

生成式人工智能背景下高校意识形态风险的三维探赜： 外在表征、内在机理与治理进阶

谭泽宇

(哈尔滨工程大学马克思主义学院, 黑龙江哈尔滨 150001)

[摘要]生成式人工智能广泛应用于高校,在其为教学、科研带来诸多变革的同时也潜藏意识形态风险。从外在表征的角度进行审视,风险主要表现为重塑高校意识形态传播样态、加剧意识形态管理难度、干扰师生意识形态认知。剖析生成式人工智能引发意识形态风险的内在机理,发现其价值预设和制造的信息茧房逐渐疏离高校主流意识形态、带来的技术崇拜与虚拟主体弱化了高校主流意识形态主导权、快捷生成海量视频与其新的传播模式加大了高校意识形态“把关”难度。对此,高校应从强化制度保障、夯实技术底座、提升主体理性等多角度发力,发掘生成式人工智能背景下高校意识形态风险的治理进阶,为中国式现代化提供坚强安全保障。

[关键词]生成式人工智能;意识形态风险;治理进阶

[中图分类号] TP18; G641; D630.8

[文献标识码] A

[文章编号] 2096-711X(2026)10-0149-04

doi:10.3969/j.issn.2096-711X.2026.10.048

[本刊网址] <http://www.hbxb.net>

习近平总书记强调:“要重视通用人工智能发展,营造创新生态,重视防范风险。”中国互联网络信息中心(CNNIC)发布的《生成式人工智能应用发展报告(2025)》显示,截至2025年6月,我国有5.15亿人使用过生成式人工智能产品,约每3名网民就有一人使用过生成式人工智能服务。从生成式人工智能用户受教育结构来看,本科及以上学历用户占比较高,约为28.9%;从应用场景来看,其广泛应用于教育与科研场景,帮助科研人员快速了解前沿科技、提炼关键信息,但因其应用于写论文、抄作业及隐含价值观偏见等而备受关注。由此可见,生成式人工智能改变了人一人、人一机交往方式和信息创造、传播范式,为我国高校教学、科研带来诸多变革的同时,也潜藏意识形态渗透的风险。高校正视生成式人工智能带来的意识形态风险,剖析内在机理、发掘治理进阶对巩固总体国家安全观、建立健全意识形态安全屏障具有重要意义。

一、审视:生成式人工智能背景下高校意识形态风险的外在表征

马尔库塞强调科学技术能增强资本主义统治在经济、政治、文化领域的全方位渗透。生成式人工智能带来的意识形态风险属于科技与意识形态互动范畴。生成式人工智能在高校教学、科研等方面的广泛应用,要求从其传播的新样态、管理难度和虚拟化意识形态阶级属性的层面认识高校意识形态风险的外在表征。

(一)新型人机关系重塑高校意识形态传播样态

1. 意识形态传播模式样态从“人一人”转向“机一人”

基于计算机、海量数据、算法的生成式人工智能已逐渐具备模拟创建人工语言的能力,逐步晋升为语言使用的新兴主体,这使原有作为人类认识传递基础的主体间性部分被“机一人”的客体间性所替代,随之,“人一人”之间的意识形态传播逐渐转向“机一人”之间的意识形态传播新样态。因生成式人工智能模型算法的规训不可避免地受训练者的影响强化或淡化某种意识形态的倾向,这种倾向将潜移默化地影响高校师生等使用者的思想。以ChatGPT为例,其给出的

答案或反馈天然带有训练者赋予的意识形态倾向,这种“机一人”的意识形态传播方式突破了原有“人一人”传播的模式,且因其传播主体为机器而使观点看似更加客观。

2. 意识形态跨时空共振风险加剧

意识形态是历史的,在人类漫长的社会发展历史与广阔的社会空间中,意识形态的主要矛盾大不相同。生成式人工智能训练过程中运用到的海量数据纵跨各阶段历史、横跨各地域思想,只要硬件支持足够的算力与存储空间,便能在数据层面实现无限扩张,进而在“答案”层面实现无限丰富。但这种无限丰富的“答案”并非人与人之间处于某个特定历史、空间范围内的对话,而是基于奖惩的训练产生,这可能使“问题”与“答案”之间产生跨越时空的鸿沟,而这种“时空错乱”的答案本身蕴含的意识形态倾向将加剧时空共振的风险。

(二)海量数据、算法推荐加大高校意识形态管理难度

1. 海量数据:加剧信息溯源难度

ChatGPT 3.0有1750亿参数,业内猜测ChatGPT 4.0的参数已达到1.8万亿。ChatGPT每一次更新换代的参数总量都呈指数增长,人类历史上产生的、互联网上存在且日益增长的数据等均转化为生成式人工智能的原料,这种原料或经过训练而产生的原料天然或后天可能带有资本主义意识形态倾向,而且这种可能带有资本主义意识形态倾向的“答案”也将重新进入模型成为新一轮的数据原料。

信息溯源是高校意识形态管理的重要内容,海量数据、二次海量数据的生成给高校意识形态管理带来巨大挑战,多数高校尚未有完备的信息溯源技术手段,加之生成式人工智能隐藏式地灌输资本主义意识形态,增加了高校意识形态内容溯源的难度和成本。

2. 算法推荐:加剧意识形态渗透风险

生成式人工智能产生“答案”不是线性的单向流程,而是“数据—算法—训练—答案(二次数据)”的闭合循环流程。人机之间的多次互动,加之生成式人工智能获取使用者的身份、偏好,灌输给使用者想看到的信息,这使主客体关系发生异化。

收稿日期:2025-12-18

基金项目:本文系2023年度教育部人文社会科学研究青年基金项目“总体国家安全观视阈下高校安全风险防范化解研究”阶段性成果(项目编号:23YJC710076)。

作者简介:谭泽宇(1996—),男,山东烟台人,哈尔滨工程大学马克思主义学院讲师,主要从事网络思想政治教育、总体国家安全观研究。

人机地位颠倒使意识形态渗透风险增加,特别是对三观尚未成型的青年学生群体。在应用过程中,其模型、算法最大限度地捕捉使用者偏好,并以对话等交互的形式潜移默化地、迎合式地灌输资本主义意识形态,而人们“只听我们选择的东西和愉悦我们的东西”,这种迎合口味的算法推荐加剧了意识形态操控风险。

(三)信息茧房、信息迷雾干扰师生对意识形态的认知

1. 信息茧房:加剧意识形态认知偏执

人使用互联网的过程是主动选择自己偏爱或有用的信息,这种使用痕迹成为生成式人工智能的数据之一,其通过与人“对话”也在分析使用者的喜好与习惯,并通过“对话”对分析结果进行验证,最终确定生成符合使用者价值选择的“答案”。

高校青年学生往往信息安全保护意识较弱,且生成式人工智能工具往往带有数据爬取的霸王或隐性条款,当使用者的偏好或研究领域被掌握,为增加用户粘性,会根据使用者的喜好主动推送符合用户价值取向的内容,看似是人工智能更加智能地迎合,实则是使用者更加麻木、被动地接受信息茧房的束缚,进一步加剧了意识形态的偏执认知。

2. 信息迷雾:虚化意识形态阶级属性

西方国家对我国意识形态攻击愈演愈烈,用于干扰公众认知的人造信息迷雾有扩散迹象。高校师生的个人隐私、行为偏好、研究领域等海量数据是西方国家制造信息迷雾的重要原料,且生成式人工智能产生的信息迷雾呈现拟真性和针对性。

生成式人工智能的“机—人”交互模式相较于“人—人”交互看似更具客观性,且其表述方式更具有“理中客”的特点,其“答案”的价值取向更具有迷惑性与政治危害性。例如:提问 ChatGPT“工人阶级在社会主义社会中的地位”,其回答倾向于“逐渐改善、关注、保护”等话题,这种去政治化、看似客观的表达实则与马克思主义基本原理相违背,且因看似客观而易于被高校学生接受,实则虚化了意识形态的阶级属性。在此基础上,生成式人工智能凭借“理中客”的形象,制造了一种看似真实的“信息迷雾”,这种“迷雾”看似逻辑自洽、细节丰富,但存在诸多虚假叙事,易使师生陷入认识论危机,即难以区分事实与虚构、真理与谬误。尤其值得警惕的是,生成式人工智能常以一种去情感化、超然、中立的“技术理性”面目出现,其生成的文本通过模拟学术话语,巧妙地将特定价值立场包装成共识性、公认的知识,从而在潜移默化中完成对意识形态阶级属性和斗争性的消解,形成了更为隐蔽的认识论层面的解构。

二、剖析:生成式人工智能诱发高校意识形态风险的内在机理

高校防范使用生成式人工智能带来的意识形态风险,须在把握其运行机制及制造信息茧房、迷雾等外在表征的基础上,向其生成的内容、形式、传播等层面深入,多角度探究其诱发高校意识形态风险的内在机理。

(一)内容层面:疏离高校中的主流意识形态

1. 价值预设内含意识形态属性

“作为资本,自动机在资本家身上获得了意识和意志。”技术是属人的行为,技术必然包含人的目的、价值、知识、观念、意志等因素。生成式人工智能技术中必然蕴含意识形态属性。

生成式人工智能蕴含的意识形态属性来源于开发者本身及训练者在训练过程中自觉或不自觉的价值预设。这种意识形态并未以显性标签的形式粘贴在生成式人工智能工具上,而是隐藏在其数据的倾向性选择、算法的选择性设计与训练中。数据与算法层面,发达资本主义国家起步较早,往往选择其意识形态体系下的数据,算法也潜在地迎合资本主义偏见和喜好,因数据与算法是生成式人工智能的核心要

素,发达资本主义国家掌握了数据与算法,就掌握了整个行业的发展模式与标准;模型训练层面,发达资本主义国家对生成式人工智能的奖励与惩戒必然基于其价值判断与选择,其中的意识形态输出不言而喻。

当前,全球大语言模型的研发与训练高度集中于少数西方科技巨头,其背后是庞大的资本投入、数据垄断与算力霸权。中文语料在训练数据中相对匮乏及依赖西方开源或闭源模型,使得国产化应用在底层逻辑上难以完全摆脱“原始设定”。如我国部分生成式人工智能工具起步较晚、数据量较小,往往基于并受制于西方发达国家的技术架构,为减少研发成本、提高运行速度和智能化程度,通常采用“语言翻译+”的开发模式,这种模式实质上是中文翻译成英文,而后进入西方发达国家生成式人工智能的模式,看似输出中立知识,实则可能已经过筛选、过滤和重构,系统性地边缘化或曲解了马克思主义理论,从而在知识生产的源头开始疏离主流意识形态,使生成的“答案”不可避免地含有资本主义的价值预设,对我国主流意识形态产生巨大冲击。

2. 个性化定制实则制造信息茧房

马斯克曾表示“Open AI 已经成为一家由微软控制的闭源、追求利益最大化的公司”,闭源的技术及逐利的公司必然在资本裹挟中抛弃社会责任,其受资本影响操纵公众、控制认知的主要途径是通过所谓个性化定制制造信息茧房。

随着生成式人工智能的广泛应用,个性化推荐已升级为更精准的个性化定制,使用者只要输入关键词或语句,生成式人工智能工具就能根据高度修正的算法和模型生成符合用户喜好的信息。长期接受这种信息的青年学生会形成思维定式,一方面不利于发展发散性和创造性思维,成为生成式人工智能的仆从;另一方面会在驯化过程中丧失主观判断,进而被长时间、隐藏式的意识形态灌输影响。

(二)形式层面:弱化高校中主流意识形态主导权

1. 技术崇拜冲击主流意识形态宣传主体权威

在高校教育者与受教育者的互动关系中,意识形态伴随知识的传递而传播。随着科技发展,受教育者获取知识的渠道已突破了传统的课堂讲授。随着生成式人工智能的快速发展、更新换代和海量数据吸纳,其在知识储备、逻辑推理、灌输方法等层面已超越个体人类的极限,在受教育者眼中成为“无所不能”的存在。加之生成式人工智能的拟人化表达,其输出的文本、图片、音频、影像、代码等在各方面均有优异表现,甚至超过该领域的专业人士,高校学生甚至教师长期使用生成式人工智能工作将形成行为习惯,放松了对其“答案”的思考与理性判断。这种技术崇拜甚至盲从极易转化为对其输出“答案”的盲从,进而形成对其“答案”蕴含意识形态的盲从。这种盲从将冲击传统主流意识形态宣传主体的权威,甚至将原有意识形态宣传主体转化成为其仆从。

2. 虚拟主体冲击主流意识形态宣传主体形象

意识形态宣传主体的形象影响意识形态工作的实效。高校思想政治工作者是高校意识形态宣传的主体,其言行关系到学生“学懂信用”。

师生在生成式人工智能工具“对话”过程中,因其海量知识数据、自然语言表达、严谨运算逻辑、生动视听觉形象,易将其视为有思维的客观实在。对其过度使用,一方面弱化高校思想政治工作者思考判断能力,迷信技术权威,从客观实际层面弱化高校主流意识形态宣传主体形象;另一方面使得高校学生沉迷于生成式人工智能的价值输出和创造的虚拟世界,从主观视角淡化高校主流意识形态宣传主体的形象,最终削弱意识形态宣传主体的凝聚力和引领力。

(三)传播层面:提升高校意识形态“把关”难度

1. 快捷生成海量音视频提高意识形态把关成本

Sora 能够根据自然语言描述生成一个逼真的物理世界。这种技术日趋成熟,标志着“有图有真相”的时代已经结束,

低级红、高级黑音视频、图片的制作成本降低。同时,相较于传统的文本把关,对音视频、图片的把关要求从业者具有更高的美学、艺术学造诣,提高了高校意识形态的把关成本。

一方面,音视频、图片内涵的意识形态及价值倾向更隐蔽,往往通过细微区别潜移默化影响人的思想,例如“毒教材”中口歪眼斜的人物形象等。而生成式人工智能制作的音视频、图片相较于传统的生产模式,效率更高,隐蔽性更强,更难被高校“把关人”识别;另一方面,随着抖音、快手、Bilibili等自媒体平台快速发展,高校师生自觉或不自觉地从这些平台上获取信息,而生成式人工智能很轻易便可生成在互联网上广泛流通的视频,其传播速度快、受众群体广、制作成本低、隐含意识形态倾向等特点,为高校意识形态把关工作带来了巨大挑战。

2. 新传播模式导致把关失效

列宁强调“社会民主主义意识只能从外面灌输进去。”但灌输什么样的意识形态必须经过选择,选择的过程实质是一种意识形态的把关过程,大量的信息只有经过把关才能进入主流传播渠道。但生成式人工智能的快速发展冲击了传统的高校意识形态把关模式。

一方面,恶意操纵下的生成式人工智能工具能够伪装成人类侵入互联网信息平台,以恶意灌水等方式冲击正常人类的讨论话题,研究机构 Imperva 发布的《2025 年恶意机器人报告》和多名学者在研究中指出,2024 年自动化流量十年来首次超越人类流量,占全球互联网流量的 51%,其中恶意机器人的占比飙升至 37%,其生成的虚假账户恶意冲击 X 平台(原 Twitter)、微博等国内外互联网平台,这些账户因经常性持续更新推文而被大众认为是人类账户,一旦这类账户被恶意操纵,将成为操纵议题、转移话题、挑唆对立情绪的主体,造成高校意识形态把关失效;另一方面,高校传统把关难以穿透闭源“算法黑箱”,传统把关是通过意识形态输出者、内容及效果进行把关,但生成式人工智能发展使意识形态输出者隐藏在算法、模型背后,输出内容多样且不可控,传统把关不能穿透“算法黑箱”从根本进行意识形态把关,易陷入“头痛医头脚痛医脚”的困境。

三、应对:生成式人工智能背景下高校意识形态风险的治理进路

“建设强大凝聚力和引领力的社会主义意识形态,是全党特别是宣传思想战线必须承担起的一个战略任务。”高校党委必须特别重视生成式人工智能这一新兴事物对高校意识形态工作带来的机遇与挑战,主动应变求变,在守好安全底线的同时谋求发展,通过强化制度保障、夯实技术底座、提升主体理性等措施治理生成式人工智能带来的高校意识形态风险。

(一) 强化制度保障:坚持马克思主义在意识形态领域指导地位的根本制度

要强化主流意识形态在人工智能领域的价值引领作用。首先,高校要严格落实意识形态工作责任制。高校各级党组织应主动应对、研究新事物,将生成式人工智能可能带来的潜在意识形态安全风险纳入党委会议、党政联席会议事范围,科学理性关注研判生成式人工智能风险,积极关注师生思想动态。

其次,要以马克思主义理论的观点正确认识生成式人工智能潜在的意识形态倾向。马克思指出:“思想、观念、意识的生产最初是直接同人们的物质活动,同人们的物质交往,同现实生活的语言交织在一起的”。生成式人工智能的使用看似是师生与机器的互动,从思想内容角度而言,是师生与整个人类群体所产生的思想、观念的互动,从“人一机一人”运作机理角度而言,是师生与算法开发者、训练实施者的互动,本质上还是人与人的互动。高校党委、师生要以马克思主义理论的观点正确认识其性质、明晰其互动本质,才能预

防意识形态安全风险。

(二) 夯实技术底座:推动生成式人工智能技术安全智能升级

“科学技术作为‘背景意识形态’构成意识形态整体结构中的隐性层面,以潜移默化的方式发挥作用”。高校在面对生成式人工智能技术渗透风险时,要发挥好科技创新策源地的作用,剖析生成式人工智能运作机理,从技术层面支撑我国生成式人工智能技术安全升级。

首先,高校要把握数据、运行模型、核心算法等生成式人工智能运行的关键环节。一方面,数据安全直接影响生成式人工智能的意识形态安全,在数据信息保护层面要做到严格保密管理,包括用户信息数据、生物数据及行为习惯偏好等,避免数据泄露造成安全风险;另一方面,运行模型、核心算法中隐含的意识形态倾向将直接影响生成式人工智能的意识形态倾向,因 Open AI 公司的闭源算法未向其他国家、公司开放,无法从其算法底层逻辑分析其是否带有或带有何种程度的意识形态倾向,因此我国高校要加快生成式人工智能技术的自主研发,将我国主流价值观融入算法、模型,避免发达国家国家算法偏见冲击我国高校意识形态工作。

其次,高校要发挥交叉学科优势,推动全链条监管。从国际视角看,发达资本主义国家注重对生成式人工智能的全过程、全链条监管,欧洲采取全面立法策略,强调基于风险的全过程监管方法;美国通过密集出台法规和政策,在顶层设计和基础设施建设、人才培养、跨部门合作等操作层面进行全过程布局。生成式人工智能技术的开发与应用涉及人工智能、网络安全、马克思主义理论、伦理学、法学等多个学科,而高校天然具有多学科、强交叉的科技创新基础与优势,要汇集校内乃至国内优秀技术研究人员与马克思主义理论研究专家,将马克思主义理论深度整合到人工智能的研发流程中,在模型构建之初,就将社会主义核心价值观、总体国家安全观等主流意识形态作为核心参数和优化目标嵌入算法逻辑、训练模型,同时在输出端进行过滤,共同推动我国自主研发可控的生成式人工智能技术,技术把关与意识形态把关相结合,摆脱西方发达国家的技术封锁与意识形态侵略。

(三) 提升主体理性:提高师生数智素养与风险防范意识

从生成式人工智能的运行机理可见,其应用给高校带来的意识形态风险主要源于研发人员自带的意识形态倾向和使用者对潜在意识形态风险认识、防范能力不到位。

首先,针对高校研发生成式人工智能技术的高水平科技工作者,要筑牢意识形态屏障,通过物质保障与思想引领确保其免受资本主义意识形态的侵扰,进而确保其研发的技术中内嵌社会主义核心价值观等我国主流意识形态。一方面,要保证关键核心研发人员的物质生活条件,建立屏障将西方渗透势力“拒之门外”,确保不被经济利益诱惑的同时调动积极性;另一方面,要强化对关键核心研发人员的思想引领和人文关怀,加强主流意识形态教育,强化职业道德、引导其坚守人民立场,推动意识形态工作与科学研究工作的有机融合,确保生成式人工智能技术在正确的道路上运行。

其次,高校要面向师生,特别是思政课教师定期开展科学的数智素养培训,提高师生识别、规避生成式人工智能带来意识形态风险的能力。一方面,要适当开办人工智能、技术伦理相关的选修课程、学术讲座,提高师生对潜在风险的认识;另一方面,要充分发挥辅导员队伍年轻化、专业化优势,常态化向学生开展风险防范教育,引导学生警惕“算法黑箱”和数据泄露可能引发的风险,切实提高高校师生数智素养和风险防范意识。

参考文献:

[1] 中共中央政治局召开会议分析研究当前经济形势和

经济工作[N].人民日报,2023-4-29(1).

[2](美)赫伯特·马尔库塞.单向度的人:发达工业社会意识形态研究[M].刘继,译.上海:上海译文出版社,2006:10.

[3](美)凯斯·R.桑斯坦.信息乌托邦:众人如何生产知识[M].毕竞悦,译.北京:法律出版社,2008:8.

[4]马克思恩格斯全集(第42卷)[M].人民出版社,2016.

[5]吴致远.有关技术中性论的三个问题[J].自然辩证法通讯,2013,35(6).

[6]习近平.论党的宣传思想工作[M].北京:中央文献出版社,2020.

[7]列宁全集(第6卷)[M].北京:人民出版社,1986.

[8]马克思恩格斯选集(第1卷)[M].北京:人民出版社,2012.

[9]俞吾金.从科学技术的双重功能看历史唯物主义叙述方式的改变[J].中国社会科学,2004(1).

[10]谭泽宇,潘强.补短板释放“人工智能+”效应[N].经济日报,2025-2-14(5).

A Tripartite Inquiry into Ideological Risks in Higher Education Institutions under Generative Artificial Intelligence: Exogenous Manifestations, Endogenous Mechanisms and Governance Pathways

TAN Ze-yu

(College of Marxism, Harbin Engineering University, Harbin Heilongjiang 150001, China)

Abstract: Generative Artificial Intelligence (GAI) has been widely adopted in higher education, revolutionizing teaching and research while posing latent ideological risks. Externally, these risks manifest in reshaping ideological communication paradigms, amplifying governance complexities, and distorting stakeholders' cognitive alignment with mainstream ideologies. Mechanistic analysis reveals that GAI's intrinsic value presuppositions and information cocooning effects gradually alienate campus ideological cohesion, while technological fetishism and virtual agency erosion undermine institutional discursive authority. Additionally, mass-produced multimodal content and novel dissemination modes challenge conventional ideological gatekeeping. To mitigate these risks, this study proposes a multi-dimensional governance framework encompassing institutional safeguards, algorithmic value embedding, and epistemological resilience enhancement, thereby securing ideological safety for Chinese modernization.

Key words: GAI; ideological risks; governance approaches

(责任编辑:桂杉杉)

(上接第148页)

[7]Zhang, Y., Xiong, F., Xie, Y., Fan, X., Gu, H. The Impact of Artificial Intelligence and Blockchain on the Accounting Profession[J]. Ieee Access, 2020(8):110461-110477.

[8]章雯华.高校财务管理信息化系统功能设计思考[J].财会通讯,2017(1):121-123.

[9]李闻一,李栗,曹菁,陈新巧.论智慧财务的概念框架和未来应用场景[J].财会月刊,2018(5):40-43.

[10]赵春荣,王天,张向阳.“双高计划”下高职院校财务管理改革与创新的思考[J].现代职业教育,2020(30):118-120.

[11]丁友刚,吴玮琼,刘阳,等.高校财务信息化建设的历程、趋势与内涵[J].会计之友,2021(9):116-124.

[12]韦德洪,陈势婷.论智慧财务管理的内涵、外延、特点与应用[J].会计研究,2022(5):40-48.

Application Practice and Risk Prevention Exploration of Cloud Data Sharing Platforms in University Financial Management Driven by Artificial Intelligence

CUI Ji

(Changzhou Vocational Institute of Industry Technology, Changzhou Jiangsu 213000, China)

Abstract: In the era of digital economy, data has emerged as a new type of production factor. Driven by artificial intelligence, the positions, processes, content and models of financial work in universities have undergone profound changes. The cloud-based data sharing platform for financial management has become a research hotspot in both academia and practice. Universities should promote dynamic alignment of organization and processes, clarify responsibilities and collaboration mechanisms, and accelerate the establishment of a high-end talent support system by building a dual-function, integrated, and risk-embedded financial management cloud data sharing platform. This will further expand intelligent application scenarios, construct a full-chain management system encompassing “data collection, intelligent analysis, decision support, and risk warning” in finance, and achieve the coordinated integration of value flows such as data flow, capital flow and talent flow, thereby enhancing the modernization of university governance systems and capabilities.

Key words: artificial intelligence; financial management cloud data sharing platform; dual-function; integrated; risk embedding

(责任编辑:陈思婷)