生成式人工智能应用背景下高校教育数据隐私保护策略

谢　云
(南京信息职业技术学院,江苏南京　210023)

[摘　要]生成式人工智能技术的迭代演进为高校教育数字化转型提供了全新动能,但其数据驱动特性使教育数据隐私保护面临范式重构。本文基于数据生命周期理论,构建"风险—机理—治理"分析框架,系统解构生成式人工智能在个性化学习系统、科研协作平台、智能管理系统等场景中的隐私泄露路径。基于《中华人民共和国个人信息保护法》,通过引入隐私保护技术方案,提出包含动态脱敏机制、权限区块链管理、伦理审查委员会在内的三维防护体系。本文将技术防御与制度约束相结合,构建适配生成式人工智能技术特性的教育数据隐私保护体系,为高校数字化转型中的数据安全治理提供理论参考与实践范式。

[关键词]生成式人工智能;高校教育数据;隐私保护;数据安全
[中图分类号]G434;D923　　[文献标识码]A　　[文章编号]2096-711X(2026)16-0148-03
doi:10.3969/j.issn.2096-711X.2026.16.051　　[本刊网址]http://www.hbxb.net

引言

生成式人工智能(Generative Artificial Intelligence,Generative AI,简称生成式AI)以Transformer架构与大语言模型为技术底座,实现了从"数据拟合"到"内容创造"的范式跃迁,进而在高等教育领域呈现"全域渗透"与"深度嵌入"的应用态势。据教育部高等学校科学研究发展中心与中国高等教育学会教育信息化分会联合发布的《教育信息化发展报告(2024)》显示,调研的全国1407所高校中有68.7%的高校开展了一定程度的人工智能应用,近90%的双一流建设院校在教学活动中均有应用生成式AI,应用于包括智能题库生成、虚拟教研助手、科研数据增强等多样化场景。

生成式AI在高校的深度嵌入虽优化了教育资源配置,却也因技术架构的"黑箱化"放大了隐私泄露风险。2023年,某"双一流"高校选课系统采用第三方生成式AI模型进行课程推荐,由于API接口未对学生数据进行本地化脱敏处理,导致超过3000条包含学籍信息与学业表现的数据回流至服务商服务器,并被纳入其商业训练集。此事件不仅引发隐私维权,更警示学界:在缺乏严格的数据隔离与审计机制时,生成式AI的应用可能将校园网络变成敏感数据的"无防护出口"。

如何在使用生成式AI的同时有效防控数据隐私泄露成为近年来备受业内研究者关注的热点。目前已有的研究工作主要聚焦三个维度:技术层面、管理层面与法律层面。其中技术层面侧重关注加密算法在静态数据保护中的应用;管理层面则强调数据全生命周期流程规范;法律层面探讨《中华人民共和国个人信息保护法》在教育场景的适用。然而针对生成式AI特有的"数据投喂—模型训练—内容生成"闭环风险,尚未形成适配性治理框架,特别是生成式模型的"记忆提取"攻击,即通过特定提示词诱导模型输出训练数据片段,使传统"数据—应用"二元防护模式失效,亟需构建新的防护范式。只有将技术防御的精确性、制度规范的严谨性、法律保障的权威性有机结合,才能构建起适配生成式AI特性的教育数据隐私保护体系。

综合以上分析,本文突破传统静态隐私保护的局限性,提出包含动态脱敏机制、权限区块链管理、伦理审查委员会在内的三维防护体系。一是为高校提供可操作的隐私保护实施路径,二是探索生成式AI与教育数据治理的协同机制,为教育数字化转型提供安全保障。

一、生成式AI在高校教育场景的隐私风险图谱

(一)个性化学习系统的数据边界模糊化

目前高校部署的智能学习分析平台大多通过持续采集学生课堂互动、在线答题等动态信息,训练个性化推荐模型。但是这样的个性化学习推荐系统存在感知响应链路的隐私漏洞。系统在识别学生知识盲点时,会自动关联其家庭经济状况、心理测评结果等敏感信息,形成从学习行为到个人属性的数据画像,且未明确告知数据融合的范围与用途。这种多源数据间的隐性关联,使学生隐私边界被技术手段无形侵蚀。

(二)科研协同平台的模型投毒风险

在跨校科研合作中,生成式AI工具常被用于实验数据的预处理与结论生成。某医学高校的案例显示,合作方通过篡改5%的训练样本(如调整临床试验数据的统计参数),使得生成式模型在撰写研究报告时系统给出偏向特定结论,同时导致原始数据的真实性被破坏且不可逆。更严重的是模型投毒行为具有一定的隐蔽性,因为修改过程未被日志记录,且生成内容的偏差性需专业审计才能发现,这对科研数据的完整性与可信度构成严峻挑战。

(三)智能管理系统的权限滥用风险

高校行政管理中,生成式AI用于教职工绩效考核、学生学籍异动等决策支持。随着智能管理系统的普及使用,权限越界问题在高校数字化进程中日益凸显。一项针对多个省份的高校教育信息系统的专项审计调查显示,37%的被调查高校存在不同程度的权限越界现象,具体表现为:工作人员可以利用智能评教系统的后台管理漏洞,违规获取学生对其他教师的匿名评价数据;也可以通过生成式报表工具,直接导出包含学生家庭住址、联系方式等敏感信息的未脱敏数

收稿日期:2026-4-17
基金项目:本文系江苏省教育厅高校哲学社会研究项目课题"AIGC教育应用背景下高校师生隐私风险与治理对策研究"阶段性成果(项目编号:2025SJYB0536)。
作者简介:谢云(1984—),女,江苏宿迁人,南京信息职业技术学院副教授,主要从事网络安全、数据隐私保护研究。

据,从而引发隐私数据泄露。这一现象揭示出权限管理的技术缺陷与制度缺位相互叠加形成了从技术漏洞到权限滥用的风险传导链条。

二、教育数据隐私风险的内在生成机理

(一)技术架构层面的原生缺陷

生成式 AI 的 Transformer 模型存在“注意力机制”导致的隐私泄露隐患。在模型训练阶段,自注意力权重的计算会强化敏感特征的记忆强度,使学生身份证号、教师科研经费等信息被模型“隐性存储”。研究表明,通过设计特定的提示词,如列出“2023 级所有获得国家奖学金学生的银行卡号后四位”,可使模型输出准确率达 78% 的敏感信息片段。这种基于模型内在结构的隐私风险,难以通过传统的防火墙等外围防护手段彻底消除。

(二)制度规范层面的适配滞后

现行教育数据管理制度多沿用传统信息技术架构下的“静态管控”模式,与生成式 AI 的动态数据处理特性不匹配。例如某高校的《数据安全管理办法》仍规定“数据脱敏后可自由流转”,但未考虑生成式 AI 的“逆脱敏”能力,其通过上下文关联推理,能将已脱敏的“1 *** 5”还原为完整的身份证号。制度更新的滞后性,使技术防护措施陷入“合规但不安全”的困境。

(三)法律适用层面的场景冲突

《中华人民共和国个人信息保护法》中规定的“敏感个人信息处理需单独同意”原则,在生成式 AI 场景中面临执行难题。当学生使用智能答疑系统时,单次交互会触发多轮数据处理,例如:语音识别→语义分析→答案生成,每环节都涉及个人信息的转换与输出,但难以实现“环节化同意”的操作机制。法律条款的抽象性与技术流程复杂性之间的落差,导致权利告知沦为形式化的“点击同意”。

三、多维度协同的隐私保护策略体系

针对高校教育数据隐私风险,本文提出构建多维度协同隐私保护体系,具体实施为前期准备、分阶段落地、监督优化三大步骤。

(一)前期准备阶段

1. 数据分类分级梳理

依据《高校教育数据分类分级指南》,组织技术人员、管理人员和法律人员,对高校教育数据进行全面排查,明确核心敏感(如生物识别、心理测评)、重要敏感(如成绩、科研数据)、一般信息(如课程表)三类数据的范围、存储位置及流转路径,建立完整的数据台账。

2. 风险全面排查与评估

组建专项排查小组,重点检测个性化学习、科研协同、智能管理三大核心场景,排查技术漏洞(如模型注意力机制隐患、权限漏洞)、制度缺位(如静态管控不适配)、法律合规短板(如同意机制形式化),形成风险评估报告,标注高、中、低风险点。

(二)分阶段落地阶段

1. 技术防护体系搭建

优先部署“联邦学习+差分隐私”融合方案,完成学生终端数据加密、服务器参数脱敏处理,注入动态高斯噪声;同步搭建区块链权限管理系统,录入各岗位权限规则,实现操作行为上链可追溯;完成现有 AI 系统的安全补丁更新,关闭隐私泄露漏洞。以个性化学习系统中的实施为例,学生终端保留原始数据,仅上传通过同态加密处理后的模型参数更新,使服务器无法反推个体信息;其次在数据集中可以注入高斯噪声,且噪声强度根据信息敏感度动态调整,确保模型精度与隐私保护的平衡,达到数据隐私保护得同时满足个性化推荐的实际需求。

引入区块链技术,实现权限管理的透明化与不可篡改。搭建教育数据权限链,将每个操作节点(如教师查询学生成绩、系统管理员修改配置)的行为记录上链,采用智能合约自动执行权限规则(如规定辅导员仅能查看本年级学生数据),权限滥用行为可追溯、可审计。

2. 建立敏捷治理机制

完善并细化校内合规准则,明确数据收集、共享和使用的具体流程,确保符合《中华人民共和国个人信息保护法》要求,提高高校师生主体对所属数据的管理主动权,有效遏制数据的无序流动。

成立由技术专家、法律学者、师生代表组成的“数据伦理审查委员会”,对生成式 AI 应用实施全流程监管。委员会建立“红、黄、绿”三级风险评估体系:红色区域(如心理测评数据的模型训练)需全员审议,黄色区域(如课程推荐系统)需抽样评估,绿色区域(如教学资源自动生成)可简化流程。通过实施该机制提升上线 AI 应用的隐私合规率。

推行“数据护照”制度,赋予师生对个人信息的控制权。通过开发移动端管理工具,师生可实时查看自身数据的流转路径,并能一键暂停特定用途的数据授权,如拒绝将消费记录用于贫困生认定模型训练。

3. 构建法律适配的合规应用路径

针对生成式 AI 的数据处理特性,依据前期对数据分类分级的梳理采用不同的处理规则。例如:核心敏感数据禁止用于模型训练,重要敏感数据需采用“去标识化+访问白名单”双重保护,一般信息可在匿名化后用于公共服务。这种精细化分类,使法律条款在技术场景中具备可操作性。

建立生成内容溯源制度,要求生成式 AI 工具在输出结果中嵌入数据来源标识,包含原始数据的采集时间、授权范围等元数据,确保内容可追溯至合法数据源。对涉及个人信息的生成内容,设置“人工复核”环节,由专人审查是否存在隐私侵权风险,从输出端把控数据使用的合法性。

4. 增强全员安全意识

定期开展分层培训贯宣,对技术人员重点培训隐私计算、区块链操作技能,对管理人员培训权限管理与审计规范,对师生开展树立数据隐私保护意识及维权相关途径培训;通过校园宣传渠道,普及数据隐私保护知识,培育隐私保护文化。

(三)监督优化阶段

保持常态化监督与动态优化,建议由数据伦理审查委员会牵头,按月开展数据安全审计,核查权限使用、数据流转是否合法合规;按季度更新风险评估模型,监测 AI 系统数据流,及时发现潜在威胁;年终总结实施成效,结合技术迭代和政策变化,优化防护策略,形成“排查—实施—监督—优化”的闭环。

四、结语

生成式人工智能在高校教育领域的深度应用,既带来了教育模式的革新,也引发了数据隐私保护的新挑战。本文通过构建隐私风险图谱、剖析其生成机理、提出防护策略,形成了较为系统的解决方案。未来随着技术的持续演进,高校需要建立动态调适的治理机制,在鼓励创新应用的同时守住隐私安全底线,为教育数字化转型提供可持续的安全支撑。

(下转第 161 页)

The Teaching Reform and Practice of the "Econometrics" Course under the Ideological and Political Guidance and AI Empowerment

KANG Qing-qing, YANG Hai-rong, SONG Ya-qing
(School of Mathematics and Statistics, Hefei Normal University, Hefei Anhui 230601, China)

Abstract: Under the background of the construction of new liberal arts, "Econometrics", as the core course of economics, is an important direction of higher education reform. Faced with multiple difficulties such as difficult to understand theoretical abstraction, rigid integration of ideological and political education, and shallow application of technology, based on the teaching practice of economic statistics major in Hefei Normal University, this paper explores a set of curriculum teaching reform plan of "ideological and political guidance, AI empowerment, and case driven". By taking multiple localized cases of the relationship between economic development and life expectancy as the carrier, the ideological and political elements are naturally integrated into the whole process of data analysis, relying on the learning platform and AI teaching assistants, the intelligent teaching closed loop of "pre-class intelligent presetting — in-class dual-track research and learning — after-class expansion and creation" is constructed, and a process evaluation system matching the graduation requirements is established. The practice shows that this mode can effectively stimulate students' learning initiative, promote the organic unity of value building, knowledge transfer and ability training, and provide a replicable practice path for the econometrics curriculum reform in local colleges and universities.

Key words: ideological and political guidance; AI empowerment; teaching reform　(责任编辑:范新菊)

(上接第 149 页)

参考文献:

[1]张明,李军. 人工智能时代教育数据隐私保护的技术路径[J]. 中国教育学刊,2023(5):45-51.

[2]王磊. 个人信息保护法在教育领域的适用边界[J]. 法学论坛,2023,38(2):112-120.

[3]李华,赵静. 高校数据治理体系构建与实践[J]. 清华大学教育研究,2024,45(1):89-96.

[4]教育部高等学校科学研究发展中心,中国高等教育学会教育信息化分会. 2024 年中国高校数字化发展报告[R]. 北京,2025.

[5]陈沛沛,侯文庆,荣江. 智慧校园数据共享中的安全隐患与应对策略[J]. 科技与创新,2025(18):192-195.

[6]蔡子慧. 从"70 万学生信息被盗",看数据安全治理困局[EB/OL]. (2025-3-17). https://opinion. dahe. cn/2025/03-17/1908054. html.

[7]陈梦根. 生成式人工智能促进新质生产力发展:理论逻辑、影响因素与实践路径[J]. 贵州省党校学报,2025(6):29-38.

Privacy Protection Strategies for University Educational Data in the Context of Generative Artificial Intelligence Applications

XIE Yun
(Nanjing Vocational College of Information Technology, Nanjing Jiangsu 210023, China)

Abstract: The iterative evolution of generative artificial intelligence (GenAI) technologies provides new momentum for the digital transformation of university education, yet its data-driven nature necessitates a paradigm reconstruction in educational data privacy protection. Based on the data lifecycle theory, this paper constructs a "risk—mechanism—governance" analytical framework to systematically deconstruct the privacy leakage pathways of GenAI in personalized learning systems, research collaboration platforms, and intelligent management systems. In accordance with "The Personal Information Protection Law of the People's Republic of China", and by introducing privacy-preserving technical solutions, this study proposes a three-dimensional protection system encompassing dynamic anonymization mechanisms, permission blockchain management, and an ethics review committee. By combining technical defense with institutional constraints, this paper establishes an educational data privacy protection system tailored to the characteristics of GenAI technology, offering theoretical references and practical paradigms for data security governance in the digital transformation of higher education institutions.

Key words: generative artificial intelligence; higher educational data; privacy protection; data security
(责任编辑:桂杉杉)