生具备自主判断能力，技术才能真正服务于其思想成长。

智能时代的发展性前提还要求健全学生的核心素养。除传统德智体美劳素养外，算法素养与数字伦理应成为重要内容。学生既要掌握基本技术原理，理解数据采集、算法推荐和内容生成的大致过程，防止被数据操控；也要养成对人工智能应用的批判性思维，包括识别深度伪造、警惕算法偏见、反思技术垄断等能力。这些素养不是附加技能，而是全面发展在数字环境中的具体体现。它们帮助学生理解技术运行逻辑，也帮助学生形成抵御技术支配的意识和能力。素养提升与全面发展因此相互支撑。

保障数字时代的全面发展条件，同样属于发展性前提的内在要求。人工智能驱动的个性化学习应服务于因材施教原则，通过精准分析学情促进教育资源优化配置。但必须规避技术路径带来的新的不平等，防止因物质条件与认知水平差异而形成数字鸿沟，警惕算法标签对学生发展潜能的限制。在此过程中，劳动教育也被赋予新的内涵。引导学生理解数据劳动的价值创造属性，认识技术平台与用户劳动之间的关系，有助于抵制技术至上主义对劳动者主体地位的消解。总之，发展性前提要求技术应用始终以学生的全面发展和素养提升为衡量标准，使生成式人工智能在思政课教学中发挥促进成长而非限制发展的作用。

3.4. 规范性前提：伦理约束与制度保障的统一

规范性前提要求伦理约束与制度保障相互配合。伦理约束提供价值边界和行为准则，制度保障提供明确规则、责任分配与监督机制。二者统一，才能使技术应用始终服务于教学育人目标，而不是反过来支配教学。伦理约束的核心是坚持以人为本和教育性原则。技术应用应尊重学生的人格与隐私，保障其知情权和选择权。数据采集应限定在教学必要范围内，避免过度收集和不当使用。算法推荐和内容生成应保持透明，使学生和教师能够理解其基本逻辑，并对输出结果进行判断。内容审核应关注知识准确性与价值导向，防止错误信息和片面观点进入课堂。教师在使用生成式 AI 时，应承担内容把关责任，不能将教学判断完全交给算法。伦理约束不是抽象原则，而应落实到具体教学行为中。

制度保障是伦理约束得以实现的条件。高校应建立生成式 AI 教学应用的准入与审查机制，明确哪些工具可以使用、在哪些环节使用、需要满足哪些条件。应建立数据保护制度，对学生的学习行为、思想动态等信息分类管理，明确授权、存储和使用规则。应建立算法透明与责任追溯机制，对生成内容保留必要记录，明确开发方、校方、教师和技术提供商各自的责任。还应建立监督反馈渠道，使教师和学生对技术应用中的问题能够及时反映并得到处理。制度保障越具体，伦理约束越可能落地。

本文的伦理框架与 OECD、欧盟等国际组织的 AI 教育伦理准则既有共通之处，也存在差异。OECD 强调人类监督与决策权保留，要求 AI 在教学中仅发挥有限且明确的作用，受影响者有权挑战算法决策 [1]；欧盟《人工智能法案》将教育领域 AI 应用列为高风险类别，要求透明度、数据质量与人类监督，并禁止情感推断类 AI [6]。本文提出的数据保护层与算法透明层在原则上与之高度一致，均强调人类最终决策权。差异在于，OECD 侧重政策协调与教育系统响应能力，欧盟侧重法律规制与制度约束，而本文框架更强调教师在教学场景中的内容审核主导权与价值引导职能。此外，国际研究亦指出 AI 系统可能存在算法偏见与结构性不平等[7]，这与本文对数据滥用和算法隐性控制的担忧形成呼应。本文框架的补充价值在于，将国际伦理原则转化为教学层面的具体操作机制，使抽象准则得以在课堂中落地。

4. 生成式人工智能时代高校思政课教学的创新路径

在明确生成式人工智能时代高校思政课教学的基本前提之后，需要进一步探讨技术融合的具体路径。基本前提为技术应用确立了方向与边界，创新路径则将前提要求转化为可操作的教学实践。创新路径主要包括四个方面：以教师主导为核心构建人机协作教学新形态，以内容审核为重点建立技术过滤与教师把关机制，以三维融合为路径优化认知诊断与沉浸体验模式，以伦理制度为保障完善数据保护与算法透明体系。四者相互衔接，共同构成技术应用的实践框架。

4.1. 以教师主导为核心构建人机协作教学新形态

生成式人工智能融入高校思政课教学，必须坚持教师主导这一核心原则。技术工具应被定位为教师能力的延伸载体，而不是教学行动的替代主体。马克思曾指出，“思想要得到实现，就要有使用实践力量的人”[8]，教师与技术的关系，应当遵循以人为本的原则，使技术服务于教学目标的实现。基于教学分工的实际需要，教师与生成式 AI 可以形成三级协作框架。在 AI 替代层，由技术承担资源检索、知识梳理、学习行为记录、互动反馈等机械性劳动，减轻教师的重复性工作负担。在人机协同层，由教师主导高阶任务，运用批判性思维设计、输入和调整指令，监督生成内容是否符合教学要求。在人类专属层，

[1]OECD 关于 AI 与教育的最直接权威文件是《OECD Digital Education Outlook 2023》第 16 章“Opportunities, guidelines and guardrails for effective and equitable use of AI in education”，该章由 OECD 秘书处与 Education International 联合制定，系统阐述了人类监督、决策权保留等原则，并明确建立在 OECD 2019 年《人工智能理事会建议》基础之上。
https://www.oecd.org/en/publications/oecd-digital-education-outlook-2023_c74f03de-en/full-report/opportunities-guidelines-and-guardrails-for-effective-and-equitable-use-of-ai-in-education_2f0862dc.html

教师坚守价值塑造、思想引导与情感交流等核心职能，引导学生在技术输出中识别算法偏见，形成独立判断能力。

这种分层协作要求教师通过能力再造掌握三项主导权。第一，培养算法辨识能力。教师应能够界定内容生成的边界，追溯模型决策的基本逻辑，开展批判性测试以识别信息偏差。部分高校的实践表明，这类能力可以将技术风险转化为思辨教学资源。第二，培养价值引导能力。教师可以利用开源模型参数可调的特性，在内容层注入思政课教学案例库，在逻辑层强化学科方法论权重，在输出层建立价值校准机制。第三，培育人机协作的教学设计能力。教师应构建完整教学闭环，如课前利用 AI 分析学情差异并生成分组方案，课中依托虚拟场景展开“技术异化”主题辩论，课后依据算法评估实施精准辅导。在此过程中，教师必须全程主导三个关键节点：教学目标的价值定向、实施过程的思辨引导、产出内容的内容审核。

为保障教师主导权落到实处，需要建立对应的技术治理机制。通过内容价值校准系统维护价值引导主导权，防止算法弱化课程的教育属性；借由决策链可视化追溯工具保障算法运行知情权，减少技术黑箱对教师掌控力的影响；建立人工复核优先介入机制，巩固教师的教学决策终审权，防止技术理性替代价值判断。这最终指向教师培育体系的转型。在培训框架中增设技术哲学研修，帮助教师理解人工智能的基本原理；增设算法审计实训，使教师掌握内容偏差检测技术；增设协作设计工作坊，开发适用于思政课教学的应用方案。通过跨学科联合教研，使教师既具备理论素养，又具备技术驾驭能力，真正实现从技术赋能到巩固教师主导地位的战略提升。教师始终是技术应用体系的建构主体与价值源泉。

4.2. 以内容审核为重点建立技术过滤与教师把关机制

生成式人工智能在思政课教学内容生成中的应用，既提高了资源生产效率，也带来了内容质量与价值导向方面的风险。仅依靠技术过滤或仅依靠教师人工审核，都难以有效应对。因此，需要建立技术约束与人工主导相结合的双轨机制，“在价值引领层面实现‘赋能’与‘赋善’相融”[9]，使内容供给既保持效率，又符合教学要求。其核心框架包括四个相互衔接的环节。

高质量语料库是内容审核的基础。高校应组织专家团队对思政课经典文献、权威教材和公开政策文本进行结构化标注，建设适用于教学场景的专用语料库。语料库应涵盖基本理论、重要概念、典型案例和时事素材，并建立定期更新机制，及时吸收公开权威信息，确保训练数据来源可靠、内容准确。语料库建设应避免单纯追求数量，而应注重内容的代表性、准确性和教学适用性，为后续模型训练提供可靠基础。

在此根基上推进专用模型开发。利用模型压缩技术实现高校本地化部署，根据院校特色设置差异化知识权重，如理工院校侧重科技发展案例，民族地区院校侧重民族团结素材。同时嵌入三类技术屏障：基于关键词匹配的实时过滤系统，拦截明显错误信息；依托语义分析的情感倾向识别模块，过滤隐性价值偏差；采用决策链追溯工具，防控系统性错误逻辑的渗透。技术过滤的作用是初步筛选，而非最终裁决。

核心环节在于建立人机协同审核流程。首轮由 AI 执行基础校验，围绕知识完整性与事实准确性进行比对。次环节由语义倾向引擎评估内容立场是否恰当，识别表述中的倾向性问题。最终由思政课教师行使终审权。教师重点研判算法难以识别的语境隐喻，校准重要理论阐释的精确度，并对潜在风险议题进行标记，形成动态监测清单。教师审核不是简单通过或否决，而是对内容进行教学化改造，使其符合课程目标与学生认知水平。

机制运行还需要配套动态更新标准。技术层面建立热点议题捕捉机制，自动抓取社会焦点；教学层面设置内容更新论证程序，严控新增教育主题；管理层面建立与公开政策文本的联动响应通道。教师是

双轨机制有效运行的关键。通过舆情分析实训，教师可掌握风险预判方法；在伦理审查培训中，可研习算法公正的判别准则；经模块化冲突管理演练，可提高价值引导精度。双轨机制的目标是将技术生产力转化为教学效能：AI 引擎保障内容传播的广度与效率，教师的专业判断则决定教学内容的思想深度与价值纯度。

4.3. 以三维融合为路径优化认知诊断与沉浸体验模式

生成式人工智能推动思政课教学创新，需要通过数据驱动的精准供给、情境化的沉浸体验、伦理自觉的素养培育三个维度融合，重塑以学生全面发展为目标的教学模式。认知诊断与供给适配是第一个维度。运用 AI 嵌入的自然语言理解引擎，解析讨论区发言、作业表述、互动日志等学习行为数据，可以构建动态认知画像，标记学生的知识结构化程度、价值判断稳定性和偏好学习模态。对基础薄弱的学生、具备一定基础的学生或者抽象思维较强的学生，可通过分层适配，使不同起点的学生都能在各自最近发展区内获得有效提升。这一维度的关键在于，数据诊断服务于教学决策，而非替代教学决策。沉浸式学习空间的重构是第二个维度。依据身体参与塑造思维的教学原理，通过 VR/AR 技术构建历史与现实情境，使抽象理论转化为可感知的叙事。内容维度要求情境素材经三级审核：学科专家核验史实准确性，教育心理学专家评估情感触动效果，技术工程师调试多模态交互逻辑。典型案例如历史文化情境沉浸课，通过体感设备模拟特定环境，引导学生在体验中理解艰苦奋斗精神。沉浸体验的价值不在于技术新颖，而在于是否服务于学生的认知深化。数字素养教育的制度化培育是第三个维度。在理论认知层，可开设生成式人工智能与劳动价值论相关课程，引导学生审视数据生成链条中的人类劳动，揭示其背后的劳动异化实质。在技能训练层，设置算法偏见识别工作坊，引导学生剖析平台推荐机制中隐含的价值偏好。在实践规范层，建立学生 AI 应用自律准则，如禁止通过刻意模糊价值立场、预设价值中立前提等方式向人工智能工具进行诱导性提问，以此消解课程内容的价值引导功能。

文华学院在思政课教学中的核心做法与本文提出的认知诊断、沉浸体验、素养培育三维融合路径具有内在一致性。在认知诊断维度，文华学院依托学习通后台数据构建学情画像，教师可精准掌握学生的预习时长、答题正确率、讨论活跃度等信息，定位理论知识的薄弱环节与兴趣点，坚持“以学定教”。在沉浸体验维度，学院建成了虚拟仿真思政课体验教学中心，通过智能化学习设备为学生提供交互式学习体验，使思政课教学从“指尖”直抵“心间”。在素养培育维度，课程打破传统“教师主讲、学生倾听”模式，围绕“人工智能是否具有意识”“996 与个人全面发展是否存在冲突”等现实议题引导学生开展小组辩论与师生对话。不可否认的是，实施过程中也遇到如下困难：部分教师对智能工具的操作不够熟练，学情数据的教学化转化需要额外投入时间；虚拟仿真设备的课堂调度与教学内容衔接尚需磨合；学生在对话思辨环节的参与深度参差不齐，个别学生仍习惯于被动接受。从反馈看，课程课堂出勤率稳定在 99%以上，互动参与率达 100%，学生满意度达 97%以上。初步实践表明，三维融合路径在提升教学实效方面具有可行性，但需要持续完善教师培训与技术支撑体系。

4.4. 以伦理制度为保障完善数据保护与算法透明体系

有学者指出，应“以技术完善与数字素养提升助推伦理与隐私保护”[10]。生成式人工智能在思政课教学中的深度应用，需要通过完善的伦理审查框架与责任追溯机制加以规范。核心架构包括数据保护层、算法透明层与应用监管层三个层面。数据保护层实施学生个人信息分类保护制度。对生理数据、思想动态与行为记录分别设定保密等级，明确不同类别数据的采集范围、存储期限和使用权限。推行教育数据调用“双重授权”原则，即学生知情同意前置与教育主管部门备案核准相结合，并依托区块链技术构建不可篡改的授权存证链条，确保数据使用可追溯、可审计。算法透明层应建立渐进式披露机制，包括三

个等级：基础级要求公开训练数据的基本构成，如教学文献在语料中的占比；解释级要求输出溯源路径，如重要理论表述须标注公开权威文献依据；审计级要求强制开放价值参数接口，供专家组检测内容生成中可能存在的价值偏差。应用监管层进行全周期风险管控。部署生成内容数字水印系统，包含模型版本、生成时间、操作者标识等可追溯信息，构建“事前准入–事中纠偏–事后追责”的闭环体系。

“思想政治教育数智化转型需要组织力量的有效介入”[11]，制度建设需重点完善三类规范。健全风险防控规程，设定多阶响应阈值。依据语义分析结果，当内容价值偏差度超过设定阈值时触发自动暂停机制，事实错误率超过设定标准时启动人工复查，敏感话题讨论频次异常波动时推送专项警报，教师拥有最高权限实施强制干预。构建法律权责体系，明确研发方需履行算法价值观年度审计义务，输出价值立场符合性报告；院校管理者承担运行过程监管责任；教师对自主启用工具生成的内容负主体责任。确立应用负面清单，严禁 AI 执行三类操作：替代价值判断、生成敏感性结论、处理未公开文献。

在实施过程方面，清华大学自 2023 年系统部署 AI 辅助教学以来，已有 395 门课程深度融入 AI 教学实践，覆盖 AI 学伴、备课助手等十大功能场景[12]。应用扩大后，学校于 2024 年启动原则制定，由多部门协同，访谈百余名师生，分析全球 25 所高校 70 份指南，2025 年正式发布，提出主体责任、合规诚信、数据安全、审慎思辨、公平包容五大原则 [2]，禁止使用敏感、涉密或未授权数据训练 AI 模型，并要求披露 AI 使用情况。实施中的主要困难是技术快速演进与制度相对稳定之间的矛盾，静态政策易滞后，因此原则预留了持续修订空间。师生反馈认为，原则使 AI 使用有据可依。初步效果上，校内 AI 应用从试点探索转向规范运行，数据安全与学术诚信风险得到制度性约束，也为其他高校提供了参照。

5. 研究局限与未来展望

本文主要采用理论分析与案例讨论相结合的方法，对生成式人工智能时代高校思政课教学的挑战、前提与创新路径进行了梳理。受研究条件限制，本文存在以下不足。其一，研究方法较为单一。文章以文献分析和逻辑推演为主，虽结合文华学院、清华大学等高校的初步实践，但案例数量有限，缺乏大规模问卷、访谈或实验数据的支撑，结论的普适性仍需进一步检验。其二，方案设计带有一定理想化色彩。文中提出的教师主导机制、内容审核流程、三维融合模式与伦理制度框架，均以较为充分的师资条件、技术支持和制度保障为前提。在部分高校，尤其是一些资源有限的院校，落地难度可能较大。其三，对技术迭代的跟进不够及时。生成式人工智能发展迅速，模型能力与应用形态不断变化，本文讨论的风险与路径可能难以完全覆盖未来出现的新情况。其四，跨学科整合尚不充分。文章涉及教育学、技术哲学、伦理学与法学等多个领域，但各部分之间的理论衔接仍有深化空间。

未来研究可从以下方向展开。第一，教师角色的深层演变值得持续关注。随着生成式 AI 承担更多知识供给与教学辅助功能，教师可能从知识传授者进一步转向价值引导者、人机协作设计者和伦理把关者。这一转变对教师素养、培训体系和评价机制提出了新要求，需要长期追踪研究。第二，需要更多实证研究。可通过多校比较、课堂观察、师生访谈和准实验设计，检验不同创新路径的实际效果，分析影响技术应用成效的关键因素。第三，跨学科研究亟待加强。应对生成式人工智能带来的复杂挑战，需要教育学、计算机科学、伦理学、法学、心理学等多学科协同，共同开发适用于思政课教学的技术工具、伦理规范和评价标准。第四，应关注技术应用的公平性问题。不同地区、不同院校之间的数字鸿沟可能影响技术赋能的实际效果，未来研究应探讨如何避免技术应用加剧教育资源不平等。总之，生成式人工智能与思政课教学的融合仍处于探索阶段，需要在实践中不断反思、调整与完善。

[2]清华大学官网于 2025 年发布《清华大学人工智能教育应用指导原则》，系国内高校首个覆盖全校、分场景的 AI 教育应用规范，明确“积极而审慎”的基本立场及五大核心原则。https://www.tsinghua.edu.cn/info/1176/122770.htm

参考文献

[1] 孙伟平. 智能系统的自主活动是否创造价值? [J]. 思想理论教育, 2025(4): 20-26.

[2] 马克思, 恩格斯. 马克思恩格斯文集: 第八卷[M]. 北京: 人民出版社, 2009: 358.

[3] 张雪. 生成式人工智能介入思想政治教育主客体关系: 机理、困境与路径[J]. 思想教育研究, 2026(8): 126-134.

[4] 史宏波, 李晶晶. 思想政治教育知识生产的新理念与新路向[J]. 思想理论教育导刊, 2026(6): 126-136.

[5] 孟偲. 大语言模型影响下思想政治教育话语传播的机遇挑战与优化路径[J]. 思想政治教育研究, 2026, 42(3): 135-143.

[6] Saarela, M., Gunasekara, S. and Karimov, A. (2025) The EU AI Act: Implications for Ethical AI in Education. In: Chatterjee, S., vom Brocke, J. and Anderson, R., Eds., *Local Solutions for Global Challenges*, Springer, 36-50. https://doi.org/10.1007/978-3-031-93979-2_3

[7] Heeg, D.M. and Avraamidou, L. (2026) Demystifying AI: The Urgency of a Critical Stance on the Use of AI Systems in Education. *Prospects*. https://doi.org/10.1007/s11125-026-09760-4

[8] 马克思, 恩格斯. 马克思恩格斯文集: 第一卷[M]. 北京: 人民出版社, 2009: 320.

[9] 辛向阳, 毛婷婷. 数智赋能思想政治教育的价值建构、现实困境与路径探索[J]. 贵州大学学报(社会科学版), 2026, 44(4): 23-32.

[10] 陈浩天, 王晶晶. 生成式人工智能嵌入教育场景的风险具象及规制逻辑[J/OL]. 湖北民族大学学报(哲学社会科学版), 1-9. https://doi.org/10.13501/j.cnki.42-1328/c.20261014.003, 2026-11-04.

[11] 罗红杰, 陈慧琳. 思想政治教育数智化转型中“技术-组织”调适关系及其优化[J]. 思想教育研究, 2026(6): 36-44.

[12] 中国科学报. 清华大学首发人工智能教育应用指导原则——不仅划出“红线”, 更要点亮“绿灯” [EB/OL]. https://cj.sina.cn/article/norm_detail?url=https%3A%2F%2Ffinance.sina.com.cn%2Ftech%2Froll%2F2025-12-02%2Fdoc-infzizmr3001782.shtml&vt=4, 2025-12-02.