

# 生成式AI在个性化学习资源智能化推荐中的伦理风险与治理框架

邹柳

黄冈师范学院教育学院, 湖北 黄冈

收稿日期: 2025年6月10日; 录用日期: 2025年7月9日; 发布日期: 2025年7月17日

## 摘要

本文以生成式AI在个性化学习资源推荐中的应用为研究对象, 探讨其引发的伦理风险及治理路径。随着生成式AI技术在教育领域的快速渗透, 数据隐私泄露、算法偏见、学术诚信危机等问题日益凸显, 亟需系统性治理框架。基于案例分析与文献综述, 研究提出多维治理方案: 技术层面融合联邦学习与差分隐私保护数据安全, 政策层面构建分级分类监管体系, 教育层面通过教师AI素养培训与学生数字公民教育强化主体责任, 行业层面倡导算法公平优化与开源社区共建。该框架为平衡AI教育创新与伦理风险提供了实践参考, 对推动教育智能化健康发展和教育公平具有重要现实意义。

## 关键词

生成式AI, 个性化学习资源, 伦理风险, 治理框架, 联邦学习, 算法公平

# Ethical Risks and Governance Framework of Generative AI in Intelligent Recommendation of Personalized Learning Resources

Liu Zou

College of Education, Huanggang Normal University, Huanggang Hubei

Received: Jun. 10<sup>th</sup>, 2025; accepted: Jul. 9<sup>th</sup>, 2025; published: Jul. 17<sup>th</sup>, 2025

## Abstract

This paper takes the application of generative AI in personalized learning resource recommendation as the research object, and discusses the ethical risks and governance paths caused by it. With the rapid penetration of generative AI technology in the education field, problems such as data

privacy leakage, algorithmic bias, and academic integrity crisis are becoming increasingly prominent, and a systematic governance framework is urgently needed. Based on case analysis and literature review, this paper proposes a multi-dimensional governance scheme: integrating federated learning and differential privacy to protect data security at the technical level, building a hierarchical and classified supervision system at the policy level, strengthening the main responsibility through teachers' AI literacy training and students' digital citizenship education at the education level, and advocating fair optimization of algorithms and co-construction with open source communities at the industry level. This framework provides a practical reference for balancing AI education innovation and ethical risks, and is of great practical significance for promoting the healthy development of intelligent education and educational equity.

## Keywords

Generative AI, Personalized Learning Resources, Ethical Risk, Governance Framework, Federated Learning, Algorithmic Fairness

Copyright © 2025 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

随着生成式 AI 技术在教育领域的快速渗透,其在个性化学习资源智能化推荐中的应用日益广泛,深刻改变着教育模式与学习生态。然而,这一技术革新也带来了一系列不容忽视的伦理风险:数据隐私泄露与滥用风险因学生敏感信息的大规模采集而凸显,算法偏见可能加剧教育不公平,学术诚信危机随着 AI 代写等现象蔓延,师生关系异化与情感疏离因 AI 介入教学互动而隐现,教育公平失衡与技术鸿沟则因技术实施成本差异而扩大。这些问题不仅威胁学生权益与教育公正性,更对教育的人文本质与价值导向提出挑战。

教育的价值导向是以培养全面发展的人为核心,注重知识传授、能力培养与价值观塑造,倡导公平、包容、创新,追求人的自由发展与社会和谐进步。将价值导向融入生成式 AI 应用,需在技术设计上确保算法公平,避免资源分配失衡;在内容生成中融入多元价值观,引导学生辩证思考;同时建立伦理审查机制,对 AI 输出内容进行价值观校准,让生成式 AI 成为传递教育价值、助力学生成长的有益工具。

在此背景下,系统探讨生成式 AI 在教育推荐场景中的伦理风险及其治理路径具有迫切的现实意义。本文基于案例分析与文献综述,从技术、政策、教育机构、行业等多维视角构建治理框架,提出融合联邦学习与差分隐私的隐私保护方案、可解释性算法设计、分级分类监管体系、教师 AI 素养培训、算法公平性优化等具体策略,旨在为平衡 AI 教育创新与伦理风险提供实践参考,推动教育智能化向公平、透明、有温度的方向健康发展。

## 2. 生成式 AI 在教育场景中的核心伦理风险

### 2.1. 数据隐私泄露与滥用风险

当前,个性化学习推荐系统所面临的核心挑战之一在于数据隐私泄露及其滥用风险。生成式人工智能在实现精准推荐时,需采集大量涉及学生敏感信息的数据,包括学习行为轨迹、答题记录、知识掌握程度以及个人偏好等。一旦此类数据泄露或遭到滥用,不仅会侵犯学生的隐私权,还可能被用于商业牟利或其他非法活动(如诈骗)。例如,未经加密的学习行为数据可能通过中间人攻击被第三方截获,或因平

台安全漏洞而遭到黑客窃取。此外，部分教育机构可能出于商业利益考虑，在未获得用户明确授权的前提下与其他机构共享数据，此举直接违反了《通用数据保护条例》(GDPR)、《中华人民共和国个人信息保护法》等相关法律法规的要求。

为了应对这些风险，研究人员提出了一系列数据安全保护措施，也就是采用联邦学习框架[1]。不过，联邦学习这种方法并不是完美无缺的。最新的研究结果显示，恶意的参与方有可能通过精心设计的梯度反演攻击方式，从共享的梯度信息当中推断出原始数据的特征。除此之外，模型的参数本身也可能隐藏着用户的行为模式，存在被推理攻击的风险。

## 2.2. 算法偏见与歧视性推荐

AI 模型的训练数据常常反映出现实世界里的结构性偏见，像性别、种族、地域或者经济背景方面的不平等情况。这些偏见可能在数据标注、特征提取和算法决策这些环节中，被悄悄编入推荐系统。就拿教育数据来说，城乡教育资源长期存在差异，这会让算法更愿意给城市学生推荐高阶学习内容，而农村学生就只能被困在基础学习路径中，进而让教育不公平的现象变得更严重[2]。美国教育部 2023 年有一份报告提到，某个自适应学习平台因为训练数据里 STEM 学科的女性样本不够，导致系统更倾向于给男生推荐理科进阶课程，这种情况被叫做“算法再生产偏见”。

为了解决这个问题，有研究提出了一种基于时空特征的动态去偏方法。这种方法通过三个技术路径来实现，分别是数据层去偏、算法层优化和建立动态反馈机制。不过，我们还是要留意技术手段存在的局限性。对抗学习办法虽然能有效减少明显的偏见，但在识别隐藏的偏见时，比如文化资本差异导致的兴趣偏好偏差，还是存在困难。所以，未来的研究需要把教育学理论结合进来。在技术方面，为了消除偏见，要采用对抗学习方法，例如构建一个能区分用户兴趣和敏感特性的对抗网络。

## 2.3. 学术诚信危机与创造力抑制

ChatGPT 和其他生成式 AI 工具在教育领域的广泛使用，正给学术原创性和学生核心能力的培养带来严重挑战。研究发现，很多学生用这类工具来代写作业、生成标准化答案，甚至完成学术论文。美国高校的调查数据显示，大约 33% 的大学生公开承认会用 AI 帮忙完成学业任务，其中文科类课程的 AI 使用比例高达 42%。

现在还有个更让人担心的情况：主流的 AI 检测工具(也就是 AI 写作识别功能)存在明显的技术问题，它们的假阳性率能达到 15%。这意味着每 7 篇被标记为 AI 生成的作业里，可能就有 1 篇是被误判的。这种技术上的不足很容易引发学术争议，还会造成信任危机。

从教育学的角度看，学生长期依赖生成式 AI，可能会让关键能力出现全面退化。具体体现为：批判性思维变弱、知识建构能力缺乏、对学术伦理的认知变得模糊[3]。实际案例显示，过度使用 AI 写作工具会让学生的论证能力下降。

## 2.4. 师生关系异化与情感疏离

AI 系统可以快速准确地回答学生的问题，提供标准化的评估和反馈，这大大提高了教学的效率和便捷性。不过，这种技术的广泛应用也带来了一些不能忽视的问题。当 AI 代替了部分原本属于教师的工作时，教师和学生之间的直接互动机会就大幅减少了。教师不再像以前那样经常和学生面对面交流、解答疑问，也没办法给予他们个性化的指导和鼓励。互动次数的减少，会让师生之间的情感联系逐渐变得疏远，教育也不再是充满人文关怀和情感交流的活动，而是慢慢变成一种机械的、按程序进行的过程[4]。

研究显示，虽然 AI 技术在处理数据和逻辑推理方面表现得很好，但在模仿人类情感和对文化的敏感

理解方面，存在明显的不足。教师的同情心和对文化的深刻理解，是教育过程中不可缺少的重要因素。教师能够根据学生的文化背景、个性特点和情感状态，灵活调整教学方法和策略，给予学生情感上的支持和理解。可是，AI系统因为没有真正的情感体验和文化认知能力，无法像教师那样敏锐地发现学生的情感变化，也不能提供真正意义上的心理辅导[5]。

如果算法推荐系统过度追求提高学习效率，这种情况可能会变得更严重。算法推荐系统通常会根据学生的学习数据和行为模式，为他们提供个性化的学习内容和建议，帮助学生更高效地学习。但这种以效率为目标的推荐系统，可能会过于关注知识的传授和技能的培养，而忽略学生的情感需求。它可能会把学生仅仅当作学习对象，忽视他们作为个体的情感和心理状态。这种对学习效率的过度追求，可能会让学生在过程中感到孤独和无助，缺乏情感上的共鸣和支持，进而影响他们的学习动力和心理健康。

## 2.5. 教育公平失衡与技术鸿沟

在教育语境下，“公平”的核心含义是确保所有学生，无论其家庭经济状况、地域、种族、性别等因素如何，都能获得平等的教育机会、资源与发展可能[6]。其衡量标准涵盖多个维度：资源分配上，学校硬件设施、师资力量、教育技术投入等应尽量均衡，避免因地域或学校性质产生巨大差距；过程公平体现在教学方法、评价体系对不同学习基础的学生一视同仁，不设不合理门槛或差别对待；结果公平则表现为学生在学业成就、能力培养、未来发展机会上的相对均衡。

生成式AI技术在教育领域的广泛应用带来了较高的实施成本，这可能导致资源分配的不平等。具体来说，私立学校和经济发达地区的学校有更充足的财政支持和先进的技术设备，往往能更早、更全面地引入生成式AI技术，成为主要受益者[7]。例如，这些学校可以聘请专业的技术团队来维护和优化AI系统，为学生提供个性化的学习体验。但这种优势可能会进一步拉大教育领域的技术差距，让资源较少的公立学校和欠发达地区的学校难以获得同等质量的教育技术资源，进而加剧教育不平等现象。

此外，生成式AI的算法推荐系统可能会产生“信息茧房”效应，在一定程度上让学生的学习路径变得固定[2]。比如，对于学习基础较弱的学生，算法可能会一直推荐低难度的学习内容，避免他们感到沮丧或挫败。这种做法虽然短期内能帮助学生保持学习积极性，但长远来看，可能会限制他们的学习潜力和进步空间。因为这些学生可能会习惯接受简单内容，缺少挑战高难度知识的机会，难以实现知识的深入拓展和能力的全面提升。

## 3. 治理框架的多维度构建

### 3.1. 技术治理：隐私保护与算法透明

#### (1) 联邦学习与差分隐私的融合应用

为了在保护数据隐私的同时实现高质量的个性化推荐，采用横向联邦学习与纵向联邦学习架构，并结合差分隐私技术，是一种有效的解决方案。

横向联邦学习(跨机构数据协作)允许这些机构在不共享原始数据的情况下，共同训练一个全局模型，从而提升模型的性能和泛化能力。在横向联邦学习中，各参与方(如不同学校)在本地训练自己的模型，并将模型更新后的梯度信息通过加密方式传输到中央服务器。中央服务器聚合这些梯度信息后，更新全局模型，并将更新后的模型参数分发回各参与方。这一过程确保了数据不出本地，同时实现了模型的协同优化[1]。

纵向联邦学习(跨数据类型整合)通过整合这些不同类型的异构数据，提升推荐系统的多样性和准确性。在纵向联邦学习中，各参与方(如不同数据源的提供者)通过加密方式共享数据特征，确保数据的安全性和隐私性。中央服务器负责协调和聚合这些特征，生成全局模型，并将更新后的模型参数分发给各参

与方。这一过程不仅提升了推荐系统的性能，还确保了数据的隐私保护。

差分隐私是一种强大的隐私保护技术，通过在数据中添加噪声，确保单个数据点的隐私性。在联邦学习中，当隐私预算( $\epsilon$ )小于或等于 3 时，推荐系统的精度达到最优，同时确保数据隐私性[8]。在联邦学习的每个训练周期中，各参与方在本地模型更新时添加高斯噪声，然后将更新后的梯度信息传输到中央服务器。中央服务器聚合这些带噪声的梯度信息，更新全局模型，并将更新后的模型参数分发回各参与方。通过这种方式，差分隐私技术在保护数据隐私的同时，确保了模型的训练效果。

在个性化学习推荐系统中，用户(如学生)的查询内容可能包含敏感信息。为了确保用户查询内容不被服务端追踪，推荐系统可以引入隐私保护信息检索(PIR)技术。PIR 技术允许用户在不泄露查询内容的情况下，从服务端获取所需的信息。PIR 技术通过加密和解密机制，确保用户查询内容在传输和处理过程中不被服务端识别。具体来说，用户将查询内容加密后发送给服务端，服务端根据加密的查询内容返回加密的推荐结果，用户再对这些结果进行解密，从而获取所需的信息。这一过程不仅保护了用户的隐私，还确保了推荐系统的正常运行。

## (2) 可解释性算法设计

智能化学习环境中资源推荐的关键影响因素，包括知识结构、易错率、个人错误率、教师推荐和学习投入时长等[9]。借助 LIME (Local Interpretable Model-agnostic Explanations) 技术，可以对推荐逻辑进行可视化展示，从而帮助教育者和学生更好地理解推荐结果的依据。LIME 技术是一种模型无关的可解释性工具，能够为复杂的机器学习模型提供局部解释。在教育推荐系统中，LIME 可以用于解释为什么向学生推荐特定的学习资源，例如数学问题。通过分析模型的局部行为，LIME 能够生成易于理解的解释，帮助用户理解推荐逻辑。例如，当推荐系统向学生推荐一个数学问题时，LIME 可以展示该问题与学生历史错误模式的相关性，从而解释为什么这个特定的问题对学生的学习最为有益。

为了进一步提升教育推荐系统的可解释性，构建了一系列透明度评价指标。这些指标从多个维度评估推荐系统的透明度，确保推荐结果不仅准确，而且易于理解和接受。主要的透明度评价指标包括：推荐结果与用户画像的匹配度、偏差检测的覆盖率、解释的可理解性、透明度的动态性。

通过构建这些透明度评价指标，教育推荐系统不仅能够提供高质量的个性化推荐，还能够增强用户对推荐结果的理解和信任。这不仅有助于提升教育推荐系统的应用效果，还能够促进教育公平，确保所有学生都能从个性化学习中受益。

## 3.2. 政策监管：法律框架与标准制定

### (1) 分级分类管理制度

基础学习行为数据，如学生的课堂参与度、作业完成情况、考试成绩等，这些属于低风险数据。通常不涉及个人隐私，可以通过匿名化处理进行分析和使用。心理健康评估数据、个人健康记录、家庭背景信息等，这些高风险数据涉及个人隐私和敏感信息，需要严格保护。在使用这些高风险数据时，必须获得伦理委员会的正式批准，并确保数据的收集、存储和使用符合严格的隐私保护标准。

为了确保教育人工智能系统的透明性和公正性，开发者被要求公开展示训练数据集的组成以及偏差修正的关键参数。这不仅有助于监管部门和公众了解系统的运行机制，还能促进开发者对算法偏见和不公平现象的自我审查和改进。例如，开发者需要公开数据集的来源、数据清洗和预处理方法、模型训练的具体参数设置等信息，以便接受外部监督和评估。

### (2) 跨部门协同监管体系

教育部、网信办和工信部共同建立了一个跨部门的审查机制，对教育领域的大型人工智能模型进行定期的安全评价。这种联合审查机制可以确保从多个角度对人工智能系统进行全面评估，包括数据安全、

隐私保护、算法公正性等方面。例如，每年对教育人工智能系统进行一次全面的安全评价，评估其是否符合国家相关法律法规和伦理标准，是否存在潜在的安全风险和隐私问题。

参照欧盟的《人工智能法案》(AI Act)为全球人工智能监管提供了重要的参考框架。该法案明确禁止在教育环境中使用“高风险”AI功能，例如基于情绪识别的学习干预。这是因为情绪识别技术可能涉及对学生情绪状态的过度监控和干预，存在隐私侵犯和心理压力的风险。我国可以借鉴这一经验，制定类似的法规，禁止在教育环境中使用可能对学生心理健康和隐私造成威胁的高风险AI功能。例如，明确规定教育机构不得使用基于情绪识别技术的AI系统进行学生情绪监控，除非经过严格的伦理审查和批准。

### 3.3. 教育机构责任：伦理实践与能力建设

#### (1) 教师 AI 素养培训

联合国教科文组织《教师人工智能能力框架》强调，教师需具备“批判性评估AI工具”和“人机协同教学设计”的能力[10]。结合这一框架，教师培训可从以下方向优化：

在模块化课程开发中，基础层增设“AI工作原理与伦理”模块，包括ChatGPT等生成式AI的技术局限性及其教学影响；应用层设计“人机协同教学法”工作坊，指导教师将AI工具整合到学科教学中。设立“AI教学伦理委员会”，参照数据隐私保护措施(如匿名化处理学生数据)，制定校本AI使用规范，审核高风险场景。

在分层能力评估方面，依据《能力框架》的“获取-深化-创造”三级水平，设计差异化培训目标：要求初级教师掌握AI工具的基本操作与风险识别(如检测ChatGPT生成的学术不端内容)；要求骨干教师能结合学科特点设计AI增强型教案(如用AI模拟历史对话，但需人工验证史实)；要求专家教师参与AI教育产品的开发与伦理审查，如“合作共善”模式，与技术开发者共同优化工具。

#### (2) 学生数字公民教育

数字公民教育聚焦于培养学生在数字时代的核心素养，其内容包含数字技术应用能力、数字伦理与法治意识、批判性思维以及社会责任担当等多个维度。在方法上，多采用理论与实践结合的模式，通过案例分析、情景模拟、项目式学习等方式，让学生在具体情境中深化认知。学生需具备“批判性使用AI”和“学术伦理”意识，将AI伦理纳入通识课程。通识课程应涵盖AI伦理的基本概念和原则，如数据隐私、算法偏见、人工智能的社会影响等。通过理论讲解和案例分析，帮助学生理解AI技术的伦理问题及其对个人和社会的影响。

培养学生对AI技术的批判意识，使其能够识别和质疑AI工具的潜在问题。例如，通过分析Deepfake技术的案例，帮助学生理解AI生成内容的真实性问题，以及如何识别和防范虚假信息。通过案例教学和讨论，培养学生的伦理决策能力。例如，讨论在使用AI工具时遇到的伦理困境，如AI生成内容的版权问题、算法偏见对特定群体的影响等，引导学生在实际情境中做出符合伦理的决策。基于技术-社会互构理论，提出从技术设计、家庭与社会责任、多方协同治理等多层面入手的治理路径，以规范情感模拟技术的发展[11]。

### 3.4. 行业自律：技术伦理与公平设计

#### (1) 算法公平性优化

在教育人工智能应用中，算法公平性是确保所有学生都能获得公平教育资源的关键。通过采用公平性约束技术，可以在推荐模型中嵌入群体平等性指标，确保不同背景的学生获得同等质量的教育资源推荐。例如，可以确保农村学生与城市学生获得同等质量的拓展资源推荐，从而缩小城乡教育差距。

在推荐算法中加入公平性约束，确保不同群体(如农村与城市学生、不同性别学生、不同种族学生等)

获得的推荐资源在质量上没有显著差异。例如，通过调整算法参数，确保推荐系统不会因学生的地理位置、家庭收入等因素而产生偏见。定期评估算法的公平性表现，通过收集和分析数据，监测不同群体的推荐结果。如果发现某些群体的推荐质量较低，及时调整算法参数，优化推荐模型，确保公平性[12]。通过实际案例分析，了解算法在不同场景下的表现，收集用户反馈，不断优化算法。例如，分析农村学生使用推荐系统后的学习效果，与城市学生进行对比，找出差距并进行针对性改进。

建立多利益相关方参与的算法审计机制，通过建立多利益相关方参与的算法审计机制，可以邀请教育专家、学生代表、技术专家等共同参与模型测试和评估，确保算法的公正性和有效性。

## (2) 开源社区共建

鼓励和支持教育大模型的开源项目，如 OpenAI 的 GPT-Edu 项目。通过开源，允许第三方审计代码和训练数据，确保算法的透明性和公正性[13]。例如，可以设立开源基金，支持开源项目的开发和维护。建立开源社区，鼓励教育工作者、技术开发者、学生等参与开源项目的开发和维护。通过社区的力量，不断优化模型，提高其公平性和有效性。例如，组织开源贡献活动，鼓励开发者提交代码改进和数据更新。促进不同机构和组织之间的合作，共享开源模型和数据资源。通过合作，可以整合各方资源，提高模型的覆盖范围和质量。

设计教育资源公平性指数(EFI)，包括多个维度，如资源覆盖范围、资源质量、用户满意度等。通过综合评估这些维度，量化推荐系统对不同群体的覆盖均衡度。通过开发 EFI，可以客观评估推荐系统在不同群体中的表现，确保资源分配的公平性。定期收集和分析数据，评估推荐系统在不同群体中的表现。通过数据分析，找出资源分配不均衡的地方，提出改进措施[14]。

## 4. 责任主体的协同机制

生成式 AI 在教育领域的应用涉及多元责任主体，他们需协同履行职责以保障技术合理运用。如图 1，具体协同机制如下。

**教育机构：**需制定校本 AI 使用政策，建立伦理审查流程，并组织师生培训。如哈佛大学要求教师提交 AI 辅助教学方案的风险评估报告，还设立了学生 AI 伦理监督员，以此规范校内 AI 应用并提升相关主体的伦理意识。

**技术开发者：**要确保算法透明性，定期发布影响评估报告，同时为用户提供数据控制权。像科大讯飞教育 AI 平台开放“偏见检测”接口，允许学校自定义公平性阈值，推动算法公平与透明。

**监管部门：**需立法规范数据采集范围，建立 AI 教育产品认证体系，并实施动态合规检查。例如欧盟通过《数字教育行动计划 2025》，强制要求教育 AI 提供可解释性报告，从政策层面加强监管。

**学生/家长：**可参与算法治理反馈，监督 AI 使用合规性，同时行使数据删除权。如澳大利亚部分学校引入家长代表参与 AI 推荐系统的优化委员会，强化了用户层面的监督与参与。

这种多元主体协同机制，通过明确各主体职责与具体案例实践，形成了覆盖政策制定、技术开发、监管执行和用户参与的完整治理链条，有助于平衡教育 AI 创新与伦理风险，推动教育智能化健康发展。

## 5. 未来治理的挑战与方向

### 5.1. 技术快速迭代与监管滞后矛盾

生成式 AI 技术发展迅速，新的模型和应用不断涌现。这种快速迭代使得监管机构难以及时跟上技术发展的步伐，导致监管措施往往滞后于技术应用[15]。并且现有监管框架和法律法规在应对生成式 AI 的复杂性和动态性方面存在不足。例如，数据隐私保护、算法偏见、内容真实性等问题的监管措施尚未完善，难以有效应对技术带来的新风险。

责任主体	核心职责	实施案例
教育机构	制定校本AI使用政策，建立伦理审查流程，组织师生培训	哈佛大学要求教师提交AI辅助教学方案的风险评估报告，并设立学生AI伦理监督员
技术开发者	确保算法透明性，定期发布影响评估报告，提供用户数据控制权	科大讯飞教育AI平台开放“偏见检测”接口，允许学校自定义公平性阈值
监管部门	立法规范数据采集范围，建立AI教育产品认证体系，实施动态合规检查	欧盟通过《数字教育行动计划2025》，强制要求教育AI提供可解释性报告
学生/家长	参与算法治理反馈，监督AI使用合规性，行使数据删除权	澳大利亚部分学校引入家长代表参与AI推荐系统的优化委员会

Figure 1. The coordination mechanism of the responsible entities  
图 1. 责任主体的协同机制

## 5.2. 全球治理协同难题

不同国家和地区在文化背景、教育理念和法律框架上存在显著差异，这使得全球范围内的治理协同面临巨大挑战。例如，某些文化可能更注重个人隐私保护，而另一些文化可能更关注技术的创新和应用。此外，各国对生成式 AI 的治理标准和监管措施不一致，导致跨国企业和技术开发者在不同国家面临不同的监管要求，增加了治理成本和复杂性。

## 5.3. 人性化教育的再定义

生成式 AI 的应用使得教育目标从传统的知识传授向个性化、多样化的教育目标转变。这种转变对传统教育理念和 method 提出了挑战，需要重新定义教育的目标和内容。生成式 AI 的推荐可能会限制学习者的自主选择，导致学习者过度依赖技术，缺乏独立思考和自主学习能力[16]。如何在技术辅助下保持学习者的自主性，是未来教育需要解决的重要问题。

## 参考文献

- [1] 李康康, 袁萌, 林凡. 联邦个性化学习推荐系统研究[J]. 现代教育技术, 2022, 32(2): 118-126.
- [2] 王佑镁, 房斯萌, 柳晨晨. 风险社会视角下教育人工智能伦理风险分类框架研究[J]. 现代远程教育, 2024(3): 28-37.
- [3] 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, 29(2): 26-35.
- [4] 张琨, 魏青松. ChatGPT 嵌入思想政治教育的伦理风险及治理路径[J]. 网络安全技术与应用, 2024(9): 102-107.
- [5] 汪金英, 潘祖贤. 隐忧与消解: 生成式 AI 教育应用的道德审视[J]. 扬州大学学报(高教研究版), 2024, 28(4): 47-57.
- [6] 扎卡里·泰勒, 王晨. 如何构建和利用大数据促进教育公平——全球化框架下的协作和应对[J]. 比较教育研究, 2025, 47(5): 41-52+62.
- [7] 刘军, 雷亮, 钟昌振, 等. 生成式人工智能的教育应用监管路线图——UNESCO《教育和研究领域生成式人工智能使用指南》解读与启示[J]. 中国教育信息化, 2024, 30(8): 13-28.
- [8] 马黛露丝, 朱海萍, 田锋, 等. 一种权衡性能与隐私保护的推荐算法[J]. 西安交通大学学报, 2021, 55(7): 117-123.

- 
- [9] 马秀麟, 梁静, 李小文, 等. 智能化学习环境下资源推荐的影响因素及权重的探索[J]. 中国电化教育, 2019(3): 110-119.
- [10] 苗逢春. 基于教师权益的自主人工智能应用——对联合国教科文组织《教师人工智能能力框架》的解读[J]. 开放教育研究, 2024, 30(5): 4-16.
- [11] 张宁, 高鹏程. 生成式人工智能情感模拟的伦理风险与治理路径: 基于技术-社会互构理论框架的分析[J]. 科学决策, 2025(2): 123-134.
- [12] 张进良, 叶求财. 大数据视阈下学习资源智能推荐模型构建[J]. 湖南科技大学学报(社会科学版), 2019, 22(4): 178-184.
- [13] 段伟文. 构建稳健敏捷的人工智能伦理与治理框架[J]. 科普研究, 2020, 15(3): 11-15+108-109.
- [14] 王佑镁, 王旦, 王海洁, 等. 基于风险性监管的 AIGC 教育应用伦理风险治理研究[J]. 中国电化教育, 2023(11): 83-90.
- [15] 罗蓉蓉, 肖攀诚. 生成式人工智能的风险审视与治理研究[J]. 咨询与决策, 2024, 4(1): 1-18.
- [16] 黄胜, 刘国权, 苏长青. 教育领域 ChatGPT 应用伦理困境与纾解对策[J]. 继续教育研究, 2024(9): 45-48.