情报杂志
Journal of Intelligence
ISSN 1002-1965,CN 61-1167/G3

《情报杂志》网络首发论文

题目：DeepSeek 驱动图书馆智慧化转型:理论逻辑、安全风险与优化路径
作者：张焕培
网络首发日期：2025-05-15
引用格式：张焕培．DeepSeek 驱动图书馆智慧化转型:理论逻辑、安全风险与优化路径[J/OL]．情报杂志. https://link.cnki.net/urlid/61.1167.G3.20250514.1704.014

网络首发时间：2025-05-15 14:50:13　网络首发地址：https://link.cnki.net/urlid/61.1167.G3.20250514.1704.014

DeepSeek 驱动图书馆智慧化转型：理论逻辑、安全风险与优化路径

张 焕 培

（广东海洋大学法政学院　湛江　524088）

摘　要：[**研究目的**]现有研究多聚焦 ChatGPT 在智慧图书馆中的传统风险，对 DeepSeek 技术特征衍生的新型安全风险关注不足。本文旨在揭示 DeepSeek 驱动图书馆智慧化转型的理论逻辑，剖析其推理可视化、模态融合高效化与场景垂直化引发的商业秘密泄露、数字古籍复合授权争议及责任穿透隐患，并提出分层治理路径，以平衡技术红利与风险管控。[**研究方法**]采用跨学科理论分析，结合技术特征解构、法律规范解释与案例实证，构建“技术-风险-规制”框架。对比 ChatGPT 与 DeepSeek 差异，梳理转型逻辑；基于《反不正当竞争法》《著作权法》及国内外典型案例，分析法律冲突；借助延伸性集体许可制度等理论工具，设计分层策略。[**研究结果/结论**]针对商业秘密泄露，构建动态保密机制与反向工程合法性边界规范，通过技术保密与用户分级管理化解冲突；针对数字古籍授权困境，引入延伸性集体许可制度，区分基础扫描、技术修复与艺术重构场景，降低授权成本；针对责任穿透风险，建立“认知干预+管理强化”双向治理体系，通过动态风险提示、责任分割及第三方认证实现权责平衡。

关键词：DeepSeek；生成式人工智能；智慧图书馆；智慧化转型；风险规制；ChatGPT

中图分类号：G250.76　　**文献标识码：**A

DeepSeek-Driven Intelligent Transformation of Libraries: Theoretical Logic, Security Risks, and Optimization Paths

Zhang Huanpei

(Guangdong Ocean University, Zhanjiang　524088)

Abstract: [**Research purpose**] Existing studies mostly focus on the traditional risks of ChatGPT in smart libraries, and there is insufficient attention to the new security risks derived from the technical features of DeepSeek. This paper aims to reveal the theoretical logic of DeepSeek driving the intelligent transformation of libraries, analyze the leakage of business secrets, disputes over composite authorization of digital ancient books, and potential hazards of liability penetration caused by its reasoning visualization, efficient modal fusion, and verticalization of scenarios, and propose a hierarchical governance path to balance the technological dividends and risk management and control. [**Research method**] An interdisciplinary theoretical analysis is adopted, combined with the deconstruction of technical features, the interpretation of legal norms, and case demonstrations, to construct a "technology-risk-regulation" framework. The differences between ChatGPT and DeepSeek are compared to sort out the transformation logic; based on the Anti-Unfair Competition Law, Copyright Law, and typical domestic and foreign cases, the legal conflicts are analyzed; with the help of theoretical tools such as the extended collective licensing system, hierarchical strategies are designed. [**Research result/conclusion**] In response to the leakage of business secrets, a dynamic confidentiality mechanism and norms for the legitimacy boundary of reverse engineering are constructed, and the conflicts are resolved through technical confidentiality and hierarchical user management; in response to the dilemma of authorization of digital ancient books, the extended collective licensing system is introduced, and scenarios such as basic scanning, technical restoration, and artistic reconstruction are distinguished to reduce the authorization cost; in response to the risk of liability penetration, a two-way governance system of "cognitive intervention + management reinforcement" is established, and the balance of rights and responsibilities is achieved through dynamic risk prompts, liability segmentation, and third-party certification.

Key words: DeepSeek; Generative Artificial Intelligence ; Smart Libraries; Intelligent Transformation; Risk Regulation; ChatGPT

作者简介：**张焕培**，男，2001 年生，硕士研究生，研究方向：数字法学。

0 引 言

近年来,生成式人工智能(以下简称 GAI)的兴起,为图书馆的智慧化转型带来了契机[1]。特别是 2025 年 1 月深度求索公司推出的推理模型 DeepSeek-R1,在多项测试中超越了 OpenAI 公司旗下的 ChatGPT 产品,成为 GAI 领域的全球瞩目事件[2]。DeepSeek-R1 凭借低至顶尖模型 1/10 的训练成本,实现了与之相媲美的性能,使得 GAI 的低成本高效率训练成为现实,并在推理效率、场景适配等关键维度树立了全新的行业标杆[3]。党的二十届三中全会提出,要评估“十四五”落实情况,重视“十五五”布局工作[4]。在“十四五”期间,智慧图书馆建设虽已取得一定进展,但 GAI 在该领域的应用仍较为有限[5]。而 DeepSeek[6] 的兴起为“十五五”布局带来了时代机遇,从 ChatGPT 到 DeepSeek 的技术迭代,实质上是 GAI 从黑箱输出到白箱推理、从基础跨模态能力到跨模态任务优化、从通用对话到提供行业适配架构的重要转变[7]。基于此种转变,DeepSeek 将助力图书馆在知识服务优化、数字内容创造,以及用户服务黏性提升等方面的突破,从而实现智慧化转型。

但在此过程中,安全风险将逐渐凸显。推理可视化可能引发商业秘密的动态泄露,模态融合高效化加剧数字古籍的复合授权争议,场景垂直化则潜藏专精服务的责任穿透隐患。值得注意的是,2025 年 3 月 7 日,国家互联网信息办公室、工业和信息化部等四部门制定《人工智能生成合成内容标识办法》(以下简称《标识办法》),其中明确提出,GAI 服务提供者需对 AI 生成内容进行显著标识,并确保标识不可篡改[8]。这一政策为图书馆应用 DeepSeek 提供了合规框架。然而,目前学界更多是基于 ChatGPT 的视角,研究智慧图书馆深化 GAI 应用过程中的安全风险,集中在数据隐私风险[9]、知识产权风险[10]、算法滥用风险[11]等传统议题,而对 DeepSeek 技术特征所衍生的新型风险关注不足。在此背景下,如何结合《标识办法》的合规要求,平衡 DeepSeek 的技术红利与风险管控,成为图书馆实现智慧化转型的关键议题。有鉴于此,本文以“推理可视化、模态融合高效化、场景垂直化”三大技术特征为切入点,系统分析 DeepSeek 驱动图书馆智慧化转型的理论逻辑与安全风险,进而提出适配公共文化服务属性的分层优化路径。

1 DeepSeek 驱动图书馆智慧化转型的理论逻辑

ChatGPT 以其流畅的通用对话能力被广泛认知,但在知识服务的可解释性、多模态协同处理,以及垂直领域适配性上存在局限。与之相比,DeepSeek 通过推理可视化实现了从技术黑箱到透明化决策的跃迁,通过跨模态任务优化提升了文图交互的生成效率与精准度,并通过提供行业适配架构支持垂直领域的知识深度挖掘(见表 1)。这些技术特性为图书馆智慧化转型提供了时代契机。

表 1 DeepSeek 区别于 ChatGPT 的技术特征

对比维度	ChatGPT	DeepSeek	技术特征
运行过程	黑箱输出	白箱推理	推理可视化
模态整合	基础跨模态能力	跨模态任务优化	模态融合高效化
应用服务	通用对话	提供行业适配架构	场景垂直化

1.1 推理可视化赋能图书馆的可信知识服务

图书馆的智慧化转型可概括为三个阶段[12]:在资源数字化阶段,通过数据库建设提升知识获取效率,解决资源可及性不足的问题。在服务智能化阶段,开始引入 GAI 技术,矛盾点转化为信息过载与精准知识服务需求之间的冲突,ChatGPT 的黑箱化的决策逻辑导致用户出现信任危机[13]。在认知增强阶段,致力于知识生产过程的透明化,旨在引导用户完成从数据到智慧的认知跃迁[14]。由此可见,如何实现可信知识服务是实现智慧图书馆转型的核心议题,而 GAI 技术的透明程度直接影响知识服务效能[15]。

作为“黑箱”GAI,ChatGPT 的知识表征建立在深度神经网络的参数之中,其知识建构过程发轫于数据驱动,呈现出隐式关联,依赖于概率生成。此种“黑箱”范式,虽能捕捉语义的深层关联,却无法解决知识溯因的不可解释性[16]。反观作为“白箱”GAI 的 DeepSeek,凭借长链推理(CoT)[17],实现了符号主义与联结主义的深度融合。其知识生产过程遵循“逻辑推演-显式关联-路径可溯”的理路,在保持深度学习表征优势的同时,引入符号逻辑的可解释性特征。而借鉴知识管理 SECI 模型[18],智慧图书馆的服务体系可分解为四个维度[19]:社会化强调用户与系统的交互信任建立,外显化着重隐性知识到显性知识的有效转化,组合化反映知识元素的逻辑重组能力,内隐化聚焦知识吸收的认知适配性。传统“黑箱”GAI 在组合化阶段存在缺陷,其知识重组过程缺乏透明路径,导致用户难以完成从数据到智慧的吸收转化。而 DeepSeek 凭借推理过程可视化技术,构建了可信知识服务体系,这可以分为三步走。第一步,廓清每个推理步骤的主张、依据与理由,使知识生产过程符合论证规范,破解传统推荐系统的黑箱化伦理困境。第二步,通过双向交互接口,支持用户质疑推理节点、回溯修正路径,将单向服务转化为协同知识建构。第三步,借助可视化推理链的线性呈现、冲突标注与元认知提示,引导用户

在自身知识巩固与新观点接纳的动态平衡之间,完成认知跃迁。

1.2 模态融合高效化支撑图书馆的数字内容创造

区别于传统图书馆,智慧图书馆不仅需要承担保存文化资源的社会功能,同时需要强化作为数字内容创造者的角色认知[20]。智慧图书馆的数字内容创造过程包括两个环节。第一个环节是资源数字化,从纸质资源扫描转向语义化知识单元重组;第二个环节是服务场景扩展,依托数字内容构建虚实融合的交互空间[21]。针对数字内容创造的第一个环节,DeepSeek 通过跨模态任务优化优势展现出独特价值。其系列多模态生成模型 Janus-Pro 支持文本到图像、图像到文本等跨模态任务。相较于 ChatGPT 依赖 DALL-E 等独立模型实现文图转换,Janus-Pro 在模态对齐效率与生成精准度上更具优势。例如,在文生图基准测试中,其准确率分别达到 80% 和 84.2%[22]。因此,此种模型可为智慧图书馆构建跨模态特征提取通道[23]。例如,在文本分支解析古籍异体字与语义逻辑,在图像分支捕捉书画作品的笔触特征与空间关系,有效实现异质文化资源的深度解析。

针对数字内容创造的第二个环节,在当下虚拟现实(VR)、增强现实(AR)等前沿技术与图书馆深度结合的浪潮中,智慧图书馆的沉浸式服务已初具规模[24]。从目前整体发展情况来看,图书馆普遍采用元宇宙技术搭建虚拟空间平台,但其中面临 3D 建模成本高、交互体验受限等瓶颈[25]。2025 年 2 月 6 日,北京大学联合香港科技大学推出了多模态版 DeepSeek-R1——Align-DS-V,通过技术创新开辟突破路径。此种模型有望将 2D 文献转化为 3D 可交互场景,助力图书馆用户多角度观察历史文物的艺术细节,同时实时对齐用户手势操作与文献内容,实现当用户凝视历史文物插图时,系统可同步触发相关动画解说[26]。目前,美国医学图书馆、中国科学院文献情报中心等机构,已将多模态应用纳入关键战略规划[15],印证了该技术路径在智慧图书馆建设中的实践价值。

1.3 场景垂直化提升图书馆的深层用户黏性

情感依恋理论揭示的用户黏性具有双重结构:表层黏性源于交互体验,深层黏性则依赖需求满足程度[27]。基于人类反馈强化学习(RLHF)的微调策略[28],ChatGPT 在通用对话中表现出色,能够保障生成文本的连贯性和流畅性,但在处理特定领域的问题时尚显乏力。因为 RLHF 依赖的反馈数据多为通用场景标注,导致专业术语使用不精准或知识更新滞后。同时,奖励模型难以量化专业逻辑严谨性,更多以流畅表达为标准来判别回答质量。2025 年 2 月 3 日,美国加利福尼亚州立大学正式引入 ChatGPT Edu,创建适用于图书馆等场景的 AI 聊天机器人[29]。此种做法通过流畅的交互体验构成表层黏性,但专业用户在检索过程中会经历认知不确定期,此时仅靠语言连贯性无法满足其深层需求[30]。

DeepSeek 通过引入混合专家模型(MoE)[31],将通用模型划分为多个专家模块,为行业版本开发提供灵活的技术架构。其通用模型通过门控网络动态分配任务至不同专家模块,支持针对法律、医疗等垂直领域的知识定制化训练。在法律领域,WhyHow. AI 团队基于 DeepSeek-R1 通用模型开发的医疗法律行业版本,通过定向训练医疗法律语料与专家模块协同,实现了患者记录与法律事件的高效关联分析。在金融领域,博时基金基于 DeepSeek-R1 行业版本完成本地化部署,通过金融数据定向优化与专家权重调整,显著提升了投资建议的专业性与精准度。由此可见,DeepSeek 的 MoE 架构为图书馆场景垂直化服务提供了底层技术支持,但实际专精功能需结合行业版本的定制开发与部署。因此,若能推进 DeepSeek 行业版本在图书馆的本地化适配,将有助于深化专业用户对图书馆服务的依赖,从而增强深层用户黏性,加速智慧化转型进程。

2 DeepSeek 驱动图书馆智慧化转型的安全风险

以 DeepSeek 为代表的新型 GAI,因推理过程可解释性增强、多模态生成能力拓展,以及垂直领域适应性深化,有望突破 ChatGPT 在图书馆应用的可信性瓶颈与服务边界。随着《生成式人工智能服务暂行管理办法》(以下简称《暂行办法》)的全面实施,图书馆在深化 GAI 应用的过程中,既需积极响应推动人工智能与公共服务深度融合的政策导向,又须落实健全技术应用风险防控机制的合规要求[32]。在此背景下,有必要结合 DeepSeek 的技术特征,探讨智慧图书馆转型中技术赋能与制度约束间的张力关系。

2.1 推理可视化之维:商业秘密的动态泄露难题

DeepSeek 的推理可视化技术,虽有利于提升图书馆知识服务的透明性与可信度,但也存在现实隐患。倘若模型在处理用户查询时涉及商业数据库、专利文献或企业定制化知识库,其白箱化的推理过程可能会泄露商业秘密。根据《中华人民共和国反不正当竞争法》(以下简称《反不正当竞争法》)第 9 条规定,法律定义上的商业秘密需满足“秘密性”“保密性”“价值性”三要件[33]。然而,由白箱推理引发的商业秘密动态泄露,可能与前两个要件相冲突。

第一重困境,推理可视化与商业秘密之“保密性”的冲突。一方面,企业通常采取代码加密、访问权限控制等措施保护商业秘密,但这些措施仅针对静态存储

的固化信息。而 DeepSeek 在实时交互中动态生成的逻辑链,难以被预先定义为保密对象,因而脱离传统保密措施的覆盖范围。同时,依据《最高人民法院关于审理侵犯商业秘密民事案件适用法律若干问题的规定》第 6 条规定,“合理保密措施”须具备有效性,但动态推理过程使保密措施无法追溯至实时生成的逻辑链,导致司法认定标准失效;另一方面,在智慧图书馆场景下,专业用户可能通过分析 DeepSeek 输出的推理路径,逆向还原企业的核心算法。此行为虽符合《中华人民共和国民法典》(以下简称《民法典》)第 123 条“侵犯技术秘密”的客观要件,但对比美国《商业秘密保护法》(Defend Trade Secrets Act,以下简称“DTSA”)第 1839(6)(B)条将“反向工程”排除于侵权范畴之外的立法模式,我国对此尚缺乏明确的禁止性规定,司法实践中易引发争议。此种争议之处,在 2024 年德国汉堡地区法院审理的 LAION 数据集案件中可以窥见[34]。德国法院虽驳回原告对数据集创建的侵权诉求,但未涉及后续反向工程行为的合法性,暴露了法律对动态推理环节规制的模糊性。

第二重困境,推理可视化与商业秘密之“秘密性”的冲突。商业秘密的“不为公众所知悉”,要求信息具有整体上的非公开性[35],而 DeepSeek 的交互式服务可能通过不同路径破坏“秘密性”。用户可通过多轮对话逐步获取技术信息的核心参数,即使单次回答未完整披露,但多次合法提问的组合可能还原商业秘密的全貌。同时,DeepSeek 基于训练数据生成的推理链条可能隐含企业未公开的决策逻辑,通过追问“为什么选择 A 方案而非 B 方案”,专业用户可逆向推导出企业保密的核心算法。此外,DeepSeek 的知识关联能力可能将零散公开信息与企业保密信息建立隐性联系。譬如,通过分析多个公开工艺步骤推测出完整的保密配方。诚然,上述泄露途径受制于《反不正当竞争法》第 9 条“以其他不正当手段获取商业秘密”的规定,但现行法缺乏对此种分步式侵权的认定标准。而且,单一合法行为与整体违法结果间的因果关系难以证成。商业秘密侵权需证明“接触+实质性相似”[36],由于用户提问未直接指向保密信息,接触证明存在困难。同时,DeepSeek 输出的是重组后的新表述,与原始保密信息是实质相同却形式不同的关系,从而使得相似性判断受阻。

2.2 模态融合高效化之维:数字古籍的复合授权挑战

DeepSeek 依托于模态融合高效化特征,促进数字内容创造,从而赋能智慧图书馆建设。然而,这一过程可能会涉及 AI 数字古籍的转化环节,这将同时触发三类人的权利主张。DeepSeek 通过提取影印本中的纹样、构图等美术元素生成模型内容。若生成成果与原作在“核心表达”上构成实质性相似,则可能侵犯原书版权人的改编权[37]。同时,数据库平台对古籍数字化版本可能主张摄影作品权。DeepSeek 生成过程中若复制光影细节、色彩校正结果等具有独创性的数字化特征,可能构成对复制权的侵犯。但此类主张需以数字化版本被法院认定为摄影作品为前提。此外,学术机构发布的残片定位数据、修复结论等成果,若经 DeepSeek 聚合后形成具有独创性的表达,如通过算法优化残片拼接逻辑,可能构成汇编作品[38]。根据《中华人民共和国著作权法》(以下简称《著作权法》)第 16 条规定,未经许可使用此类数据组合将侵犯学术机构的汇编权。

由此导致的后果是,现有授权机制将会受到冲击,致使图书馆陷入两维成本困局。就横向谈判成本而言,同一古籍因涉及原作者继承人、数字化平台、学术机构等多方权利人,需同步处理改编权、复制权、汇编权等异质权利主张。譬如,某古籍的原作者继承人可能要求高额的改编权费用,而数字化平台可能对复制权有严格的限制,学术机构则可能要求对汇编权进行特别约定。而且,生成物可能同时包含受版权保护元素与公共领域内容,会涉及“孤儿作品”难以授权的问题。在此情形下,传统“点对点”授权模式将产生高额合规成本。就纵向审查成本而言,图书馆除了需要逐一核查各数据库平台的禁止性条款之外,还需要系统性比对不同权利人的义务性条款,更要持续监控多源数据的授权变更状态。此类审查成本在跨机构协作场景中过于高昂,严重制约古籍数字化效率。除此之外,现行法缺乏对 AI 数字古籍的独创性认定标准,此种权属模糊性也将进一步激化复合授权挑战。

2.3 场景垂直化之维:专精服务的责任穿透隐患

DeepSeek 因具备应用场景垂直化特征,成为图书馆提升用户深层黏性的关键底牌。但其中涉及的专精式对话服务,也成为法律责任穿透风险的重要诱因。GAI 在处理专业知识时,可能因运作机理的局限性而产生机器幻觉[39]。比如,2023 年美国律师史蒂文·施瓦茨因轻信 GAI 的专业功能,在法庭文件中引用了并不存在的法律案例。这背后的根源在于,GAI 更多是基于概率分布而非语义逻辑,去组织输出内容[40]。即便相较于 ChatGPT 而言,DeepSeek 具有一定的技术突破。但在 GAI 底层原理未被改变的情况下,仍然无法避免图书馆用户的权益因 AI 幻觉而受到损害,遑论图书馆的角色定位还可能放大用户对 DeepSeek 专精服务的信任。出于对公共服务主体的认知惯性,部分用户可能会默认图书馆提供的法律或金融信息具备权威性。如果仅是简单地标注“AI 生成”,这些用户仍倾向

于视其为专业意见。同时,上述信任强化所带来的不利后果还具有长尾效应,其往往不是在图书馆场景下即时产生的,而是在特定的社会环境中才逐渐显现,如法律建议的错误将导致用户败诉,金融分析的失误可能引发投资亏损,而图书馆难以弥补此类不可逆的损害。

正因如此,在 DeepSeek 赋能智慧图书馆转型的过程中,责任穿透隐患将被激化。作为开发者的深度求索公司,主要负责提供基础模型与 API 接口,并通过协议手段将法律责任转移至图书馆。这从《DeepSeek 开放平台服务协议》3.2 与 7.3 条款中可以窥见,深度求索公司强调其开放平台提供的是中立、基础的模型技术服务[41],该技术服务仅为价值链下游的系统、应用或功能的一部分,无法决定服务的最终目的和用途。下游系统、应用或功能的提供者、运营者,应对该系统、应用或功能负责,并承担相应的法律责任。这意味着,图书馆作为模型部署方,由于将 AI 服务嵌入自身业务系统,直接面向用户提供服务,成为实际责任主体。与之相对的,用户与图书馆而非开发者建立服务合同关系,其权益受损时将直接追究图书馆责任。由此导致的后果是,图书馆在专业领域服务中的注意义务显著加重,对图书馆的管理效率与能力提出了严苛要求。

3 DeepSeek 驱动图书馆智慧化转型的优化路径

DeepSeek 通过推理可视化、多模态融合与垂直场景应用构建的技术优势,在提升知识服务精准度的同时,亦催生出了图书馆智慧化转型的三重规制命题:商业秘密保护的保密性与秘密性要件在透明化推理场景中的冲突消解、数字古籍生成引发的复合版权授权困境,以及专精服务属性导致的责任穿透风险。这些挑战映射出 AI 技术迭代与制度供给之间的动态博弈,要求探索以风险分层识别为基础、技术法律协同为手段、多方利益平衡为导向的优化路径。

3.1 商业秘密泄露层面:形塑以保密性与秘密性要件为锚点的控制机制

2023 年三星电子公司允许其员工使用 ChatGPT,以提高工作效率。然而,在此之后连续发生了三起数据泄露事件,导致该公司的相关设备信息以及会议内容流出[42]。这表明,GAI 应用本身就存在商业秘密泄露的可能[43]。相较于 ChatGPT,DeepSeek 衍生出了推理可视化特征。这将进一步激化其在智慧图书馆应用中的商业秘密泄漏风险,与商业秘密之“保密性”“秘密性”要件形成冲突。为避免上述风险泛滥,应形塑针对性的控制机制。

3.1.1 保密性要件冲突之控制

一是保密措施有效性认定标准的动态衔接。一方面,可以依据《反不正当竞争法》第 9 条的“保密措施”要件,在司法解释中引入“技术过程保密”理念。最高人民法院在相关司法解释中已经明确了保密措施的具体要求[44],如限定涉密信息的知悉范围、对涉密信息载体采取加锁等防范措施、在涉密信息的载体上标有保密标志等,这些措施可以为“技术过程保密”提供具体的法律依据;另一方面,明确以预见性防护与实时性阻断为核心的动态保密认定义务。对于预见性防护而言,要求模型开发者对可能涉及商业秘密的推理路径进行风险预标注,如采用 NLP 语义标记技术对敏感决策节点进行加密。对于实时性阻断而言,当用户查询触发预设风险阈值时,模型开发者内置的系统应自动启用逻辑链截断机制。除此之外,在建立可验证的动态保密技术清单的前提下,应允许模型开发者在个案中举证其采取的技术防护措施满足商业秘密保护的最低标准。

二是反向工程合法性边界的规范衔接。通过司法解释对《民法典》第 123 条“不正当手段”进行类型化解释,将利用 GAI 推理路径实施的算法还原行为纳入“技术秘密侵害”范畴,并参考美国《统一商业秘密法》(Uniform Trade Secret Act)第 1 条注释的实质性步骤标准,确立“三步认定法”——判断反向工程对象是否涉及保密算法片段,并分析技术手段的专业复杂程度,最后评估还原结果与源技术的相似度。同时,应考虑到主体行为的差异,对图书馆场景中专业用户的常规分析行为与恶意商业破解行为进行区分。若企业已对 AI 推理路径采取逻辑链片段化、非连续化处理等技术隔离措施,则用户通过常规分析手段获取片段信息的行为不视为侵权。

3.1.2 秘密性要件冲突之控制

就风险预防而言,应要求图书馆考量用户类型、管理要求、访问权限三要素,构建 AI 交互权限的用户分级管理制度,以避免出现法律纠纷。针对进行实名认证的普通公众,允许其进行公共数据库检索,输出结论层摘要。针对通过机构认证与研究计划备案的学术研究人员,可以在隐藏核心参数的前提下,为其开放逻辑层框架。针对签署保密协议与通过动态令牌认证的企业合作方,图书馆可以为其提供定制化知识库访问。

就风险控制而言,应通过合理的法律调适,应对分步式侵权的规制失灵。在 DeepSeek 赋能智慧图书馆场景中,分步式侵权表现为,用户通过多次独立查询,分别获取商业秘密的零散片段,最终通过组合分析复现商业秘密整体。此种侵权行为引发的法律争议有二:一是单一查询行为可能合法,但系统性组合构成侵权。二是因果关系证明困难,需证明分散查询与商业秘密泄露的实质关联。针对第一点,应对《反不正当

竞争法》第9条“其他不正当手段”进行扩张解释。具体可以如下设计,通过多次合法查询系统性获取商业秘密片段,且具备下列情形之一的,应当认定为以不正当手段获取商业秘密:①查询行为具有明确指向性,集中于特定技术或经营领域;②查询结果经组合可实质性复现商业秘密核心内容;③查询目的与用户身份或常规需求明显不符;针对第二点,廓清分步式侵权的司法认定标准。先通过查询日志时间序列分析报告,验证查询内容是否逻辑递进关系,以此判断行为关联性。再参考专家出具的信息组合可行性鉴定书,验证各次查询结果是否可相互印证形成完整信息链,以判断信息互补性。最后对比分析用户历史行为数据,考察查询频次是否偏离正常使用模式,以判断目的非常性。

3.2 复合授权挑战层面:形塑AI数字古籍的类型化版权授权规则

在智慧图书馆场景中,AI数字古籍的技术特性,使得传统“一刀切”的版权授权模式面临实践困境。古籍数字化过程中,生成内容与原作的相似性梯度、技术介入深度,以及成果创造性维度存在显著差异,导致单一授权规则难以适配多元场景。因此,有必要构建类型化的版权授权规则,以满足多源作品复合使用的批量授权需求,并破除技术修复与艺术创作交织的权属认定难题。

第一种情形,AI数字古籍触发改编权、复制权、汇编权复合主张,生成内容与原作高度相似。为解决授权效率与公平性问题,可以借鉴北欧的延伸性集体许可制度(Extended Collective Licenses,以下简称ECL),通过集体管理组织与使用者签订一揽子协议,覆盖多权利人的授权[45]。事实上,我国存在相关的制度渊源。2011年7月13日,我国《著作权法》第三次修订工作正式启动,在修订草案中引入了ECL。2012年3月31日,国家版权局发布的《著作权法》修改草案第一稿中,明确了ECL的相关规定,并在第二稿适当缩小了适用范围。虽然立法者最终在新修订的《著作权法》将审议稿中关于ECL的相关条款删除,但过往的立法考量一定程度上也证明了ECL在我国适用的正当性。因此,为应对AI数字古籍的复合授权挑战,未来可以在修订《中华人民共和国公共图书馆法》时,增设古籍数字化复合授权特别条款。推进ECL的费率动态调整,对高频作品与低频作品进行支付区分。同时,配套异议退出机制,允许权利人申请退出并单独授权,但需举证证明其作品具有特殊市场价值。除此之外,加强统一版权数据库的合规性保障。数据库应公开作品版权的保护期、权利人联系方式,减少“孤儿作品”问题。图书馆可在线提交使用清单,系统自动生成授权协议与费用清单,缩短谈判周期。

第二种情形,AI数字古籍仅调用如单一影印本等少量数据,但生成内容与原作高度相似。为避免该象限下的AI数字古籍独创性难以认定,可以将其生成情况,划分为基础扫描、技术修复、艺术重构三个梯度。其一,基础扫描属于对古籍的机械复制,未添加任何如无构图调整、色彩修正等创造性表达。根据《北京市高级人民法院侵害著作权案件审理指南》第2.5条规定,对已有作品的精确复制,不产生新的著作权[46]。因此,基础扫描梯度下的AI数字古籍不构成作品,可依据《著作权法》第24条合理使用。其二,技术修复指的是,通过技术手段修复古籍破损部分,但未改变原作品表达形式。修复行为仅恢复原作原貌,属于技术性劳动,不产生新权利。因此,技术修复梯度下的AI数字古籍一般不构成新作品,无需额外授权。其三,艺术重构是在古籍数字化过程中加入创造性表达以形成作品的行为。在此梯度下的AI数字古籍,可能构成摄影作品或演绎作品,需取得原权利人的复制权、改编权许可。

3.3 责任穿透隐患层面:构建“认知干预+管理强化”的双向治理体系

2023年2月23日,华尔街投行摩根大通已限制员工使用ChatGPT,体现了金融领域对GAI服务的担忧[47]。而与ChatGPT相比,DeepSeek在法律与金融领域的专精式服务属性相对凸显。这将进一步激化其在智慧图书馆应用中的责任穿透隐患,对图书馆的管理效率与管理能力提出挑战。为破除这一难题,可以从用户认知干预、图书馆管理强化两个层面入手,设计双向应对机制。

从用户认知干预层面来看,可借鉴《中华人民共和国个人信息保护法》中“告知-同意”的分层设计理念,构建动态风险告知机制。针对初次使用时的轻量提示,图书馆可在服务入口设置浮动标签,如“AI生成内容可能存在误差”,字号应与界面元素保持视觉协调,避免遮挡核心内容。针对高危场景的强化告知,当用户触发法律咨询、投资建议等高风险服务时,图书馆应通过模态对话框展示结构化风险清单,嵌套核心风险要点与扩展解释链接。如此一来,既满足强制阅读要求,又保留信息获取便捷性。针对持续性警示的嵌入。在AI生成内容的显示界面固定位置添加水印式提示,并设置错误举报入口。不仅如此,应弱化用户对AI生成内容的权威性感知。避免使用具有专业指向的服务命名,改用“法律信息参考工具”“金融数据解析助手”等中性表述,禁用法槌、天平、货币符号等易引发权威联想的图标,采用中性色彩与扁平化设计,与人工服务形成视觉区分。加强对图书馆用户在AI时代下的信息素养教育,使其合理地理解和使用Deep-

Seek,这在国外已有先例。宾夕法尼亚大学比德尔图书馆提供了 GAI 在法律学习和实践中的案例,普林斯顿大学费尔斯通图书馆开展了关于 GAI 与人文科学的讨论。

从图书馆管理强化层面来看,关键在于两点。一是结合有限审查与第三方认证。依据《暂行办法》第 14 条,图书馆需部署内容过滤与追溯技术。除此之外,还需要对涉及法律结论、投资建议等高危输出,按比例抽取并由合作律所或金融机构进行合规性审查,建立人工复核标签以增强可信度。同时,引入第三方认证接口,与裁判文书网等权威平台对接,实时验证 AI 生成内容的时效性与准确性,并通过 API 返回验证结果标识。在此基础上,针对法律、金融等领域的知识滞后性问题,联合行业协会与专业机构建立动态语料更新联盟。例如,与最高人民法院合作接入最新司法解释数据库,或与证券公司共享合规研报,通过定向数据授权解决训练数据时效性问题,并鼓励专业用户提交修正反馈,并将其纳入模型优化奖励机制,形成闭环迭代;二是构思 AI 专精服务的责任分割。图书馆可将 AI 服务划分为基础信息检索与专业分析建议两类,前者由图书馆承担有限责任,后者需用户签署额外免责协议并跳转至 DeepSeek 官方界面完成,实现责任分流。此外,可参考金融科技领域经验,购买 AI 责任险,将因内容错误导致的用户损失纳入保险赔付范围,降低图书馆直接赔偿责任[48]。

4 结 语

DeepSeek 作为新一代 GAI 技术的代表,凭借推理可视化、模态融合高效化与场景垂直化的技术特征,为图书馆智慧化转型注入了新动能。其凭借透明化推理逻辑、高效跨模态处理与提供行业适配架构,有效提升了知识服务的可信度、数字资源的创造力及用户黏性,为“十五五”时期公共文化服务的智能化跃迁提供了实践路径。然而,技术革新亦伴生新型风险。商业秘密的动态泄露挑战、数字古籍的复合授权困局,以及专精服务的责任穿透隐患,揭示了技术赋能与制度约束之间的张力关系。针对上述风险,本文提出分层治理框架。通过动态保密机制与反向工程边界的规范衔接,平衡商业秘密保护与技术透明化的冲突;依托类型化版权授权规则与延伸性集体许可制度,破解 AI 数字古籍的复合授权难题;构建“认知干预+管理强化”的双向治理体系,实现用户权益保障与图书馆责任分流的协同。未来,随着 GAI 技术的持续迭代,智慧图书馆的建设需进一步关注技术动态性与制度滞后性的矛盾,探索更加灵活的风险预警机制与多方协作治理模式。同时,如何将伦理原则嵌入技术设计、如何在全球化背景下协调版权规则差异,亦是亟待深化的研究方向。唯有在技术创新与制度完善的动态平衡中,图书馆方能真正迈向可信、普惠、可持续的智慧化未来。

参 考 文 献

[1] 柯 平. 面向“十五五”规划的智慧图书馆建设[J/OL]. 图书馆理论与实践,1-12[2025-03-28]. https://doi.org/10.14064/j.cnki.issn1005-8214.20250207.001.

[2] 方兴东,王 奔,钟祥铭. DeepSeek 时刻:技术—传播—社会(TCS)框架与主流化鸿沟的跨越[J/OL]. 新疆师范大学学报(哲学社会科学版),1-11[2025-03-28]. https://doi.org/10.14100/j.cnki.65-1039/g4.20250218.001.

[3] 邓建鹏,赵治松. DeepSeek 的破局与变局:论生成式人工智能的监管方向[J/OL]. 新疆师范大学学报(哲学社会科学版),1-10[2025-03-28]. https://doi.org/10.14100/j.cnki.65-1039/g4.20250214.001.

[4] 中共二十届三中全会在京举行[N]. 人民日报,2024-07-19(001).

[5] 杨新涯,戴立伟,钱国富. DeepSeek 在图书馆的应用场景构架研究[J/OL]. 图书馆论坛,1-8[2025-03-28]. http://kns.cnki.net/kcms/detail/44.1306.g2.20250218.1639.004.html.

[6] 本文所指的 DeepSeek,包括 DeepSeek 通用模型本身,与由其衍生的人工智能大模型。

[7] 郭亚军,徐苑茜,梁艳丽,等. 从 ChatGPT 到 DeepSeek:生成式人工智能迭代对图书馆的影响[J/OL]. 图书馆论坛,1-9[2025-03-28]. http://kns.cnki.net/kcms/detail/44.1306.G2.20250226.1616.005.html.

[8] 袁 璐. 添加标识辨别 AI 视频以假乱真[N]. 北京日报,2025-03-20(007).

[9] 郭利敏,付雅明. 以大语言模型构建智慧图书馆:框架和未来[J]. 图书馆杂志,2023,42(11):22-30.

[10] 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(08):190-197.

[11] 童云峰,张学彬. 智慧图书馆领域生成式人工智能技术应用的法律容错机制[J]. 图书馆工作与研究,2024,(10):23-32.

[12] 周 纲,朱雯晶,张 磊. 开放、合作、智慧的图书馆未来——2024 年世界开放图书馆基金会会议(WOLFcon)综述[J]. 图书馆杂志,2024,43(12):77-88.

[13] 王鹏涛,徐润婕. AIGC 介入知识生产下学术出版信任机制的重构研究[J]. 图书情报知识,2023,40(05):87-96.

[14] 赵 杨,张 雪,范圣悦 AIGC 驱动的智慧图书馆转型:框架、路径与挑战[J]. 情报理论与实践,2023,46(07):9-16.

[15] 李佳轩,储节旺,杜秀秀. 关联、黑箱与赋能:AIGC 驱动智慧图书馆的转型路径[J]. 图书情报工作,2023,67(23):18-27.

[16] 陆 伟,刘家伟,马永强,等. ChatGPT 为代表的大模型对信息资源管理的影响[J]. 图书情报知识,2023,40(02):6-9.

[17] 长链推理(Chain of Thought)技术是一种改进的 Prompt 技术,旨在提升大模型在复杂推理任务上的表现。CoT 通过要求模型在输出最终答案之前显式输出中间逐步的推理步骤,从而增强模型的算数、常识和推理能力。其工作流程从传统的输入直接到输出,转变为输入经过思维链再到输出。

[18] SECI 模型是由野中郁次郎和竹内弘高提出的知识创造模型,

旨在解释隐性知识和显性知识如何在组织中相互转化和创造新知识。该模型将知识创造过程分为四个阶段:社交化(Socialization)、外化(Externalization)、组合(Combination)和内化(Internalization)。

[19] 陈廉芳. 图书馆知识管理 SECI 模型探讨[J]. 图书馆学研究, 2008(01):39-42.

[20] 曾成敏. AIGC 嵌入智慧图书馆建设:功能、风险及规制[J]. 新世纪图书馆,2024,(09):12-182.

[21] 卢　恒,陈章杰,周　知. 基于知识图谱的虚拟学术社区用户生成内容知识共聚框架研究[J]. 情报理论与实践,2023,46(12):157-166.

[22] 黄　鑫. 全产业链能力是 AI 竞争关键[N]. 经济日报,2025-02-21(006).

[23] 段玉聪. 抢占 AI 话语权:DeepSeek 的技术优势、战略布局与未来生态图景[J/OL]. 新疆师范大学学报(哲学社会科学版),2025,(04):1-17[2025-03-02].

[24] 郎林芳,黄世晴,王　珏,等. 元宇宙图书馆阅读推广服务创新发展研究[J]. 图书馆杂志,2023(10):55-63.

[25] 李　涛. 生成式人工智能 Sora 赋能智慧图书馆的探索与法律规制[J]. 图书馆建设,2024(5):128-137.

[26] 多模态版 DeepSeek-R1:评测表现超 GPT-4o,模态穿透反哺文本推理能力,北大港科大出品,已开源[EB/OL]. [2025-03-02]. https://36kr.com/p/3154441699891712.

[27] Bowlby J, Solomon M. Attachment theory[M]. Los Angeles, CA: Lifespan Learning Institute, 1989.

[28] 人类反馈强化学习(Reinforcement Learning from Human Feedback)是一种将人类反馈与强化学习相结合的方法,旨在通过引入人类偏好来优化模型的行为和输出。这种技术与训练动物的过程相似,给动物一个指令,比如"坐下",如果它照做了,你就给它一块零食(奖励);如果它没坐,你就说"不"(负奖励)。动物通过不断尝试和犯错,逐渐学会如何根据你的指令行动,以获得更多的食物。

[29] ChatGPT 接入美国加州州立大学系统,为 50 余万名师生提供 AI 服务[EB/OL]. [2025-02-08]. https://www.163.com/dy/article/JNQH6C7V05349YKB.html

[30] Kuhlthau C C. Inside the search process: Information seeking from the user's perspective[J]. Journal of the American society for information science, 1991, 42(5): 361-371.

[31] 混合专家模型(Mixture of Experts)是一种先进的神经网络架构,其包含多个专家网络和一个门控网络。专家网络并行存在,负责处理不同部分的数据或任务。门控网络则根据输入数据的特征,为每个专家分配权重,决定各专家对最终输出的贡献。

[32] 王　旭,程光辉,赵鸿玉,等. 技术性压力视角下智慧图书馆人工智能风险防控体系研究[J]. 图书馆学研究,2024,(07):56-69.

[33] 陈耿华,陆　睿. 侵犯商业秘密罪规制疑难问题及破解研究[J]. 竞争政策研究,2023,(06):5-21.

[34] 包赛君,肖冬梅. 生成式人工智能训练数据的著作权法因应:欧盟版权例外规则及其对我国的启示分析[J/OL]. 图书馆论坛,1-11[2025-02-27].

[35] 新明,叶　悦. 论反不正当竞争法"商业数据专条"的构建与完善[J]. 知识产权,2024,(06):93-110.

[36] 王立梅,张军强. 商业秘密刑民交叉案件审理模式的再思考[J]. 江淮论坛,2020,(01):116-123.

[37] 刘智锋,吴亚平,王继民. 人工智能生成内容技术对知识生产与传播的影响[J]. 情报杂志,2023,42(07):123-130.

[38] 郑海味,张伟君. 人工智能生成图片中用户的独创性贡献——以"AI 文生图"著作权案判决为例[J]. 中国出版,2024,(12):52-57.

[39] 陈　智. 韧性视角下 ChatGPT 应用的技术特性、演化过程与治理方略[J]. 科技进步与对策,2023,40(23):111-120.

[40] 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. 上海交通大学凯原法学院;2024:11.

[41] DeepSeek 开放平台服务协议[EB/OL]. [2025-02-12] https://platform.deepseek.com/downloads/DeepSeek%E5%BC%80%E6%94%BE%E5%B9%B3%E5%8F%B0%E6%9C%8D%E5%8A%A1%E5%8D%8F%E8%AE%AE.html.

[42] 李希梁,张钦昱. 生成式人工智能的反垄断规制[J]. 电子政务,2024,(05):53-63.

[43] 斜晓东. 论生成式人工智能的数据安全风险及回应型治理[J]. 东方法学,2023,(05):106-116.

[44] 最高人民法院关于审理不正当竞争民事案件应用法律若干问题的解释[EB/OL]. [2025-02-17] https://ipc.court.gov.cn/zh-cn/news/view-427.html.

[45] 吴　高. 数字环境下图书馆孤儿作品合理使用规则设计研究[J]. 图书馆建设,2023,(01):84-94.

[46] 北京市高级人民法院侵害著作权案件审理指南[EB/OL]. (2019-09-04)[2025-02-19] https://www.chinacourt.org/article/detail/2023/06/id/7335115.shtml.

[47] 黄宝菊. ChatGPT 生成式 AI 对金融行业的影响、挑战及应对[J]. 征信,2024,42(12):86-92.

[48] 曾子明,孙守强. 智慧图书馆人工智能风险分析与防控[J]. 图书馆学研究,2020(17):28-34.