

生成式人工智能融入高等教育生态系统的风险表征、预警及化解*



黄蓓蓓 宋子昀 钱小龙^[通讯作者]

(南通大学 未来教育研究所, 江苏南通 226019)

摘要: 在生成式人工智能迅猛发展的背景下, 认识、防范和化解生成式人工智能给高等教育生态系统带来的教育风险迫在眉睫。对此, 文章首先剖析了生成式人工智能融入高等教育生态系统的风险表征, 主要表现为师生主权消解、知识内容扭曲、安全包容破坏和人才培养异化。随后, 文章构建了由风险预警主体、风险预警内容、风险预警保障、风险预警流程四个模块组成的生成式人工智能融入高等教育生态系统的风险预警机制。最后, 文章提出从激发主体意识、鉴别知识内容、优化数字环境、培养创新人才等方面入手, 来化解生成式人工智能融入高等教育生态系统的风险。文章的研究可为生成式人工智能教育风险治理提供理论指导, 并推动生成式人工智能在高等教育中的安全应用, 从而为高等教育数字化转型和高质量发展保驾护航。

关键词: 生成式人工智能; ChatGPT; 高等教育生态系统; 教育风险; 风险预警

【中图分类号】G40-057 【文献标识码】A 【论文编号】1009—8097 (2024) 05—0016—11 【DOI】10.3969/j.issn.1009-8097.2024.05.002

生成式人工智能 (Generative Artificial Intelligence, GAI) 的蓬勃发展, 促进了互联网时代知识生产模式的变革, 也为教育创新与变革增添了新动力。生成式人工智能是一种人工智能技术, 其利用深度学习模型生成类似人类的内容, 以响应复杂多样的提示^[1]。ChatGPT、Google Bard、文心一言、通义千问等, 都属于典型的生成式人工智能工具。以 ChatGPT 为代表的生成式人工智能粗暴、直接地融入教育生态系统中, 引发了系统中不同要素间的“连锁性反应”^[2], 而高等教育生态系统更是首当其冲, 技术融入所导致的教育风险已开始涌现。例如, 《大西洋月刊》刊发的一篇文章直接用“大学论文已不复存在”作为标题^[3], 来表达对大学生课程论文作弊的担忧。此外, ChatGPT 会生成伪造的参考文献^[4]、捏造的历史事实等^[5], 容易误导学生。随着生成式人工智能在高等教育领域应用的深入, 认识、防范和化解其给高等教育生态系统带来的教育风险迫在眉睫。对此, 本研究从风险表征、风险预警、风险化解三个方面展开研究, 重点回答生成式人工智能给高等教育生态系统带来的教育风险有哪些表现、如何提前防范风险以及如何有效化解风险这三个问题, 以期实现新一代人工智能技术在高等教育生态系统中的安全应用。

一 生成式人工智能融入高等教育生态系统的风险表征

生态系统是生物之间以及生物与非生物环境之间通过物质循环、能量流动、信息传递形成的统一体。而高等教育生态系统是高等教育系统内部各要素之间及其与自然、社会环境之间交互所形成的开放性、非线性复杂系统^[6]。以生态系统的结构和功能为分析维度, 生成式人工智能融入高等教育生态系统的风险主要表现为生态主体层面的师生主权消解风险、生态介体层面的知识内容扭曲风险、生态环境层面的安全包容破坏风险、生态功能层面的人才培养异化风险。

1 师生主权消解, 弱化高等教育生态主体

高等教育生态系统的核心主体是学生和教师, 其中学生是消费者, 教师是生产者。师生主

权消解风险是指生成式人工智能在教育领域的过度应用,可能会使学生和教师过度依赖技术,从而消解其教育主体地位,让学生和教师逐渐成为被技术支配的延伸客体,具体表现为:①学生学习主体性消解风险。与基础教育相比,高等教育更强调学生的自主学习。生成式人工智能采用便捷的对话问答方式,能简化信息的获取过程,技术的频繁使用容易让学生产生一定的依赖思想和惰性心理,继而沉迷其中。这种沉迷会影响学生的学习自主性、批判性思维和问题解决能力,使其容易屈服于技术,逐渐丧失学习主权。此外,学生的身体还可能被技术控制并导致身体隐退^[7]。②教师教学主体性消解风险。目前,高校的学术科研压力可能会在一定程度上逼迫大学教师部分甚至完全借助生成式人工智能来撰写课程方案、生成教学材料等。当过度依赖生成式人工智能时,大学教师便容易失去教学实践的主动性,成为智能技术的执行者。此外,凭借快捷的智能响应和广泛的信息储备,生成式人工智能可以实现评价、反馈等部分教育教学职能,这会在一定程度上冲击教师的知识权威地位,使其逐渐丧失教学实践主权。

2 知识内容扭曲,影响高等教育生态介体

高等教育生态介体是指连接不同生态主体以及生态主体与环境之间的中介,其中以知识为代表的教育内容是一种重要的介体。传播专业化的知识是高等教育的重要使命,这就要求知识内容必须科学、权威。知识内容扭曲风险是指生成式人工智能产生的知识内容可能存在虚假信息,扭曲知识内容的科学性和权威性,具体表现为:①知识内容科学性存疑风险。训练数据的失真,会让生成式人工智能成为虚假信息的制造者。由于生成式人工智能生成的文本与人类生成的文本相似,故大学生很难区分真实的知识和未经验证的信息。在高等教育中,事实性错误、概念误用、高深知识出错等情况会误导大学生^[8],进而对其价值观念造成负面影响。②知识内容权威性受损风险。生成式人工智能生产的知识是一种“组合式创新”,是对已有数据和信息的重组,并非实质性的知识生成。这种知识内容并不遵循“专家或教育行政部门开发、审定→学校传授”的知识传播逻辑^[9],因而人机交互直接获得的知识貌似“高大上”,但权威性却难以保证。若认识不到这种幻象,就容易引发将生成式人工智能视为知识权威和道德权威的风险^[10]。

3 安全包容破坏,侵害高等教育生态环境

高等教育生态环境既包括现实环境,也包括虚拟环境。伴随高等教育数字化转型的加快,数据日益成为数字化教学的核心资源。在此背景下,数字化的虚拟环境更加强调保障数据安全与实现数字包容。安全包容破坏风险是指生成式人工智能破坏虚拟环境的安全性和包容性,具体表现为:①教育数据安全和隐私风险,主要包括隐私泄露和隐私侵犯^[11]。一方面,开发商可能会从教师、学生等用户与生成式人工智能的交互中收集数据并进行分析,以提高生成式人工智能的性能,至于交互留下的历史数据将会被如何处理尚不明晰;另一方面,生成式人工智能可能会根据大学师生提供的输入数据进行训练,输入的部分隐私信息可能会以某种形式提供给后续用户,以上均有可能造成师生个人隐私数据的泄露和隐私侵犯。②偏见歧视风险。生成式人工智能使用来自互联网的大量数据进行预训练,如果训练数据存在偏见,那么生成的结果也会存在偏见。这些偏见包括性别偏见、种族偏见、文化和语言偏见、意识形态偏见等^[12],会强化先前存在的不平等和刻板印象,破坏高等教育生态环境的包容性。

4 人才培养异化,扰乱高等教育生态功能

除了具备生态系统的一般功能,高等教育生态系统还承载了人才培养、学术科研、社会服务三大重要功能,而生成式人工智能的融入给人才培养带来了风险。人才培养异化风险是指生

成式人工智能在高等教育生态系统中的融入会让原本具有创造性、道德性、协同性的人才培养活动异化,具体表现为:①人才培养异化为同质的标准生产。通过培养创新型人才来助推教育强国建设,是高等教育的重要使命。生成式人工智能使用深度学习算法,根据已有的训练数据生成内容,这限制了生成内容的创造性。当完全依赖生成式人工智能进行教与学时,高等教育教学便会异化为标准化的工厂生产。②人才培养异化为泛滥的学业失信。培养有道德的数字公民,是高等教育人才培养的重要目标之一。若不进行技术检测,将很难区分生成式人工智能产生的文本和人类产生的文本^[13]。当大学生利用此类工具完成论文或作业、通过作弊或抄袭获得更高的分数时,便会导致学业失信。③人才培养异化为恶性的人机博弈。伴随机器的加入,实现人机协同的教育教学成为高等教育的重要追求。然而,当大学生利用生成式人工智能进行学习时,若只关注机器反馈的结果而忽视知识本身及其价值,就有可能出现人机博弈^[14]。

二 生成式人工智能融入高等教育生态系统的风险预警

生成式人工智能融入高等教育生态系统存在诸多风险,需要进行风险预警,以保障高等教育生态系统的安全运行与可持续发展。风险预警旨在预测和警示生成式人工智能给高等教育生态系统带来的教育风险,重点回答“谁来预警”“预警什么”“怎样保障”“如何预警”四个问题。对此,本研究构建了由风险预警主体、风险预警内容、风险预警保障、风险预警流程四个模块组成的生成式人工智能融入高等教育生态系统的风险预警机制,如图1所示。具体来说,该机制是以高校为主和以政府、教育行政部门、家庭、社会为辅的风险预警主体,针对生态主体、生态介体、生态环境、生态功能四个层面的风险,通过组织建设、制度构建、技术运用、人才培养等风险预警保障,按照“风险识别→风险分析→风险评价→风险警示→风险追踪→风险识别……”的风险预警流程,对生成式人工智能带来的教育风险进行预警。

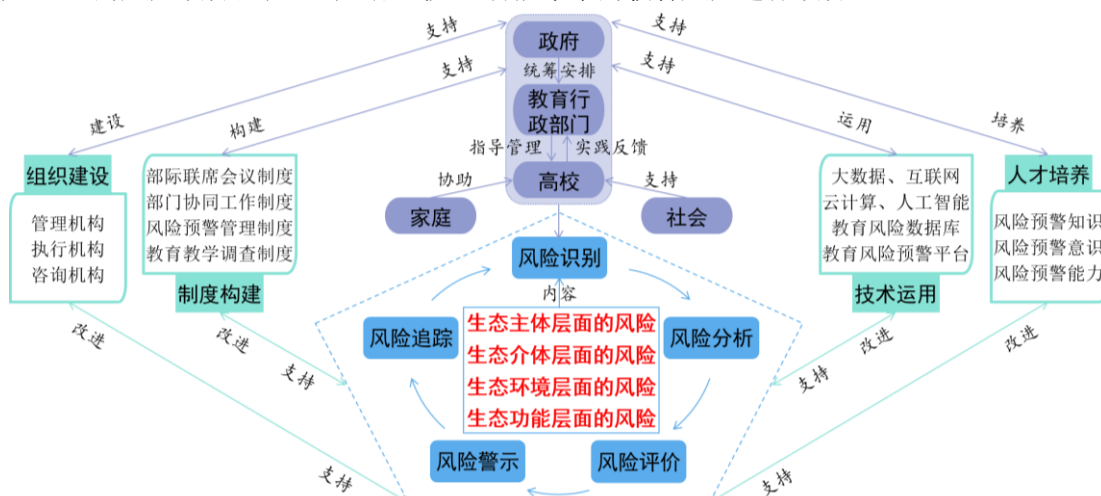


图1 生成式人工智能融入高等教育生态系统的风险预警机制

1 风险预警主体：以高校为关键行动者的多元主体协同参与

结合风险预警的功能实现,风险预警主体包括政府、教育行政部门、高校、家庭、社会。其中,高校直接参与风险预警行动,是风险预警的关键行动者;其他主体通过与高校协作互动,

间接参与风险预警行动：①政府是“总体牵头者”，主要负责对风险预警工作进行统筹安排，包括构建生成式人工智能教育风险预警制度、推动制定生成式人工智能风险预警防控标准文件、为高校教育风险的预警和防控提供政策支持、调动各职能部门参与风险预警并明确各部门的职责分工等。②教育行政部门是“直接管理者”，主要负责制定风险预警的具体政策、规章、规划等，确定风险预警的基本流程，并对各高校的预警工作进行指导、管理等。③高校是“关键行动者”，主要根据政府和教育行政部门的要求开展具体的风险预警工作，对教育风险进行识别、分析、评估、警示、追踪；同时，根据通用的教育风险框架，结合本校的特点和实际，将教育教学实践中具体的风险预警情况及时反馈给上级教育主管部门。④家庭是“重要协助者”，其与高校联系密切，主要配合高校开展相关的教育风险调查或主动向高校反馈观察到的教育风险因素，辅助高校开展风险预警工作。⑤社会是“有力支持者”，其中的行业协会、科技企业、各大媒体等都可可为高校开展相关的风险预警工作提供支持，如行业协会能为生成式人工智能标准规范的制定提供智力支持、科技企业能为教育风险预警平台的搭建提供技术支持、各大媒体能为教育风险预警信息的广泛发布提供宣传支持等。

2 风险预警内容：技术引发生态系统结构与功能的嬗变

生成式人工智能融入高等教育生态系统，引发了师生主权消解、知识内容扭曲、安全包容破坏、人才培养异化等风险，也引发了高等教育生态系统结构（包括主体、介体、环境）与功能的嬗变。因此，风险预警内容不仅要聚焦微观层面的四大风险表征，更要从宏观入手，关注以下四个层面的风险：①生态主体层面的风险，是指生成式人工智能给教育部门、高校及其管理者、教师、学生等主体带来的风险，包括教育权益、职能、角色方面的侵害、弱化与改变等。②生态介体层面的风险，是指生成式人工智能对高等教育的内容、方法、媒介等造成损害，影响生态主体之间以及生态主体与生态环境之间的互动。高校是教育、科技、人才的重要汇聚点^[15]，而知识是支撑高校担当使命的重要基础。因此，高等教育生态系统中的知识内容扭曲风险需要得到重点关注。③生态环境层面的风险，是指生成式人工智能破坏高等教育发展、产生制约和控制作用的多维空间与环境^[16]。生成式人工智能技术的发展不仅会导致安全包容破坏风险，还有可能使教育环境被资本利益操控^[17]，让商业化凌驾于公益性和人本价值之上。④生态功能层面的风险，是指生成式人工智能扰乱高等教育生态系统的相关教育功能。除了异化人才培养功能，生成式人工智能还会给高等教育生态系统的学术科研功能带来危害，包括分散学术责任、催生学术歧视等^[18]。整体而言，这些风险均与高等教育密切相关，具有风险来源复杂多样、风险类型动态变化、风险影响持续深远等特点。

3 风险预警保障：整体智治助力生态系统安全稳定运转

针对生成式人工智能的教育风险预警是一项系统工程，需要以智能的方式整体推进。“整体智治”主张运用数字技术来推动整体性治理的升级，重点强调整合组织部门、重塑制度体系、应用数字技术和提升主体能力^[19]。基于此，可从组织建设、制度构建、技术运用和人才培养四方面提供风险预警保障：①组织建设，即政府、教育行政部门、高校都要建立专门负责或兼管教育风险预警的管理、执行、咨询等机构并组织实施。对此，高校可以借鉴麻省理工学院的做法^[20]，包括成立风险指导小组，为风险预警提供战略指导和方向；成立风险咨询团队，调集多个领域的专家、学者来识别和评估风险；成立风险管理委员会，监督和管理风险预警工作，宣传风险管理文化。此外，高校还可以设立风险预警研究中心，为预防数字技术带来的教育风险

提供智力支持。②制度构建,即构建“国家(中央政府和教育部)-地方(地方各级政府和教育行政部门)-高校”风险预警体系,重点建设部际联席会议制度(国家各部委对教育风险预警决策进行协商)、部门协同工作制度(多元主体对风险预警工作进行统一执行与有效落实)、风险预警管理制度(明确各主体的风险预警责任,规范风险预警管理流程,建立风险预警信息报告机制)、教育教学调查制度(定期调查生成式人工智能在高校教学中的应用情况和潜在风险)。

③技术运用,即高校充分发挥科研的优势,与相关企业合作,应用大数据、互联网、云计算、人工智能等技术,建立教育风险数据库,打造教育风险预警平台。例如,应用机器学习、智能统计引擎、图像识别等技术,对教育数据方面的风险进行捕捉、识别、定性、分析、评估^[21]。

④人才培养,即高校和教育行政部门开展生成式人工智能风险预警的相关教育和培训活动,通过开发风险预警课程、采用场景式教学模式、搭建风险预警实训平台、组织风险预警专题讲座、实施风险能力评估、深化校企协同育人机制等,激发学生、教师、教育管理者、技术人员等的风险预警意识,夯实其风险预警知识,不断提高其风险预警能力。

4 风险预警流程: 人机协同推动实现敏捷化的信息传递

风险预警属于风险管理的一部分。国际标准化组织(International Organization for Standardization, ISO)在《2018 风险管理指南》中指出,风险管理的过程包括沟通与咨询、建立环境、风险评估(包括风险识别,风险分析和风险评价)、风险应对、追踪和检查、记录和报告等活动^[22]。此外,风险警示重在让预警信息得到有效传递,其也是风险预警的关键一环。基于此,本研究设计了“风险识别→风险分析→风险评价→风险警示→风险追踪→风险识别……”的风险预警流程,如图2所示。这是一个循环往复的过程,多元主体相互联动,以风险保障为支持来实施风险预警流程。其中,最核心的是运用人机协同的风险预警方法与手段,将人对风险的认识、研判和机器在风险数据收集、分析、评价、监测等方面的作用进行融合,实现对风险预兆的敏捷响应,促进风险预警信息在高等教育生态系统中的敏捷传递。

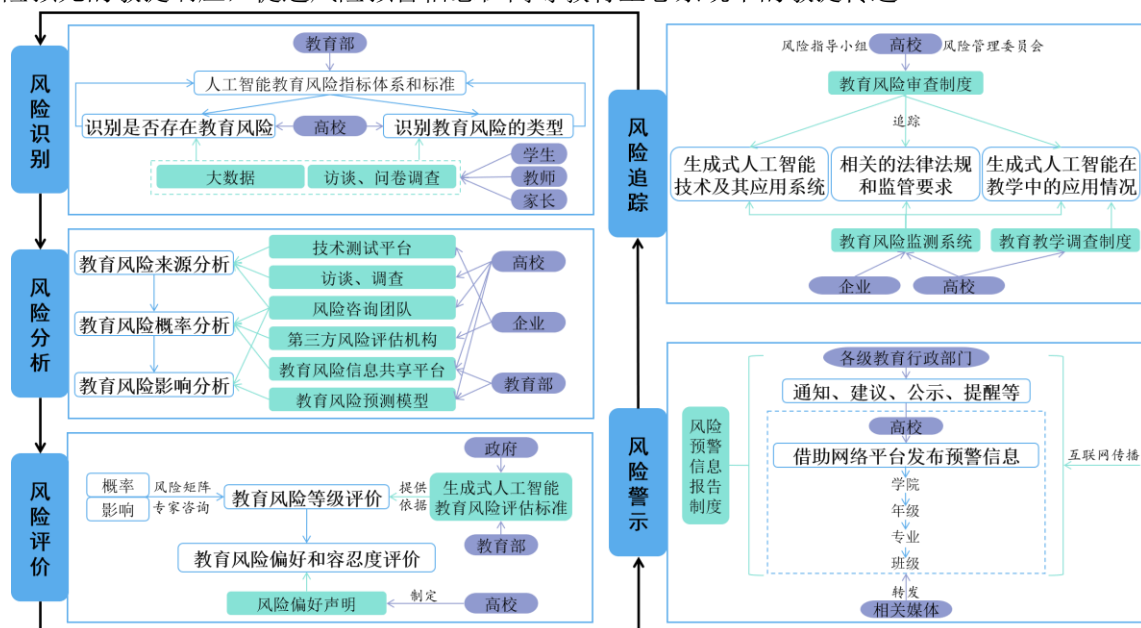


图2 生成式人工智能融入高等教育生态系统的风险预警流程

（1）风险识别

进行教育风险预警的第一步是对教育风险进行识别，即识别生成式人工智能带来的教育问题中有哪些存在教育风险，并分析判断该教育风险的类型。对此，教育部首先要发布通用的人工智能教育风险指标体系（下文简称“通用指标体系”），为高校进行教育风险识别提供抓手和依据。在此基础上，高校要结合自身特点，依托通用指标体系，应用大数据技术充分收集并分析相关信息来对教育风险进行实时甄别与筛选^[23]；再结合访谈、问卷调查等传统方法，了解潜在的教育风险，最终以人机协同的方式进行教育风险识别。风险识别主要包括两方面的工作：①识别是否存在教育风险。高校可以通过审查相关文件、分析已有数据，来识别存在的风险；也可以通过访谈、问卷调查等形式，定期从学生、教师、家长等相关人员中获得反馈信息，了解其在应用生成式人工智能过程中遇到的问题与挑战，并识别出其中的教育风险；还可以召开专家研讨会，并结合技术预测，对可能出现的风险进行提前预判与确定。②识别教育风险的类型。高校以教育部发布的通用指标体系为基础，对识别到的教育风险进行分类。然而，通用指标体系并不能覆盖所有高校的教育风险类型。因此，不同层次水平、办学类型、保密级别的高校需要进一步构建适合本校的教育风险指标体系，对个性化的教育风险进行定位与识别。之后，高校还需将相应的教育风险记录在风险登记手册中，并及时上传至教育风险数据库。

（2）风险分析

识别教育风险之后，接下的一步便是分析教育风险，具体包括：①教育风险来源分析。可以从生成式人工智能技术本身、教育主体的数字素养、技术在教育教学中的应用情况等方面，来分析教育风险的成因。在此过程中，高校可以与企业合作，借助技术测试平台（如新华网联合中国科学院计算技术研究所等行业机构研发打造的“生成式人工智能内容安全与模型安全检测平台”^[24]），对生成式人工智能进行测试、对教育主体开展访谈和调查、依托风险咨询团队对风险来源进行科学研判等。②教育风险概率分析。一方面，高校可以借助校内的风险咨询团队和第三方风险评估机构，咨询不同领域专家和技术人员的意见，从而预估教育风险发生的可能性。另一方面，不同高校、教育部可以共同搭建教育风险信息共享平台，连接教育风险数据库，汇集相关数据，在大数据、人工智能等技术的支持下分析风险概率。③教育风险影响分析。可以从生态主体、生态介体、生态环境、生态功能等要素出发，分析生成式人工智能给高等教育生态系统带来的潜在危害。在此基础上，高校可以依托风险咨询团队，利用机器学习、深度学习等技术构建教育风险预测模型，来分析教育风险的影响，并关注教育风险影响的规模、范围、时效性等。从长期来看，还可以重点关注生成式人工智能给教育伦理带来的影响。

（3）风险评价

教育风险评价旨在根据教育风险分析的结果来评价生成式人工智能的教育风险等级，并评价其教育风险偏好和容忍度，具体包括：①教育风险等级评价。教育风险等级反映的是教育风险的严重程度，大致可以分为低风险、一般风险、较大风险、重大风险四个等级。可以根据教育风险发生的概率及其影响，在专家咨询的基础上应用风险矩阵法，计算得到风险的等级数值，之后与相应的风险标准进行比较，最终确定教育风险的等级。目前，国内尚未有针对生成式人工智能和教育风险的统一评估标准。为科学评估教育风险，政府和教育部亟需制定生成式人工智能教育风险评估标准，明确风险等级划分，提供评价方法指导，为地方各级教育行政部门和高校开展教育风险评估提供依据和支持。②教育风险偏好和容忍度评价。风险偏好（Risk

Appetite)是高校愿意承担和接受的风险水平;而风险容忍度(Risk Tolerance)是高校在管理风险时愿意接受的结果的变化程度,是对风险偏好的进一步量化。高校需要设置不同的风险偏好声明,明确本校为实现风险预警目标而愿意接受的风险等级,并界定相应的风险承受能力,以指导高校的风险管理人员进行决策,使决策与风险承受能力相一致。在生成式人工智能驱动的教学创新方面,高校需要声明较高的风险偏好;而在技术引发的数据隐私安全方面,高校需要声明较低的风险偏好,甚至声明无风险偏好。

(4) 风险警示

教育风险警示是教育风险预警的关键一环,主要通过层级化下达和网络化传播相结合的方式实现,警示的内容包括教育风险的发生概率、影响情况和严重等级。具体而言,各级教育行政部门、高校和相关媒体通过由上而下层级化的风险预警信息报告制度,借助互联网及时传播教育风险警示信息,以确保每个利益相关者知晓警示信息:①各级教育行政部门主要以通知、建议、公示、提醒等方式进行教育风险警示。例如,教育部发布生成式人工智能教育风险提示清单,对一定时间内可能出现的普遍性教育风险进行警示,并提醒各高校做好风险防范工作。②高校借助网络平台发布预警信息,确保各学院、年级、专业、班级的相关师生及时获知风险预警信息,从而促使其在使用生成式人工智能进行教与学的过程中能够防范教育风险。除了在校内发布生成式人工智能带来的普遍性教育风险,高校还要结合本校的具体情况,额外提醒本校所面临的特殊风险。③相关媒体应积极转发政府、教育行政部门等发布的教育风险警示信息,进一步扩大教育风险警示的范围。各大媒体需重视信息来源的真实性和可靠性,避免传播虚假信息。在此过程中,权威的官方教育媒体发挥着不可或缺的作用。

(5) 风险追踪

随着高等教育生态系统内外部环境的变化和风险应对措施的实施,教育风险可能会出现、改变或消失,因而需要对教育风险进行动态追踪,以增加、调整或取消相应的教育风险警示。在此之前,高校的风险指导小组和风险管理委员会需先建立教育风险审查制度,实时记录风险,定期召开风险审查会议,为风险追踪提供制度保障。风险追踪的内容主要包括:①生成式人工智能技术及其应用系统。继发布 ChatGPT-3.5 之后,OpenAI 公司又迅速发布了 ChatGPT-4。生成式人工智能技术的迅猛发展,可能给高等教育生态系统带来新的风险。因此,高校需掌握生成式人工智能技术的发展情况,及时对技术性能进行测试。②相关的法律法规和监管要求。2023 年 7 月,国家互联网信息办公室等七部门联合发布了《生成式人工智能服务管理暂行办法》^[25]。高校需结合此暂行办法的规范要求来定位教育风险,为重新确定风险的性质和等级提供依据。③生成式人工智能在教学中的应用情况。高校需建立教育教学调查制度,监测教育风险的演变和发展情况,并监测是否有新的教育风险发生。在此过程中,高校可以联合科技企业开发教育风险监测系统,借助大数据、人工智能等技术对教育风险开展大规模的持续监测和风险重估。

三 生成式人工智能融入高等教育生态系统的风险化解

面对师生主权消解、知识内容扭曲、安全包容破坏、人才培养异化等风险,不能仅限于通过风险预警机制来防范,还需要高校这一关键主体联合政府、教育行政部门等主体,从激发主体意识、鉴别知识内容、优化数字环境、创新人才培养等方面入手来有效管控和化解这些风险。

1 激发主体意识：提高师生智能素养，制定技术应用指南

防止生成式人工智能剥夺师生的教育主体身份、逐步消解师生的教育主体性地位，需要切实激发师生的主体意识，这一方面可从智能素养入手，在认知层面让大学师生意识到生成式人工智能服务于高等教育的工具属性；另一方面可从应用指南入手，在应用层面为大学师生合理使用生成式人工智能提供指导：①提高师生智能素养。按照“初步接纳→增强认同→深度思考”的内在逻辑^[26]，以开展人工智能素养通识教育和培训为抓手，帮助大学师生树立正确的技术观，掌握生成式人工智能的原理，发展生成式人工智能难以替代的批判性思维、创造力、伦理道德意识等，打破技术权威并建立教育主体自信。同时，将智能素养测评纳入大学教师专业发展和大学生综合素养评价的范畴。②高校要制定生成式人工智能技术应用指南，说明如何在教学中正确使用这些工具和用其作弊的后果等^[27]。史翠珊效应（Streisand Effect）显示，全面禁止反而会增加 ChatGPT 的使用^[28]，而技术与教育的融合无法阻挡。因此，高校要制定应用指南，促进生成式人工智能有效融入大学课堂并防止师生滥用技术。例如，美国罗切斯特大学制定了《艺术、科学和工程教师课堂使用 ChatGPT/AI 指南》^[29]，用于教师选择是否、何时、如何在课堂上使用生成式人工智能；同时，规定了教师需要遵守的务实、透明、始终如一、负责等伦理原则。

2 鉴别知识内容：实施教育提示工程，设计人机协作任务

知识内容扭曲风险源于生成式人工智能本身的缺陷，且与知识的传递方式密切相关。为此，可将生成式人工智能生成的知识内容作为教学材料，让大学生通过使用提示来实现知识的迭代传递，在人机交互的过程中有效鉴别知识内容，具体包括：①实施教育提示工程，聚焦批判性思维培养，帮助大学生学会提问与反思。教育提示工程（Educational Prompt Engineering）聚焦教育领域提示的设计，旨在让学习者和教育者能够使用适合机器理解的自然语言来获得准确的响应^[30]。大学需要实施教育提示工程，具体可从传播全面的内容知识、培养批判性思维和运用迭代设计方法入手^[31]，来开发教育提示工程相关的课程，构建深度对话学习模式，让学生在问答过程中对技术生成的内容进行批判性思考。②设计人机协作任务，运用评估验证策略，实现对知识的科学建构。大学教师可以设计学生与生成式人工智能共同参与的学习任务，让学生批判性地评估这些工具的价值，并有效和负责任地使用这些工具。例如，教师可以要求学生与 ChatGPT 共同撰写一篇文章，并反思共同写作的过程和内容。但需注意的是，任务结束后教师必须及时为学生提供反馈，评估生成式人工智能生成的内容的可信度，以减少虚假信息的传播。

3 优化数字环境：加强数据安全治理，完善算法监管机制

数字化转型是撬动高等教育生态系统实现高质量发展的重要杠杆，而这有赖于构建安全包容的数字环境。安全包容破坏风险大多由技术本身的数据和算法引起，因此加强对数据安全的治理和对算法的监管十分重要，具体包括：①加强对数据安全的整体性治理和敏捷治理，提升高校数据安全治理能力。针对教育数据安全，需要构建政府、教育行政部门、高校三位一体的数据安全治理体系，通过整体性治理来破解教育数据安全治理的碎片化难题，并通过敏捷治理来实现对教育数据安全风险的快速识别与响应。具体到高校，一要明确教育数据安全治理的目标和战略愿景；二要建设教育数据安全的管理、执行机构，并明确其职责分工；三要制定高校教育数据安全政策文件、管理办法等；四要采用数据脱敏技术、访问控制等，以保障教育数据安全；五要提升大学教师、学生、管理人员的数据安全意识与能力。②完善算法监管机制，构建差异化、多主体、全周期的算法监管机制，包括参考欧盟在《人工智能法案》中提出的基于

风险等级的监管思路^[32],针对不同等级的教育风险采取不同的管控措施;调动政府、高校、社会等多元主体参与监管的积极性,并明确各主体的监管责任和权力,为社会组织和用户群体参与人工智能治理创造条件^[33];从输入端与输出端入手,强化对生成式人工智能的全周期监管^[34]。

4 培养创新人才:开发交叉课程,打造大学融合课堂

为化解人才培养异化风险,可以借助生成式人工智能技术创新人才培养的内容和模式,具体包括:①利用生成式人工智能整合课程资源,开发跨学科的交叉课程。在生成式人工智能支持的人机互补式课程资源开发模式^[35],不同学科领域的大学教师和专家可以利用生成式人工智能整合课程资源,并对课程资源进行二次设计,开发出跨学科的交叉课程。在此基础上,大学教师可以将相关的课程资源提供给学生,学生也可以根据自身的学习需求进一步定制课程。同时,大学教师要做好相应的审查评估工作。此外,要按照常规的知识生产逻辑来开发体系化、权威化的专家课程,并注重技术生成课程与专家课程的互相补充^[36],从而为开展规模化的跨学科教学提供支持,并为应用技术生成课程提供参考和依据。②构建人机协同的教学模式,打造大学融合课堂。具体来说,一要分割教学责任,明确生成式人工智能机器与教师各自的角色、任务、责任^[37];二要基于生成式人工智能支持的内容生成和个性化辅导,通过设计人机协作学习任务,来提高学生的创造力和批判性思维,促进学生的高意识生成式学习^[38];三要以现实社会交互中蕴含的情感来引导学生与生成式人工智能机器进行正确互动,构建“学生-教师-机器”之间的良性关系;四要以元宇宙学校为抓手^[39],构建线上线下融合式学习空间,让真实的人际互动与智能的人机互动相融合、让现场的实践操练与虚拟的仿真模拟相融合。

四 结语

生成式人工智能的融入,有望从教育教学、学术科研、社会服务、教育治理等方面重塑高等教育生态^[40],但也给高等教育生态系统带来了潜在的教育风险。本研究从风险表征、风险预警、风险化解三个方面,分析了生成式人工智能融入高等教育生态系统所产生的教育风险,可为生成式人工智能教育风险治理提供理论指导,并为高等教育生态系统的数字化转型和高质量发展夯实安全底座。为促进生成式人工智能深度融入高等教育生态系统,本研究建议从以下方面入手:①以安全为底线,遵循可信人工智能的基本框架,立足高等教育需求,面向高等教育应用场景,以基础能力层、专业能力层和应用服务层为架构^[41],开发安全可信的教育大模型,服务于高等教育;②关注生成式人工智能给高等教育生态系统带来的伦理风险,制定生成式人工智能教育应用的伦理原则,开展科技伦理教育,让大学师生合乎道德地使用生成式人工智能工具;③将生成式人工智能与元宇宙相融合,构建具身智能型的高等教育环境,开发动态适应的沉浸式学习资源,探索构建新型的人机协同学习模式^[42],让大学生在深度对话、具身体验、有效移情的过程中掌握专业知识,培养高阶思维,强化伦理道德意识,进而实现智慧成长。

参考文献

- [1][28]Lim W M, Gunasekara A, Pallant J L, et al. Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators[J]. The International Journal of Management Education, 2023,(2):1-13.
- [2][17]詹泽慧,季瑜,牛世婧,等.ChatGPT 嵌入教育生态的内在机理、表征形态及风险化解[J].现代远距离教育,2023,(4):3-13.

- [3]Stephen M. The college essay is dead[OL].
<<https://www.theatlantic.com/technology/archive/2022/12/chatgpt-ai-writing-college-student-essays/672371/>>
- [4][7]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.
- [5]申丽萍,何朝帆,曹东旭,等.大语言模型在中学历史学科中的应用测评分析[J].现代教育技术,2024,(2):62-71.
- [6]胡涌,柳小玲,王玲,等.高等教育生态系统的基本构架[J].中国林业教育,2009,(1):1-5.
- [8]钟秉林,尚俊杰,王建华,等.ChatGPT对教育的挑战(笔谈)[J].重庆高教研究,2023,(3):3-25.
- [9][36]李会春.ChatGPT的智慧生成特征及对高等教育的挑战[J].江苏高教,2023,(8):1-12.
- [10]段伟文.积极应对 ChatGPT 的伦理挑战[N].中国社会科学报,2023-3-7(7).
- [11][37]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]Ray P P. ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope[J]. Internet of Things and Cyber-Physical Systems, 2023,(3):121-154.
- [13]Cotton D R, Cotton P A, Shipway J R. Chatting and cheating: Ensuring academic integrity in the era of ChatGPT[J]. Innovations in Education and Teaching International, 2024,(2):228-239.
- [14]张志祯,张玲玲,米天伊,等.大型语言模型会催生学校结构性变革吗?——基于 ChatGPT 的前瞻性分析[J].中国远程教育,2023,(4):32-41.
- [15]许韬,李宇航,代东旭.高等教育数字化的理论检视、现实挑战与进阶路径[J].中国高等教育,2023,(23):18-21.
- [16]贺祖斌.中国高等教育系统的生态学分析[D].武汉:华中科技大学,2005:5.
- [18]路飞,马雨璇.人工智能生成内容对学术生态的影响与应对——基于 ChatGPT 的讨论与分析[J].现代教育技术,2023,(6):15-25.
- [19]赵玲玲.整体智治:基层治理模式创新的实践逻辑与实现路径——以浙江省“大综合一体化”行政执法改革为例[J].地方治理研究,2023,(1):12-24、78.
- [20]MIT Risk Management & Compliance Services. Risk management system[OL].
<<https://riskandcompliance.mit.edu/risk-management-mit/risk-management-system>>
- [21]赵磊磊,张黎,代蕊华.智能时代教育数据风险治理:实然困境与实践路径[J].湖南师范大学教育科学学报,2021,(6):94-102.
- [22]ISO. ISO 31000: 2018-risk management guidelines[OL].
<<https://pecb.com/whitepaper/iso-310002018-risk-management-guidelines>>
- [23]付昌奎,曾文婧.强化风险治理 护航教育改革[N].光明日报,2019-12-17(13).
- [24]新华网.新华网举行“生成式人工智能内容安全与模型安全检测平台”(AIGC-Safe)邀请测试发布会[OL].
<<https://baijiahao.baidu.com/s?id=1771219051791186742&wfr=spider&for=pc>>
- [25]国家互联网信息办公室等七部门.生成式人工智能服务管理暂行办法[OL].
<http://www.cac.gov.cn/2023-07/13/c_1690898327029107.htm>
- [26]钱小龙.面向高质量发展的基础教育数字化转型行动[J].南通大学学报(社会科学版),2023,(5):122-132.
- [27]Rudolph J, Tan S, Tan S. ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?[J]. Journal of Applied Learning & Teaching, 2023,(1):342-363.
- [29]University of Rochester. AS&E instructors' guide to using ChatGPT/AI in the classroom[OL].
<<http://www.rochester.edu/college/honesty/assets/pdf/chatgpt-ai-guidance-for-instructors.pdf>>
- [30]赵晓伟,祝智庭,沈书生.教育提示语工程:构建数智时代的认识论新话语[J].中国远程教育,2023,(11):22-31.

- [31]Cain W. Prompting change: Exploring prompt engineering in large language model AI and its potential to transform education[J]. TechTrends, 2024,(1):47-57.
- [32]EU Artificial Intelligence Act. EU artificial intelligence act[OL]. <<https://www.aiact-info.eu/articles/>>
- [33]张欣.生成式人工智能的算法治理挑战与治理型监管[J].现代法学,2023,(3):108-123.
- [34]吕健,陆宣.ChatGPT 为劳动者带来的机遇、挑战及其应对[J].当代经济管理,2023,(12):1-8.
- [35]苗逢春.生成式人工智能技术原理及其教育适用性考证[J].现代教育技术,2023,(11):5-18.
- [38]祝智庭,戴岭,胡姣.高意识生成式学习:AIGC 技术赋能的学习范式创新[J].电化教育研究,2023,(6):5-14.
- [39]钱小龙,张奕潇,宋子昀,等.打开元宇宙学校之门:发端、现状与走向[J].现代教育技术,2023,(3):15-26.
- [40]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.
- [41]曹培杰,谢阳斌,武卉紫,等.教育大模型的发展现状、创新架构及应用展望[J].现代教育技术,2024,(2):5-12.
- [42]翟雪松,楚肖燕,焦丽珍,等.基于“生成式人工智能+元宇宙”的人机协同学习模式研究[J].开放教育研究,2023,(5):26-36.

Risk Representation, Early Warning and Resolution of Integrating GAI into the Higher Education Ecosystem

HUANG Bei-Bei SONG Zi-Yun QIAN Xiao-Long^[Corresponding Author]

(Institute of Future Education, Nantong University, Nantong, Jiangsu, China 226019)

Abstract: Under the backdrop of the rapid development of generative artificial intelligence (GAI), it is urgent to recognize, prevent, and resolve the educational risks that GAI brings to the higher education ecosystem. Therefore, this paper firstly analyzed the risk representation of integrating GAI into the higher education ecosystem, which mainly manifested as the dissolution of teachers' and students' sovereignty, the distortion of knowledge content, the destruction of security inclusion, and the alienation of talent cultivation. Subsequently, the risk warning mechanism that integrating GAI into the higher education ecosystem was constructed, which consisted of four modules of risk warning subject, risk warning content, risk warning guarantee, and risk warning process. Finally, this paper proposed to address the risks of integrating GAI into the higher education ecosystem by stimulating subject awareness, identifying knowledge content, optimizing digital environment and cultivating innovative talents. The research in this paper could provide theoretical guidance for risk governance of GAI education, promote the safe application of GAI in higher education, and safeguard the digital transformation and high-quality development of higher education.

Keywords: generative artificial intelligence; ChatGPT; higher education ecosystem; educational risk; risk warning

*基金项目: 本文为国家社科基金一般项目“全民终身学习视野下的国家在线教育体系发展研究”(项目编号: 20BSH053)的阶段性研究成果。

作者简介: 黄蓓蓓, 研究员, 博士, 研究方向为高等教育信息化, 邮箱为 963326617@qq.com。

收稿日期: 2023 年 10 月 9 日

编辑: 小米