Advances in Social Sciences 社会科学前沿, 2025, 14(12), 29-38
Published Online December 2025 in Hans. https://www.hanspub.org/journal/ass
https://doi.org/10.12677/ass.2025.14121062

Hans汉斯

高校图书馆应用生成式人工智能的风险与对策研究

吴　进，张艳艳，尹　晖，庞　萍，王　栋*

中国海洋大学图书馆，山东　青岛

收稿日期：2025年8月25日；录用日期：2025年11月27日；发布日期：2025年12月5日

摘　要

文章聚焦高校图书馆应用生成式人工智能的风险与对策，首先，阐述生成式人工智能赋能高校图书馆建设的场景，包括参考咨询服务升级、信息素养教育创新、科技查新效能提升、阅读推广服务优化及古籍信息处理革新。接着，分析应用过程中面临的风险，如AI幻觉风险、可版权性不明、权利归属争议、用户信息被不合理使用以及系统稳定运行风险。最后，提出对策：积极履行告知义务、协同发力缓解AI幻觉风险、关注跟踪立法动态和司法实践、做好权利义务分配、加强用户信息保护以及开展分类专项培训。

关键词

生成式人工智能，高校图书馆，风险，对策

Study on Risks and Countermeasures of Generative AI Application in University Libraries

Jin Wu, Yanyan Zhang, Hui Yin, Ping Pang, Dong Wang*

Library of Ocean University of China, Qingdao Shandong

Received: August 25, 2025; accepted: November 27, 2025; published: December 5, 2025

Abstract

This paper focuses on the risks and countermeasures of applying generative artificial intelligence (GAI) in university libraries. Firstly, it elaborates on the scenarios where GAI empowers the development of university libraries, including the upgrading of reference and consultation services, the

*通讯作者。

文章引用: 吴进, 张艳艳, 尹晖, 庞萍, 王栋. 高校图书馆应用生成式人工智能的风险与对策研究[J]. 社会科学前沿, 2025, 14(12): 29-38. DOI: 10.12677/ass.2025.14121062

innovation of information literacy education, the improvement of sci-tech novelty search efficiency, the optimization of reading promotion services, and the innovation of ancient book information processing. Secondly, it analyzes the risks encountered in the application process, such as the hallucination risk of generated content, unclear copyrightability, disputes over rights ownership, unreasonable use of user information, and the risk of unstable system operation. Finally, it puts forward countermeasures, including actively fulfilling the obligation of notification, making joint efforts to mitigate the risk of AI hallucination, paying attention to and tracking legislative developments and judicial practices, properly allocating rights and obligations, strengthening the protection of user information, and carrying out classified and specialized training.

Keywords

Generative Artificial Intelligence, University Libraries, Risks, Countermeasures

1. 引言

自美国 OpenAI 公司推出生成式人工智能产品 ChatGPT 以来，人工智能生成内容技术(AIGC)凭借强大的功能性迅速成为科技领域的现象级技术，其应用场景从自动生成新闻稿件、文学作品，到辅助制定个性化医学诊疗方案、生成高质量音视频素材，催生了大量创新成果，推动着社会各行业转型升级。在教育信息化的背景下，作为服务学校人才培养和科学研究的学术性机构，高校图书馆也受到了生成式人工智能的深刻影响。一方面，生成式人工智能凭借强大的自然语言理解能力、海量信息组织处理能力和内容生成能力，为高校图书馆信息服务提升带来前所未有的机遇。另一方面，其应用也面临 AI 幻觉、权利归属争议、信息安全及法律伦理等诸多挑战，亟需系统性回应。

目前，学界就生成式人工智能在图书馆的融合应用开展了系列研究，主要集中在两个维度：一是探讨技术赋能场景，如参考咨询、智能问答、知识服务、信息搜索、阅读障碍服务、数字阅读推广等[1]-[8]；二是识别技术应用带来的潜在风险，涵盖阅读推广、科研数据服务中的版权风险，参考咨询服务、知识服务中的伦理风险，数据收集、存储、处理等阶段的数据安全风险，智慧图书馆建设中的入侵风险、网络舆情风险、超量下载风险以及馆员面临的岗位替代风险等[4] [9]-[15]。提出的应对策略包括，对合理使用条款的适用范围做扩张解释，增设人工智能科研数据适用著作权例外，构建面向人机协同学习关系的服务体系、多元协同参与的敏捷-韧性治理模式、全生命周期风险治理框架，完善网络信息安全制度，限制应用范围，提高相关人员数字素养，树立馆员人工智能意识等举措[4] [10] [13]-[16]。上述研究覆盖面较为广泛，但仍存在一定局限：部分研究仍停留在一般性的风险列举层面，缺乏可落地的应对路径；与技术应用密切相关的部分风险如生成内容的可版权性认定、权利归属认定及系统稳定运行风险等缺少系统性研究。本文旨在填补现有研究在部分风险领域的空白，聚焦五大核心风险，将法律、伦理、管理与技术服务相融合，提出更具针对性和可操作性的对策，以期不断深化高校图书馆应用生成式人工智能相关研究。

2. 生成式人工智能赋能高校图书馆建设的场景

2.1. 参考咨询服务升级

参考咨询作为连接用户与知识资源的纽带，承担着解答信息咨询、提供学术支持等重要职能。传统

服务模式受限于岗位配置、馆员素养及服务时长等因素，难以满足师生个性化、专业化、多元化、即时性的信息需求。生成式人工智能凭借快速检索、高效组织、精准生成信息和友好交互等能力，为参考咨询服务形态重塑与效能提升提供了新路径。

基于生成式人工智能的“AIGC＋ 参考咨询”模式具备以下优势：一是生成式人工智能依托海量数据资源和卓越的自然语言处理能力，能够精准理解用户的咨询意图，可向读者提供较为准确、全面的咨询服务。二是不同于人工咨询，“AIGC＋ 参考咨询”突破了时间和空间限制，可随时随地响应用户咨询需求，实现 7 × 24 小时不间断在线服务。在寒暑假、论文开题以及课题申报等特殊时期，能够有效缓解咨询高峰时段咨询馆员的工作压力。三是以 DeepSeek 为代表的生成式人工智能产品还支持本地数据投喂和数据本地存储，图书馆可将学校学科设置、馆藏资源、自建特色数据库及咨询服务数据等各类信息投喂接入模型，为读者提供更精准的专业化服务。

2.2. 信息素养教育创新

文献检索教学是高校图书馆提升学生信息素养的重要载体。传统的文献检索教学方式存在教学形式单一、内容枯燥、评估机制不完善等问题，难以适配学生多元学习需求。生成式人工智能为文献检索教学质量提升注入了新动能。

以 DeepSeek 为代表的主流产品支持文本、图像、语音等多模态检索，具备多语言处理、智能推荐、深度思考及联网检索等功能，在信息搜寻与组织领域展现出显著应用潜力。教学馆员可以将生成式人工智能工具的使用融入教学课程体系，借以拓展教学内容维度，提升教学效果。相关调查显示，此举可显著提升教学满意度[17]。教学馆员还可借助生成式人工智能辅助完成课程优化设计，生成课程大纲，提供针对性强的教学案例，开发智能教学助手等，推动教学模式的创新与变革。

2.3. 科技查新效能提升

科技查新是指查新机构根据查新委托人提供的有关科研资料查证其研究的结果是否具有新颖性，并做出结论，为科研立项、成果鉴定、奖励申报、专利申请等提供客观依据。高校科技查新站一般内设在图书馆，主要面向本校及其他单位从事课题研究的教师、研究生及科研人员提供有偿科技查新服务。作为高校图书馆信息服务体系的重要一环，当前查新工作面临人力成本高、专职查新员数量有限、查新员学科背景单一等问题，难以应对社会公众的大量委托和科研立项、奖项申报等特定时间段的集中委托。生成式人工智能技术为科技查新服务效能提升带来了新机遇。

查新员可借助生成式人工智能快速了解查新项目的基本情况，例如在“体外膜肺氧合技术”查新中，快速梳理该技术的原理、类型、风险等内容[17]；可借助生成式人工智能完成同义词扩展工作，例如在“肿瘤基因检测”相关查新中，人工智能给出的“基因测序”“DNA 检测”“遗传检测”“分子诊断”等均为有效的核心同义词[17]；可借助生成式人工智能完成查新比对总结，例如，将检索到的相关文献以附件形式上传后，让人工智能执行相关比对指令并生成总结。同时，还可将收费标准、工作时限、发票开具等常见问题投喂给大模型，开发查新咨询智能助手，进一步承担科技查新咨询任务。

2.4. 阅读推广服务优化

《普通高等学校图书馆规程》第三十二条明确要求图书馆应积极参与校园文化建设，开展阅读推广等文化活动。各高校图书馆纷纷响应政策号召，开展了多种形式的阅读推广活动。然而，传统模式存在活动形式同质化，用户参与度低等问题，生成式人工智能技术为图书馆阅读推广服务优化提供了新助力。

生成式人工智能技术能够构建用户画像，根据用户的借阅记录、阅读偏好及学科专业等信息，推荐符合用户兴趣的图书、期刊、论文等信息资源，提升资源推广精度。借助生成式人工智能技术构建虚拟

书友交互系统，可通过与用户分享阅读心得等方式增强用户体验与粘性。阅读推广馆员借助生成式人工智能可快速形成特定主题书单，支撑特色阅读活动开展，已有高校图书馆开展了相关实践[18]。馆员还可借助生成式人工智能对阅读趋势进行预测，例如热门书籍、新兴学科、社会热点等，帮助图书馆及时调整资源建设和阅读推广服务策略。

2.5. 古籍信息处理革新

高校图书馆作为古籍保护与传承的重要阵地，拥有丰富的古籍资源。据文化和旅游部 2020 年统计，全国 203 个古籍重点保护单位中包含 55 所高校图书馆[19]。古籍是历史文化的重要载体，具有独特的公共文化服务与社会教育价值。《“十四五”公共文化服务体系建设规划》明确提出加强古籍资源应用，并倡导高校图书馆向公众开放古籍服务。古籍资源的有效应用离不开精准高效的信息处理，核心在于通过技术手段实现对古籍内容的智能化标注、深度揭示与系统分析，从而充分挖掘其中蕴含的历史知识与文化价值。生成式人工智能在这一领域的应用价值主要体现在三个方面：其一，凭借卓越的总结与摘要生成能力，可快速提炼古籍核心内容，生成精准易懂的摘要，帮助读者把握文献主旨；其二，依托强大的语言转化能力，可实现文言文与白话文的精准转换，降低读者古籍阅读的语言门槛；其三，拥有多语种翻译功能，可将古籍内容转化为多种外语版本，方便具有不同文化背景的读者阅读，有效拓展古籍的传播范围与影响力。

3. 高校图书馆应用生成式人工智能的风险

高校图书馆在引入生成式人工智能技术助推智慧图书馆建设的同时，需警惕潜在的应用风险。

3.1. 生成内容的幻觉风险

生成式人工智能凭借强大的内容生成能力，在参考咨询、科技查新、信息素养教育、古籍信息处理等场景中展现出良好的赋能前景。然而，其幻觉(Hallucination)现象，即生成看似合理却与事实不符的虚假信息[20]，成为影响高校图书馆应用生成式人工智能技术的关键风险之一。

AI 幻觉被广泛认为是大语言模型的一个重大缺陷，XU.ZW 等学者通过构建形式化世界，从理论层面论证了大语言模型无法完全消除幻觉[21]。随着生成式人工智能技术的广泛应用，AI 幻觉缺陷对信息的真实准确性带来严重挑战，引发广泛担忧。2025 年 7 月，一篇关于 DeepSeek 就不实信息道歉的声明，经记者核查发现该声明系虚假信息，DeepSeek 官方从未发布有关道歉声明[22]。另外该声明中所引用的裁判文书亦无法从“中国裁判文书网”等官方渠道得到检索论证。部分自媒体未辨真伪加以引用，促成二次传播，引发广泛误读。对引入生成式人工智能技术的高校图书馆而言，馆员在科技查新或情报分析中，生成式人工智能可能虚构研究机构、学者姓名或实验数据；读者在使用相关服务时，生成式人工智能可能生成虚假的文献信息、馆藏数据或学术结论。曾有多起美国律师因使用 ChatGPT 等人工智能生成的虚构案例参与诉讼活动而受到法院制裁或驳回动议的案例[23]。与司法有关的法律材料，对事实和法律的准确性有着极为严格的要求，AI 幻觉所带来的相关风险于法律活动而言是极其致命的，对于追求严谨科学的高校学术科研而言亦是如此：生成的虚假信息可能误导师生科研决策，降低读者对图书馆的信任度，甚至破坏学术规范，引发学术不端。

3.2. 生成内容的可版权性不明

读者使用图书馆提供的生成式人工智能服务时，生成内容的可版权性认定成为绕不开的问题。目前，人工智能生成内容的可版权性尚未明确，致使图书馆、读者在使用、传播这些生成内容时面临“是否构

成侵权”“能否主张权利”的不确定性，影响着生成式人工智能在高校图书馆的应用。

(1) 从立法视角看，生成内容可版权性认定不明确。《著作权法》是认定人工智能生成内容是否具有版权的直接依据。《著作权法》第二条指出“作品”享有著作权，根据《著作权法》第三条的规定，“作品”是指文学、艺术和科学领域内具有独创性并能以一定形式表现的智力成果，包含文字作品、摄影作品等八种具体的形式，对于其他符合作品特征的智力成果亦可认定为作品。人工智能生成内容是否属于作品需要进行作品构成四要件认定，即生成内容是否属于文学、艺术和科学领域，是否具有独创性，是否具有一定的表现形式，是否属于智力成果。人工智能生成内容这一表现形式伴随科技发展而出现，现行著作权法在制定时未对此加以明确规制，现有司法解释亦未就此问题进行回应。因此，从著作权法的视角来看，现行版权法律体系以“人类创作”为核心构建权利框架，而人工智能生成内容的“非人类主体性”使其难以被现有法律概念所涵盖，导致权属认定从源头陷入模糊。

(2) 从司法实践视角看，生成内容可版权性判定不一致。虽然著作权法无明文规定人工智能生成内容是否具有版权，但是人民法院作为国家审判机关需要回应这一司法问题。通过检索“中国裁判文书网”，筛选出三个人工智能生成内容是否构成作品的相关案例。在北京菲林律师事务所诉北京百度网讯科技有限公司侵害署名权、保护作品完整权、信息网络传播权纠纷案中，法院认为涉案图形不符合图形作品的独创性要求，涉案文字作品虽具有独创性，但其非由自然人创作，故涉案分析报告不构成作品。该案系人民法院首次回应涉计算机软件智能生成内容是否构成作品的案例[24]。在深圳市腾讯计算机系统有限公司诉上海盈讯科技有限公司侵害著作权及不正当竞争纠纷案中，法院认为涉案文字满足著作权法对文字作品的保护条件，属于受著作权法保护的文字作品。该案系全国首例认定人工智能生成内容构成作品的案例[25]。在李某诉刘某侵害作品署名权、信息网络传播权纠纷案中，审理法院认定涉案图片属于受著作权法保护的图片作品。该案系我国首例人工智能文生图案例[26]。综合来看，人民法院在司法实践中对人工智能生成内容是否具有可版权性的判决标准尚未达成一致，这种司法裁判差异进一步加剧了可版权性认定的复杂性。

3.3. 生成内容的权利归属争议

借助图书馆提供的生成式人工智能技术生成内容的权利归属认定较为复杂，主要涉及三方主体：图书馆(学校)、技术供应商和用户。

从用户视角看，读者是生成式人工智能的具体使用者，提示词的确定、提示词的出现顺序、相关参数的设定、生成内容呈现的方式等均来自用户的构思，无不体现着用户个性化的选择和安排。在李某诉刘某侵害作品署名权、信息网络传播权纠纷案中，审理法院亦认定涉案图片从构思到选定，均体现着李某的智力投入，涉案图片具备了“智力成果”要件，属于受著作权法保护的作品。

从图书馆视角看，一方面，图书馆代表学校是生成式人工智能的购买引进方，作为合同双方之一，享有相关权利，承担相应责任。另一方面，对于本地化部署的生成式人工智能而言，往往需要进行二次训练优化。既需要图书馆提供馆藏资源等本地数据进行数据投喂，这些数据不乏独有的自建馆藏资源，又需要图书馆对人工智能生成内容的整体框架、应用场景等进行针对性设计。图书馆在资源数据提供与模型适配方面的付出，为内容生成提供了必要条件，间接影响着生成内容的输出，隐含着图书馆应当在权利分配中有所体现。

从技术供应商视角看，算法决定人工智能生成内容的逻辑与方式，作为算法开发者，供应商在模型的训练与优化上投入了大量人力、物力与技术资源，在内容生成中扮演着重要角色。而且作为技术出让一方，是合同主体之一，对权利义务条款的拟制具有非常大的话语权，供应商可在服务协议中保留对生成内容的某些权利。

3.4. 用户信息被不合理使用风险

模型优化的内在需求驱动厂商搜集、使用用户信息。人工智能技术的一大优势是可以通过不断学习来训练改进模型。OpenAI 在发布的早期技术报告中指出，对于在同用户交互过程中搜集的用户信息及隐私，公司不保证 ChatGPT 在与其他用户的聊天过程中不使用该素材[1]。OpenAI 在“隐私和政策－政策常见问题－如何使用您的数据来提高模型性能”中强调，“当您与我们分享您的内容时，它有助于我们的模型更准确、更好地解决您的特定问题，还有助于提高其总体功能和安全性。我们不会使用您的内容来营销我们的服务或创建您的广告资料，我们会使用它来使我们的模型更有帮助。例如，ChatGPT 可以通过进一步训练人们与它的对话来改进，除非您选择退出。”[27]对于“退出”设置，OpenAI 亦给出了操作说明：针对网页版需要点击“个人资料图标－选择设置－转到数据控件－关闭‘为所有人改进模型’”；针对移动设备需要“打开侧边栏菜单－点击个人资料图标－选择数据控件－关闭‘为所有人改进模型’”[28]。再看主流国产大模型代表性产品之一——“豆包”，在软件“设置”里并没有直接的个人信息使用同意与否的选项，经笔者再三尝试，发现相关设置放置于“账号设置－个性化内容推荐”项下，而且隐私政策也不是直接明示，而是放在“个性化内容推荐－了解更多”里面，具有相当强的隐蔽性。因此，这两款国内外主流生成式人工智能产品关于个人信息的关闭设置对公众用户而言不是显而易见的。同时，两款应用的默认设置均是用户同意使用个人信息。在此情形下，公司使用用户信息的行为难以认定为符合《个人信息保护法》第十四条的规定：“基于个人同意处理个人信息的，该同意应当由个人在充分知情的前提下自愿、明确作出。”这极易造成用户个人信息被不合理使用，甚至涉及隐私侵犯。

3.5. 系统稳定运行风险

生成式人工智能系统的稳定运行是图书馆依托该技术拓展服务边界、提升服务效能的核心前提，但图书馆在推动系统落地与长期运营过程中，可能面临技术运维升级成本高企与馆员应用能力不足等风险。

(1) 技术运维升级的经费压力。图书馆引入生成式人工智能系统并实现提供持续稳定服务，除需要支付软硬件设备费、技术授权使用费等费用外，还需要承担持续性的维护升级费用，且该类支出存在不确定性。首先是维护成本。随着用户使用频次的增加、服务场景的拓展，数据存储容量、算力等需求将同步增长，可能导致服务器扩容、云端资源租赁等维护成本超出经费承受范围。其次是版本迭代升级费。当前生成式人工智能技术发展迅速，主流大模型产品不断推出新版本，如果不及时跟进升级，可能出现系统功能滞后、生成内容质量不理想等问题，直接弱化服务效果。最后是故障应急处理费用。在日常运行维护过程中，不可避免会产生数据传输失败、模型响应异常、系统管理漏洞等技术故障问题，需要技术供应商提供技术支持，部分供应商如果凭借技术垄断地位收取高额应急服务费会导致图书馆陷入被动局面。

(2) 馆员应用能力不足风险。生成式人工智能作为科技发展前沿成果，对馆员的应用和管理能力提出了新要求。系统操作层面，如果馆员不能掌握本地数据投喂、参数调整设置等关键技能，可能导致系统无法充分发挥作用，形成“技术闲置”；故障处理层面，如果馆员不具备常见故障排查处理能力，过度依赖供应商技术支持，可能延长服务中断时间，加剧用户体验落差；服务匹配层面，如果馆员对生成式人工智能技术功能边界认知模糊，可能无法准确评估系统功能与用户需求的匹配度，导致“过度承诺”或“服务缺位”，均会影响用户对图书馆服务的满意度和信任度。

4. 高校图书馆应用生成式人工智能的对策

4.1. 积极履行告知义务

高校图书馆在提供生成式人工智能服务过程中，需以保障用户的知情权与选择权为核心，构建全流程、透明化的告知机制，确保用户对服务运行机制、数据使用规则及潜在风险等形成清晰认知。采取的

措施包括但不限于：

(1) 编制规范化服务说明文档。文档应详细阐述生成式人工智能服务的功能、使用方式，以及在服务过程中用户信息的收集、存储、使用和共享情况，包括是否会将用户数据用于模型训练等关键信息。

(2) 建立强制化首次告知流程。在用户首次使用生成式人工智能服务时，应通过全屏弹窗等不可规避的方式触发告知程序，弹窗内容应以突显方式提示用户阅读服务说明和隐私政策等与用户有关的关键信息，并设置分步确认机制，杜绝一键同意模式的模糊授权。

(3) 实施动态化信息更新机制。当模型功能迭代、数据政策调整或风险等级变化时，应在服务界面置顶公告，同步通过图书馆主页、微信公众号等渠道推送更新通知，确保用户能够及时了解服务变化情况。

4.2. 协同发力缓解 AI 幻觉风险

理论上 AI 幻觉系属固有缺陷，难以完全消除。高校图书馆可通过构建“数据优化、信息溯源、二次训练、建立幻觉数据库”的四维协同体系，降低 AI 幻觉对服务质量的影响。

(1) 数据优化：强化源头质量防线。大模型本地部署时，生成内容的准确性高度依赖图书馆本地数据的质量。如果馆藏编目数据存在字段缺失、学位论文研究方向标注错误、学科分类体系混乱等问题，模型可能基于错误样本进行不合理推测，进而生成虚构信息。对此，中国海洋大学在部署本地化 DeepSeek 时明确要求由各二级单位提供相关政策、制度等信息，以确保底层数据的准确性和权威性。因此，高校图书馆需建立标准化数据预处理流程，组建由学科馆员、技术人员及数据专家构成的工作组，对拟投喂的数据进行多轮清洗，尤其对于古籍数字化文本等特殊资源需独立进行 OCR 识别校对。

(2) 信息溯源：强化生成内容可验证性。通过技术配置对内容生成过程施加硬性约束，要求系统在输出学术性内容等要求严格的数据时标注信息来源，对原创性生成内容，详细说明推理依据，从机制上减少“无依据生成”，同时也方便读者溯源验证、参考引用。

(3) 二次训练：提升场景适配度。依托图书馆的核心资源，如权威数据库、学科知识库、机构知识库等数据，构建专业化训练数据集，进行二次训练，让模型优先学习经过学术验证的内容，减少对通用网络数据的依赖。针对科技查新、文献综述生成等服务场景，设计开展专项训练，以适配特定场景需求。

(4) 建立幻觉案例库：构建风险共治生态。在用户使用窗口设置醒目的“幻觉举报”入口，提供“事实错误”“逻辑矛盾”“来源虚构”等标准化分类标签，方便用户将使用过程中遇到的各类型幻觉情形反馈给图书馆。图书馆应定期汇总分析幻觉案例库，总结错误类型，分析涉及领域，研判影响程度，为模型后续优化提供数据支持。同时推动跨主体协作：与技术供应商合作，共享幻觉案例数据库，助力供应商开发针对性修复算法，共同推动模型改良升级；联合同类型高校图书馆组建学科应用联盟，通过专题研讨会等形式共享幻觉应对经验，提升领域内幻觉风险管控能力。

4.3. 关注跟踪生成式人工智能的立法动态和司法实践

法律作为社会关系的调整工具，其制定与修订具有天然的滞后性，难以完全覆盖新技术催生的新型社会关系。生成式人工智能技术作为新质生产力的一种表现形态，在重塑服务模式的同时，也带来了权利界定、责任划分等新型法律问题，现行法律法规体系尚未形成对其应用场景的全面规制。随着技术应用的不断深入、扩展，行业对人工智能专项立法及现有法律修订的需求将持续增强，立法机关也将结合技术发展趋势与社会治理需求，推进相关法律制度的完善。在此背景下，高校图书馆需建立常态化立法跟踪机制，指定专人定期梳理法律法规及规章制度，重点关注生成内容版权归属、用户信息保护等与图书馆业务直接相关的条款，确保图书馆的应用实践始终与法律规定保持一致。

司法实践是法律适用的具体体现，随着最高人民法院《关于统一法律适用加强类案检索的指导意见

(试行)》的发布，各级人民法院在审理案件时愈发注重类案的参考价值。与生成式人工智能有关的司法案例中所体现的法律适用标准、权利义务界定等裁判规则，能为图书馆规范应用生成式人工智能提供更为具体的指引。因此，高校图书馆需要对相关的司法实践保持高度关注，应建立专门的案例收集机制，定期从中国裁判文书网、北大法宝等权威平台检索涉及生成式人工智能版权纠纷、用户信息保护、服务合同争议等与图书馆服务相关的案例，必要时可邀请法学专家共同分析裁判要旨，提炼总结对图书馆管理工作具有指导意义的要点。

4.4. 做好权利义务分配

为化解生成式人工智能应用中的权利归属争议与责任纠纷，高校图书馆需联合技术供应商、用户共同构建权责清晰、利益平衡的三方权利协调框架，通过契约约定与规则制定明确各方权利义务。

(1) 技术供应商层面：拟制专项权利义务条款。一是明确权利归属。在协议中明确约定图书馆提供的本地馆藏数据、自建特色数据库等资源所衍生的生成内容权益，归图书馆(高校)及用户共同所有；技术供应商仅基于算法开发与服务支持获得约定的技术服务费，不得主张对生成内容的所有权或使用权。二是界定责任边界。在售后服务条款中明确供应商的义务，包括但不限于：订购期内提供 7 × 24 小时故障响应服务，确保系统故障修复时长不超过 12 小时；根据技术迭代情况，提供免费版本升级服务，且升级内容需提前与图书馆沟通确认；因技术缺陷导致生成内容侵权或用户信息泄露的，供应商需承担全部赔偿责任及法律后果。

(2) 用户方面：建立多方参与的权利规则制定机制。组建由图书馆负责人、法律专家、信息技术专家、师生代表组成的专项工作委员会，结合《民法典》《著作权法》等相关法律规定，研究确定图书馆生成式人工智能服务条款，明确：生成内容权利义务归属、学术诚信和非商业使用条款、图书馆在合理范围内的使用权等。

4.5. 加强用户信息保护

高校图书馆需从“技术选型、协议约束、流程规范、安全管控”四个维度构建用户信息保护体系，防范信息被不合理使用或泄露。

(1) 在技术选型上，优先采用 DeepSeek 等支持本地化部署的生成式人工智能产品，通过数据存储与计算过程的本地化，减少用户信息向第三方传输的风险；如需使用云端服务，需对供应商的隐私保护能力进行严格评估和约束，确保其符合国家信息安全有关标准。

(2) 在协议约束上，与供应商签订专项《用户信息保护补充协议》，明确核心条款：用户信息的使用范围仅限于图书馆服务场景，不得用于模型训练、商业推广等其他用途；供应商需对收集的用户信息进行匿名化处理，且存储期限不得超过服务合同有效期；供应商委托第三方处理用户信息时，需事先征得图书馆书面同意，并要求第三方承担同等保密责任；协议还需同时约定违约责任，供应商如违反信息保护义务，需支付违约金并赔偿图书馆(学校)及用户的全部损失。

(3) 在流程规范上，优化用户信息授权环节。在服务注册或功能开通页面，将 “是否允许个人信息用于模型训练”“是否同意信息跨场景使用” 等授权项设置为独立勾选框，采用红色字体标注，避免嵌套在“服务条款”等冗长文本中；提供授权变更通道，用户可随时在个人中心调整授权选择，且系统需记录每一次授权变更的时间与内容，形成可追溯的授权日志。

(4) 在安全管控上，建立用户信息分级管理制度。将用户信息分为“普通信息”与“敏感信息”，对敏感信息采用加密算法存储，并设置访问权限白名单，仅授权图书馆信息安全专员、指定服务馆员等必要岗位人员访问；明确安全管控责任人，定期开展信息安全排查。

4.6. 开展分类专项培训

为提升师生与馆员对生成式人工智能的应用能力与风险防控意识，高校图书馆可构建面向不同群体的专项培训体系。

(1) 面向学生群体的阶梯式培训。根据本科生与研究生的不同学术需求，设计差异化内容。针对本科生，开设“生成式人工智能工具入门与学术规范”系列培训，采用“理论讲解 + 案例演示”模式讲解主流 AI 工具的基础功能、操作流程、学术场景适用边界及学术诚信体系对人工智能生成内容的引用要求。针对研究生，开展“AI 生成内容验证与科研效率提升”进阶培训，重点讲解 AI 幻觉的典型特征及辨别方法，指导学生如何使用 AI 工具辅助完成文献综述撰写等科研工作。

(2) 面向馆员群体的专业化培训。按岗位职能分为技术馆员与服务馆员两类培训。针对技术馆员，开展生成式人工智能系统运维相关专项培训，内容涵盖本地数据投喂的标准化流程、模型参数调整技巧、常见系统故障排查方法等，还可邀请供应商技术专家进行实操指导，组织技术馆员参与模拟故障处置演练，确保其能独立完成系统日常维护与突发问题解决。针对服务馆员，开展 AI 服务咨询相关专项培训，包括生成式人工智能技术原理、适配服务场景、高风险场景应对等内容。

5. 结语

生成式人工智能为高校图书馆提升服务效能、满足师生多元信息需求提供了全新路径。然而，图书馆在应用过程中亦面临着一系列风险，这些风险不仅可能影响图书馆服务的质量，还可能给师生的学术科研活动带来潜在困扰，甚至引发法律与伦理层面的问题。高校图书馆需积极采取应对策略，在充分发挥生成式人工智能技术优势的同时，有效规避潜在风险。

高校图书馆与人工智能的交叉研究只有进行时没有完成时，随着技术的迭代升级以及应用场景的不断深化，不可避免地会有更多风险挑战被识别挖掘出来，这也是本文写作的局限。同时，这亦是对未来学界研究的新期待，期待学界深化对图书馆与人工智能领域的交叉研究，为图书馆应用生成式人工智能提供更坚实的理论支撑，共同促进生成式人工智能在高校图书馆领域的规范应用。

基金项目

本文系中央高校基本科研业务费专项“高校图书馆应用生成式人工智能的风险与对策研究”(项目编号：202553004)、CALIS 全国农学文献信息中心研究项目：DeepSeek 赋能高校图书馆信息服务提升研究(项目编号：2025084)、中央高校基本科研业务费专项“面向 AIGC 时代人才培养的信息素养教育——基于素质模型的多维探索”(项目编号：202553005)阶段性研究成果之一。

参考文献

[1] 吴进, 冯劭华, 咎栋. Chat GPT 与高校图书馆参考咨询服务[J]. 大学图书情报学刊, 2023, 41(5): 25-29.

[2] 卢燕群. 人工智能赋能图书馆智慧服务——基于多源融合数据的智能问答实践[J/OL]. 图书馆论坛, 2025. https://link.cnki.net/urlid/44.1306.g2.20250812.0913.002, 2025-08-12.

[3] 刘丹. AIGC 嵌入高校图书馆未来学习中心的价值、风险与防控策略[J]. 传播与版权, 2025(7): 64-67.

[4] 胡安琪, 吉顺权. AIGC 嵌入图书馆知识服务的价值、风险及其防控策略[J]. 图书馆工作与研究, 2024(5): 63-70.

[5] 赵宇翔, 付振康, 曾鹏翔, 等. 生成式人工智能驱动下智慧图书馆信息搜索的技术框架及服务模式研究[J]. 中国图书馆学报, 2025, 51(5): 54-66.

[6] 朱玉萍, 张伟. AIGC 赋能图书馆阅读障碍服务研究[J]. 图书馆学刊, 2025, 47(2): 15-19.

[7] 张炎坤. DeepSeek 赋能图书馆数字阅读推广: 机遇、挑战与对策[J]. 山东图书馆学刊, 2025(3): 1-8+53.

[8] 王军飞. ChatGPT 在图书馆新媒体服务中的应用与思考[J]. 新世纪图书馆, 2024(1): 69-74.

[9] 田欣雨. 图书馆利用生成式人工智能进行阅读推广的版权风险[J]. 图书馆界, 2024(4): 75-80.

[10] 闫宇晨. ChatGPT 应用背景下图书馆科研数据服务版权风险研究[J]. 国家图书馆学刊, 2024, 33(3): 25-36.

[11] 覃振明. 生成式人工智能在图书馆参考咨询服务中的应用与伦理风险探究[J]. 江苏科技信息, 2025, 42(12): 132-136.

[12] 马雨珊. 论图书馆应用生成式 AI 数据风险的法律规制[D]: [硕士学位论文]. 衡阳: 南华大学, 2024.

[13] 翟羽佳. 大语言模型应用于图书馆建设未来学习中心的适应性、风险与策略[J]. 图书馆学研究, 2024(7): 77-85.

[14] 张军雄. Chat GPT-4 对图书馆网络信息安全的挑战与对策[J]. 图书馆学刊, 2024, 46(9): 85-88.

[15] 辛海滨, 王晓倩, 赵璨. 生成式人工智能环境下的图书馆员: 机遇、风险与对策[J]. 图书馆学刊, 2024, 46(10): 36-40.

[16] 王静, 王鹏. 智慧图书馆生成式 AI 大模型风险治理机制研究[J]. 情报杂志, 2024, 43(8): 190-197.

[17] 吴进, 王立杰, 咎栋. DeepSeek 赋能高校图书馆信息服务提升研究——以中国海洋大学图书馆为例[J]. 高校图书馆工作, 2025, 45(2): 22-29.

[18] 中国海洋大学图书馆. 同学们, DeepSeek 为你推荐书单啦! [EB/OL]. 2025-03-05. https://mp.weixin.qq.com/s/8e7inMkKMCrc344NtVfP5A, 2025-08-24.

[19] 童雨萱, 完颜邓邓. 我国高校图书馆古籍资源开放服务现状调查与分析[J]. 图书馆, 2022(11): 78-84.

[20] 杨奇光, 韩佳序. AI 幻觉对新闻真实性的冲击及其风险防范研究[J/OL]. 新媒体与社会, 2025. https://link.cnki.net/urlid/CN.20250807.1624.002, 2025-08-07.

[21] Xu, Z., Jain, S. and Kankanhalli, M. (2025) Hallucination Is Inevitable: An Innate Limitation of Large Language Models. https://arxiv.org/abs/2401.11817v2

[22] 南方都市报. “DeepSeek 给王一博道歉”声明热传!多方核查: 是假的[EB/OL]. 2025-07-04. https://mp.weixin.qq.com/s/_2cNLnFWJASv57t6BJ8iKw?scene=1&click_id=7&poc_token=HITQqmijjgX3aj4tHo4ise-iWlSSG1Fyh2ueNDT4, 2025-08-24.

[23] 姚立. 生成式人工智能 ChatGPT 编造法律案例的风险与启示——评 2023 年美国“马塔诉阿维安卡公司案”[C]//《新兴权利》集刊 2024 年第 2 卷——人工智能背景下的新兴权利研究文集. 2024: 192-202.

[24] 北京互联网法院判决计算机软件智能生成文字内容不构成作品[EB/OL]. 2019-05-29. http://legal.people.com.cn/n1/2019/0529/c42510-31108379.html, 2025-08-24.

[25] 中国法院网. 腾讯诉盈讯科技侵害著作权纠纷案[EB/OL]. https://www.chinacourt.cn/article/detail/2021/01/id/5709690.shtml, 2025-08-24.

[26] 中国法院网. 破冰: 首例人工智能文生图案生效[EB/OL]. https://www.chinacourt.cn/article/detail/2024/02/id/7796864.shtml, 2025-08-24.

[27] OpenAI (2025) How Your Data Is Used to Improve Model Performance. https://help.openai.com/en/articles/5722486-how-your-data-is-used-to-improve-model-performance

[28] OpenAI (2025) Data Controls FAQ. https://help.openai.com/en/articles/7730893-data-controls-faq