

# Research on the application of big data technology in archives management

Haijun Huo

Dongsheng District Employment Service Center, Ordos City, Ordos, Inner Mongolia, 017000, China

## Abstract

Big data technology, with its strong capabilities in data storage, processing and analysis, has brought new opportunities for the field of archives management. The traditional archives management mode is limited by the problems of resource dispersion, low retrieval efficiency and low information utilization rate, and the introduction of big data technology can effectively improve the intelligence and efficiency of archives management. This paper starts with the basic characteristics of big data technology, analyzes its application advantages in archival information collection, storage and retrieval, data mining and sharing and collaboration, discusses the technical bottlenecks and management challenges in the current application process, and puts forward targeted solutions. The research shows that the wide application of big data technology will promote the development of digitalization, intelligence and refinement of archives management, but its promotion still needs the coordinated support of system, technology and talents.

## Keywords

big data technology; file management; informationization; data mining; resource sharing

# 大数据技术在档案管理中的应用研究

霍海军

鄂尔多斯市东胜区就业服务中心, 中国·内蒙古 鄂尔多斯 017000

## 摘要

大数据技术以其在数据存储、处理和分析方面的强大能力,为档案管理领域带来了新的机遇。传统档案管理方式受限于资源分散、检索效率低及信息利用率不高等问题,而大数据技术的引入则能有效提升档案管理的智能化和效率。本文从大数据技术的基本特征入手,分析其在档案信息采集、存储与检索、数据挖掘及共享协作等方面的应用优势,探讨当前应用过程中存在的技术瓶颈及管理挑战,并提出针对性的解决方案。研究表明,大数据技术的广泛应用将推动档案管理的数字化、智能化与精细化发展,但其推广仍需制度、技术与人才多方面协同支持。

## 关键词

大数据技术; 档案管理; 信息化; 数据挖掘; 资源共享

## 1 引言

档案是记录和传承人类社会活动的重要载体,其管理水平直接影响信息资源的开发与利用。传统档案管理主要依赖人工操作和纸质媒介,存在数据分散、检索效率低下及存储成本高等问题。在信息化浪潮的推动下,档案管理逐步向数字化方向发展。然而,随着档案数量的爆发式增长和复杂化程度的提高,现有的数字化档案管理系统在处理大规模数据时力不从心,难以满足快速、精准、高效的信息需求。

大数据技术作为一种新兴的信息处理工具,凭借其强大的存储、计算和分析能力,为档案管理的现代化提供了技术支撑。通过引入大数据技术,档案管理可以实现数据的实时采集与处理,提升信息的精准性和利用率,构建全新的智

能化档案管理模式。然而,在技术应用的过程中,也面临数据安全、标准化不足及专业人才短缺等挑战。本文旨在通过分析大数据技术在档案管理中的具体应用,探讨优化档案管理效能的路径,为推动档案管理的智能化转型提供理论依据和实践参考。

## 2 大数据技术在档案管理中的作用

### 2.1 提升档案信息的采集效率

大数据技术通过自动化采集工具和智能感知技术,能够大幅提高档案信息的采集效率。在传统模式中,档案的采集通常需要依赖人工分类整理,费时费力且容易出错。而大数据技术引入后,依托传感器网络和光学识别技术,可以实现对纸质档案的批量扫描和数字化处理。与此同时,通过互联网和物联网技术的结合,档案管理机构可以实时获取动态数据流,例如社会事件的音视频记录、科研成果的数字化存储等。这些技术的应用显著缩短了档案采集的周期,保障了

【作者简介】霍海军(1979-),男,中国内蒙古鄂尔多斯人,研究生学历,从事档案管理研究和人力资源服务。

数据的完整性和时效性。

## 2.2 优化档案存储与检索系统

档案的存储和检索是管理工作的重要环节，大数据技术通过分布式存储和快速检索算法，为这一环节提供了有力支持<sup>[1]</sup>。在存储方面，传统单一数据库模式难以应对多源异构数据的存储需求，而大数据技术通过构建分布式存储系统，实现了对海量数据的高效管理。特别是在应对档案格式多样化（如文本、图像、音频、视频）时，大数据技术能够自动分类存储并压缩冗余信息。在检索方面，大数据技术依托全文搜索、语义分析及自然语言处理技术，可以快速定位目标信息，提高档案利用效率。例如，档案管理系统可以通过关键词关联、内容匹配和智能推荐，实现精准、全面的信息检索，为用户提供更便捷的服务。

## 2.3 挖掘档案数据的潜在价值

档案数据的价值不仅在于保存和检索，更在于通过深度分析挖掘出隐藏的信息模式和趋势。大数据技术中的数据挖掘和机器学习算法，可以对历史档案进行多维分析，揭示其内在关联。例如，通过对历史档案数据的时间序列分析，可以预测某些领域的未来发展趋势；通过社交网络分析技术，可以揭示不同主体之间的联系及其演变规律。此外，数据可视化技术能够将复杂的档案数据直观呈现，帮助决策者更好地理解信息，从而为公共政策制定和社会治理提供有力支持。

# 3 大数据技术在档案管理中的应用实践

## 3.1 智能化档案分类与整理

档案分类与整理是档案管理的基础工作，大数据技术通过智能化算法优化了这一过程。在具体应用中，可以利用文本分类算法对电子档案进行自动化分类，将不同主题和类型的档案归档存储。这种方法不仅减少了人工操作的烦琐性，还提高了分类的精准度和一致性<sup>[2]</sup>。例如，在高校档案管理中，系统可根据档案内容的关键词和主题模型，将学生档案、科研成果及行政文件自动分类存档，为后续的管理与利用提供便利。

## 3.2 构建多元化档案检索平台

多元化的档案检索平台能够满足用户对档案信息的多样化需求。通过大数据技术的支持，档案管理系统可以将传统的目录检索扩展为全文搜索及多维分析。例如，用户在检索过程中可以通过自然语言提问的方式获取目标信息，系统则通过语义分析和关联挖掘为其生成匹配结果。此外，多元化检索平台还支持多模态数据的联合查询，用户可以同时搜索文本、图像及视频档案，获取全面的背景信息。这种检索方式极大地提升了档案利用效率。

# 4 大数据技术应用中的挑战

## 4.1 数据安全与隐私保护

大数据技术的应用使得档案管理面临严峻的数据安全

挑战。例如，海量档案数据在采集、存储和传输过程中，可能因技术漏洞或人为干预导致信息泄露。这不仅威胁到档案管理机构公信力，还可能对个人隐私和社会稳定产生不利影响。特别是在云计算和分布式存储广泛应用的背景下，数据跨平台传输的复杂性进一步增加了安全隐患。为此，需要通过完善的加密技术、访问控制及监测系统，加强档案数据的安全保护。例如，可在数据传输环节引入端到端加密技术，确保数据在传输过程中不被拦截或篡改。同时，利用人工智能技术加强网络威胁检测，实时监控系统异常活动并及时采取应对措施。此外，还应通过法律法规明确数据安全责任，对违规操作或管理不当的行为实施严厉处罚，从而构建多层次的安全保障体系，全面提升档案管理中的数据保护能力。

## 4.2 技术适配性与标准化不足

尽管大数据技术具有广泛的应用潜力，但其适配性和标准化问题在实际操作中较为突出。例如，现有的档案管理系统可能无法直接兼容新的大数据技术，需要进行复杂的系统改造，这不仅增加了实施成本，还可能导致数据迁移中的信息丢失或格式混乱。此外，不同机构之间的数据格式和管理标准缺乏统一性，制约了档案的互联互通和共享利用。在多源异构数据的处理中，缺乏统一的分类和标签规范，导致不同机构之间的数据交换困难甚至不可能。因此，应推动建立统一的数据管理标准体系，包括文件格式、存储结构和元数据定义等内容。同时，鼓励研发适配性强的档案管理系统，支持多类型数据的灵活存储与分析。通过标准化和技术适配性的双重提升，不仅能够简化大数据技术的落地应用，还能提升跨部门档案资源的共享与协作创造条件，从而提高档案管理的整体效能。

## 4.3 专业人才的短缺与培养

大数据技术的应用对档案管理人员提出了更高的专业要求。目前，许多档案管理机构面临技术人才短缺的问题，制约了大数据技术的推广应用。一方面，传统档案管理人员通常缺乏数据分析、算法开发等技术能力，难以胜任复杂的技术操作；另一方面，具备大数据技术背景的人才对档案管理工作又相对陌生，导致跨领域复合型人才严重不足。为解决这一问题，应从教育和实践两个方面着手。一是加强相关领域的职业教育，将大数据技术应用纳入档案管理专业课程体系，并通过模拟实验和实操训练提升学生的实际能力<sup>[3]</sup>。二是推动档案管理机构与高校及科研机构合作，建立联合培养机制，为现有从业人员提供继续教育和技能提升机会。

# 5 应对策略与发展建议

## 5.1 构建数据安全保障体系

针对数据安全挑战，应建立全方位的数据保护机制，覆盖档案管理的各个环节。在档案数据的采集与存储环节，可以通过加密算法、区块链技术和权限管理系统等多重手段限制数据的非法访问，从根本上提高数据保护水平。例如，

采用非对称加密算法对关键数据进行加密存储,并结合身份验证机制确保只有授权人员能够访问档案数据。档案管理机构还应制定数据安全操作规程,包括数据的传输加密、备份与恢复策略,避免因技术漏洞或意外事故导致的数据丢失或泄露。此外,应定期进行网络安全测试与评估,及时发现并修复系统漏洞,防范恶意攻击和信息泄露的风险。为提升整体安全意识,档案管理机构可组织全员安全培训,让工作人员掌握基础的数据安全知识和应急处理技能。在政策层面,政府应出台相关法规,为档案数据安全提供法律保障,从而形成“技术+制度”的双重防护体系,确保档案数据的安全性和完整性。

### 5.2 推动标准化与互联互通建设

为提升大数据技术在档案管理中的适配性,应积极推动档案管理的标准化建设与互联互通。首先,政府和行业协会应牵头制定统一的档案数据格式、分类编码和存储标准,解决目前各机构档案数据格式多样、管理流程不统一的问题,从而为大数据技术的广泛应用奠定基础。例如,可以建立覆盖全国范围的档案分类与编码体系,确保数据在不同平台间的无缝转换和共享。其次,应通过构建开放的数据共享平台,推动跨部门、跨地区的档案协作与资源共享。这样的平台不仅能提升数据的利用效率,还能为社会公众提供更加便捷的档案服务。同时,政府应提供政策支持和资金投入,帮助经济欠发达地区提升档案管理的基础设施水平,解决技术设备不足和网络覆盖有限的问题。通过多方协作,档案管理系统可以形成统一、高效的资源网络,不仅推动档案管理的信息化进程,也为社会治理和科研活动提供强有力的支持。

### 5.3 加强专业人才培养与技术支持

档案管理的数字化转型离不开专业人才的支撑,因此应从教育和实践两方面加强人才培养和技术支持。在教育层面,各大高校和职业培训机构应将大数据技术与档案管理相结合,开设相关课程或专业模块,培养既懂档案业务又具备技术能力的复合型人才<sup>[4]</sup>。例如,可以通过模拟档案管理场景,让学生掌握数据采集、存储、检索及分析的全流程操

作。在职业培训方面,档案管理机构可以定期举办技术讲座和实操培训,让现有工作人员了解最新的大数据技术及其应用方法。此外,档案管理机构还可以与科研机构和企业合作,建立联合实验室或人才培养基地,为技术创新和实践提供试验田。在技术支持方面,应组建专业的技术团队,提供日常系统维护和技术咨询服务,确保档案管理平台的稳定运行。通过教育培训、实践合作和技术支持的综合推进,档案管理领域的人才短缺问题将得到有效缓解,为大数据技术的深入应用提供强有力的保障。

## 6 结语

大数据技术的引入为档案管理领域带来了前所未有的发展机遇,同时也提出了诸多挑战。在采集、存储、检索及数据分析等方面,大数据技术推动了档案管理向智能化、精细化的方向发展,显著提升了工作效率和信息利用价值。然而,其全面推广应用仍需克服数据安全、标准化建设及专业人才匮乏等问题。在未来的发展中,通过制定数据安全保护机制、推动档案管理的标准化建设及加强专业人才培养,档案管理将实现更加高效、安全和创新的技术应用。此外,政策支持、技术创新和多方协作的进一步深化,将为大数据技术在档案管理中的全面应用提供保障。未来,随着大数据技术的不断成熟和普及,档案管理有望成为推动社会发展、传承历史文化的重要技术驱动力,为社会提供更加高效、高质量的信息服务,同时也为数字化社会的构建贡献力量。通过技术、制度与人才的共同发力,档案管理的现代化进程将迈向新的高度,实现更大范围的资源共享和社会价值。

### 参考文献

- [1] 高严,于大海,江晨,等.大数据技术的特点与应用分析[J].电子技术,2022,51(10):142-144.
- [2] 张虹.探讨信息时代医院综合档案管理的新思路[J].兰台内外,2024,(36):63-65.
- [3] 王宇飞,尚震.药品生产线档案管理智能化解决方案研究[J].流程工业,2024,(12):50-53.
- [4] 张琦.人工智能时代计算机应用技术发展研究[J].数字通信世界,2024,(12):37-39.