

The application of DeepSeek large model to enable file search and use of innovation work research

Wei Li

Xining Archives, Xining, Qinghai, 810000, China

Abstract

With the rapid development of information technology, the archival work is facing new opportunities and challenges. This study focuses on the application of DeepSeek large model in the field of archival search and utilization, and explores its important role in innovative archival work. Through the analysis of large model technology and DeepSeek, file file using target and the present situation of the social phenomenon, expounds the DeepSeek model powerful natural language processing ability, knowledge understanding and reasoning ability how to break through these difficulties and operation and implementation, with the help of the model, file information intelligent classification, accurate retrieval and the depth of the user requirements, greatly improved the file efficiency and service quality. Research process, using the actual case analysis, theoretical analysis, literature review method, verify the DeepSeek big model can assign file file file using work in improving government service efficiency, optimize the allocation of resources, improve file experience, and provide scientific basis for government decision-making remarkable achievements, for the modernization of archives management transformation provides new ideas and practical reference.

Keywords

file search and utilization, DeepSeek big model, information security

应用 DeepSeek 大模型赋能档案查档利用创新工作研究

李薇

青海省西宁市档案馆, 中国·青海 西宁 810000

摘要

随着信息技术的飞速发展,档案工作面临着新的机遇与挑战。本研究聚焦于将DeepSeek大模型应用于档案查档利用领域,探索其对创新档案工作的重要作用。通过分析大模型技术与DeepSeek、档案查档利用目标及现状等社会现象,阐述DeepSeek大模型强大的自然语言处理能力、知识理解与推理能力如何突破这些困境及其操作与实施,借助该模型,档案信息的智能化分类、精准检索以及用户需求的深度理解得以实现,大幅提升了查档效率和服务质量。研究采用实际案例分析、理论分析、文献综述等方法,验证了DeepSeek大模型赋能下档案查档利用工作在提升政务服务效率、优化资源配置、改善查档人的体验,并为政府决策提供科学依据等方面取得了显著成效,为档案管理工作的现代化转型提供了新思路与实践参考。

关键词

查档利用; DeepSeek大模型; 信息安全

1 引言

随着人工智能技术的快速发展,大模型在多个领域展现出强大的应用潜力。DeepSeek 作为先进的大模型,具备自然语言处理、知识推理和智能推荐等功能,能够为档案管理带来创新解决方案,更好服务和支撑基层社会治理。本文将探讨如何利用 DeepSeek 大模型赋能档案查档利用的创新工作模式。

2 背景与现状

2.1 AI 大模型技术与 DeepSeek

2017年“人工智能”一词首次出现在《政府工作报告》,2024年政府工作报告提出,持续推进“人工智能+”行动,将数字技术与档案资源优势、群众需求更好结合起来,支持大模型广泛应用。《报告》中还强调要发展哲学社会科学、新闻出版、广播影视、文学艺术和档案等事业,加强智库建设。创新性表述的背后往往蕴含着深远的战略意义,每当科技行业的新名词出现在《政府工作报告》中,它们都会在不同阶段推动产业的转型升级^[1]。过去一年,中国的人工智能产业跑出了“加速度”。DeepSeek 在全球火爆出圈,创造了中国人工智能乃至中国创新产业的“高光时刻”。通过

【作者简介】李薇(1987-),女,中国青海西宁人,本科,馆员,从事档案研究。

DeepSeek 这类公司的兴起，我们可以看到中国在科技发展上的创新性和包容性，在数字化浪潮席卷全球的今天，档案行业正站在变革的十字路口，传统的档案管理方式面临着诸多挑战：海量数据的存储、检索效率低下、信息安全性不足。然而，随着人工智能技术的飞速发展，尤其是像 DeepSeek 这样的前沿科技公司的崛起，档案行业似乎找到了破局的关键，我们要抓住这次人工智能技术带来的机遇，使我们的数字技术与中国制造的优势、市场规模的优势能够充分结合，推动档案领域人工智能大模型的广泛应用，聚焦存史资政育人根本任务，努力让人工智能真正赋能档案服务管理，走进千家万户。

全国档案动态显示：档案研究逐步跟上节奏，牛力、金持、黎安润泽在《档案学通讯》2024年第6期发表《大模型在档案工作数智转型中的应用：新机遇、新模式和新转变》一文提到：目前，图书馆以及情报领域已率先开启对大模型应用的探索之旅，档案工作完全可以从中学取经验，为自身发展所用。档案部门应当积极主动作为，提升自身主观能动性，大力开展大模型等前沿技术的研究与实践应用，稳步推进档案工作的数智化升级进程^[2]。在具体实践中，上海市提出“抢抓战略机遇，加快建立与智能经济相适应的政策和制度体系”；银行业推进“AI+金融”融合发展；北京东营市成功推出了“DeepSeek+个人健康档案”应用。部分档案部门积极主动适应新形势新机制，守正创新、实干争先，率先行动了一批，如3月浙江宁波市档案馆举办《Deepseek大模型基础与AI档案业务交流》培训讲座；北京市昌平区档案馆依托数字档案馆建设，在国内档案领域率先接入 DeepSeek 大模型进行本地化部署，实现了档案检索功能重塑；福建三明市档案馆在“三明市档案信息网”和“三明档案”微信公众号上线 AI 智能问答服务等。

2.2 档案查档利用目标及现状

《“十四五”全国档案事业发展规划》提出“到2025年档案科技创新实现新突破，科研成果转化应用加速推进，科技力量在档案事业发展中的支撑作用显著增强”发展目标，提出“加快推进对重要档案数字化成果进行文字识别和语音识别”等几个方面的具体发展路径。2025年1月召开的全国档案局长馆长会议做的《守正创新 真抓实干 全面实现“十四五”档案事业发展目标任务》报告，指出全国档案查询利用服务平台累计接入2200余家档案馆、注册用户超20万人。目前全国各级档案部门正在在持续优化查档利用服务上功夫，积极扩大资源共享范围，国家查档利用的平台已在逐步完善，提供到馆查、跨馆查、网上查、等多渠道查档服务，但是仍有部分档案由于数字化程度和历史原因，查档利用依靠人工操作或简单技术手段，不能很好的凸显档案的专业化、精细化、特色化，当面对海量且类型繁杂的电子档案，以及多样化的用户需求时，传统方式暴露出效率低下的弊端。网友们公布的现有查档方式无外乎这么几种

1. 官网查询：登录12333全国人社政务服务平台 2. 公众号查询：关注户籍地或工作地的人力资源公众号，服务大厅里档案查询 3. 电话咨询：拨打12345便民热线，获取档案部门电话 4. 支付宝查询：搜索“档案查询”，或拨打市民热线。从侧面看出网民对传统档案查档方式简单、效率低、均衡性和可及性不高，且难以应对复杂查询的问题有所反映和吐槽。

2.3 DeepSeek 大模型对于档案查档利用的展望

DeepSeek 作为一家专注于人工智能和机器学习的科技巨头，DeepSeek 的核心技术之一是自然语言处理（NLP），它能够精准地理解和解析文本内容，其强大的技术实力和创新能力为档案行业带来了前所未有的机遇。在档案管理中，这意味着海量的纸质档案、电子文档等可以被快速而准确地分类、索引和检索^[3]。DeepSeek 大模型的有几个核心能力，能很好的与档案工作联系，如自然语言处理（NLP）：理解用户查询意图，提供精准回答；如知识推理：通过上下文和历史数据推理出最佳答案；如智能推荐：根据用户行为和偏好，推荐相关档案；如自动化处理，自动生成档案摘要、分类和标签。

3 操作与实施

3.1 数据准备

档案数据作为单位或组织信息资产的重要组成部分，其准确性、完整性和可用性对于业务决策、历史追溯、合规审查等至关重要，为充分发挥档案数据价值，必须要首先收集并整理档案数据，确保数据质量和完整性。需要进行如下几个步骤：对数据进行清洗和预处理，便于模型训练。①组建团队：抽调档案管理、信息技术、业务专家等人员，明确分工与职责。②资源筹备：准备数据收集与整理所需的软硬件设备，如存储服务器、数据采集软件等。③制定标准：确定档案数据分类标准、元数据规范和数据格式要求，做好数据初步筛选，剔除重复、明显错误的信息，记录问题数据等。特别是要注意涉密档案信息、敏感档案信息、不易公开档案目录条目的筛查。同时做好部分档案目录的扩展，如《关于王某某等5名同志职务任免的通知》文件目录，需要补齐王某某意外的其他4人名字，做到元数据要素齐全完整。

3.2 模型训练与优化

使用 DeepSeek 大模型进行训练，优化模型参数。以热线为例，将 DeepSeek 大模型与查档热线系统后台对接，整合地区多个部门的政策信息、查档清单及历史工单数据，构建统一的数据池，为智能化服务奠定坚实基础。如借助于 DeepSeek 大模型，实现工单智能处理（工单智能处理是指通过坐席助手实时语音转写后，根据通话内容自动总结需求，自动推荐知识内容；实时质检，自动会话小结；一键填单，智能分类和智能派遣）。工单智能处理可以适用于综合档案馆、大型企业多库房分头管理派单查询的情况，能够显

著提升政务服务效率、优化资源配置、改善查档人的体验，并为政府决策提供科学依据。

3.3 通过实际应用反馈，持续优化模型性能

以热线为例，通过智能助手功能开发，搭建政策问答、工单分类、诉求预判等核心功能，增加“**档案助手”（可以自命名）功能，在辅助话务员接听查档人诉求方面，能准确识别来电意图，实时调取知识库生成规范答复和快速解答群众疑问，满足机关、企业和群众个性化、多样化需求，减少人工客服的工作负担，释放人力资源用于处理复杂问题，显著提升政务服务效率、优化资源配置、改善查档人体验，是政府数字化转型的重要组成部分，推动政务服务向智能化、精准化方向发展。同时可以基于历史工单数据，训练 DeepSeek 大模型，使其能够精准识别查档人诉求语义，并根据训练情况，动态优化模型性能，确保系统能够不断适应新的诉求场景，在反复多遍的智能应答中，收集数据源逐步进行自我优化。

3.4 系统集成

整合全国档案查询利用服务平台、地区档案信息动态和鉴定后易于公开的档案目录库，并将训练好的模型集成到档案热线和搜索目录系统中，完善相关知识图谱，实现跨省市、跨部门信息语义关联检索。同时，建立动态更新机制，确保最新信息能够实时同步，打造大模型智能知识库。

3.5 确保系统稳定性和可扩展性

借助于大模型，实现智能预警功能，自动抓取集中性事件，根据敏感字（扬言报复、安全隐患等）自动分析紧急事件，并对以上事件进行告警通知，能够实时分析查档人诉求数据，识别潜在风险（如群体性事件、舆情危机），并在问题发生前发出预警。通过自动化预警机制，确保的档案部门能够快速响应，避免问题扩大化。自动收集数据，分析来话热点、用户偏好渠道、话务分布变化趋势等形成日周月报，为领导实时掌控系统运行情况及政府决策提供数据依据。智能报告功能能够生成跨部门的数据分析报告，推动多部门协同治理，提升整体服务能力，同时有助于推动全宗介绍、组织沿革、大事记、全宗卷编制工作，提升内容管理水平。

4 用户培训与反馈

4.1 对档案管理人员进行培训，确保熟练使用新系统

构建档案保密责任制，将档案保密责任具体化、层级化，覆盖到每一个涉及档案保密工作的组织和个体单元。组织学习《中华人民共和国档案法》《档案法实施条例》《保守国家秘密法》及其他法律法规，在“深、细、钻”上下功夫，带着信念学、带着使命学，把牢“档案工作姓党”政治属性，筑牢保密意识，锻造忠诚干净担当的政治品格，从而确保国

家秘密的安全与档案信息的妥善管理。同时通过举办实操演练、交流学习等方式，让档案管理人员详细了解 DeepSeek 的基础知识、大模型发展历程，结合实际案例等等，从而增强自信心，勇于接受新挑战，敢于尝试新方法，实现自身价值与工作业务的双提升。

4.2 收集用户反馈，持续改进系统功能

逐步构建覆盖“智能知识库—智能生成工单—智能辅助转办—智能生成工作报告”的全链条 AI 赋能体系，努力实现“智能分派零延迟、诉求解答全天候、服务效能再跃升”目标，为深入推进档案智库建设提供有力支撑。同时，各级档案部门还要时刻保持知识本领上的恐慌感和理论素养上的危机感，紧跟时代要求继续探索、用心学习、潜心钻研其他 AI 技术在档案查档利用中的应用。

5 挑战与对策

在数字化时代，档案信息的安全至关重要。我们要加强数据加密和访问控制，确保档案数据安全，坚持安全底线，强化风险防控，加强应急管理。DeepSeek 在数据安全方面投入了大量资源，采用了先进的加密技术、访问控制技术和数据备份策略，为档案数据筑牢了坚固的“防火墙”。无论是纸质档案的数字化过程，还是电子档案的存储和传输，DeepSeek 都能确保数据的安全性和完整性。例如，在金融行业的档案管理中客户，信息和交易记录等敏感数据需要严格保密。DeepSeek 的数据安全解决方案能够防止数据泄露、篡改和丢失，让档案管理者 and 用户放心使用。但在档案结合具体实操中，我们还是要高度关注大模型嵌入档案工作过程中可能出现的应用风险与技术瓶颈，诸如数据治理难题、结果控制困境以及数字素养提升需求等问题。适时要考虑档案部门 AI 技术人才短缺现状，借助得力放心的第三方进行策划、检测、评估。

6 结论

应用 DeepSeek 大模型赋能档案，不仅仅是技术的简单应用，更是一场深刻的变革，它为档案管理者带来了全新的思路 and 工具，为档案用户提供了更加便捷和高效的服务，为档案行业的发展注入了强大的动力。通过合理的实施步骤和应对挑战，档案查档利用工作将迎来智能化、自动化的新时代，推动档案工作向智能化、高效化方向发展。

参考文献

- [1] 牛力、金持、黎安润泽.大模型在档案工作数智转型中的应用：新机遇、新模式和新转变[J].档案学通讯,2024(第6期)
- [2] 档案行业迎来“智慧革命”：DeepSeek如何重塑档案管理的未来？[恒小智聊档案EB/OL].[2025-02-12].
- [3] TQC在档案管理中的运用[J].档案学研究,1990,(02):24-30.