

# 大数据时代档案管理模式变化研究

刘海彦

准格尔旗矿区事业发展中心 内蒙古 鄂尔多斯 017000

**摘要:** 大数据时代, 数据呈爆炸式增长, 档案管理模式迎来深刻变革。传统档案管理在数据存储、检索利用等方面面临挑战, 而大数据技术为其带来新契机。本研究聚焦大数据时代档案管理模式变化, 分析大数据对档案收集、整理、存储、检索及利用等环节的影响, 探讨档案管理在理念、技术、方法上的创新, 旨在构建适应大数据环境的档案管理新模式, 提升档案管理的效率与质量, 充分发挥档案价值, 为相关领域发展提供有力支撑。

**关键词:** 大数据时代; 档案管理模式; 变化与创新

引言: 在数字化浪潮的强力推动下, 大数据时代已然来临, 海量数据以前所未有的速度和规模产生与传播。档案管理作为信息管理的重要分支, 其传统模式在大数据的冲击下, 逐渐暴露出存储容量有限、检索效率低下、信息利用不充分等诸多弊端。为顺应时代发展, 提升档案管理的效能与价值, 深入探究大数据时代档案管理模式的变化成为当务之急。本研究旨在剖析大数据对档案管理的影响, 探寻新模式构建路径, 助力档案管理实现现代化转型。

## 1 大数据时代档案管理的核心特征与挑战

### 1.1 大数据时代档案管理的核心特征

(1) 数据规模海量性与多样性: 全球档案数据占比超15%, 且随数据爆炸持续增长, 如济南市档案馆仅婚姻档案数据就达3.5T。数据类型涵盖结构化的政务凭证、企业ERP日志, 与非结构化的电子证照、社交媒体记录等多元形态。(2) 数据价值密度低与动态性: 档案数据多以碎片化存在, 需深度挖掘方可释放价值, 例如企业客户档案中的行为数据需专业分析才能支撑决策; 同时数据随业务开展实时更新, 呈现动态流转特征。(3) 数据生成与传播的实时性: 政务服务、企业运营等场景中, 电子文件实时生成并需即时归档, 如银川市档案馆通过智能系统实现档案数据的快速采集与同步。(4) 跨平台与跨领域的数据关联需求: 民生服务需打通民政、税务等多部门档案, 济南市已实现164万条婚姻档案与不动产登记等系统的关联共享, 体现跨领域关联的核心需求。

### 1.2 传统档案管理模式的局限性

(1) 存储与检索效率低下: 依赖人工分类编目, 某省级档案馆年处理100万份电子文件需耗时6个月, 且物理存储占用大量空间, 检索需人工翻阅。(2) 数据安全与隐私保护不足: 缺乏加密与权限管控体系, 易出现内部滥用或外部攻击风险, 传统实体档案也存在失窃、损

毁隐患。(3) 缺乏智能化分析与决策支持能力: 侧重“存数据”而非“用数据”, 无法通过关联分析实现客户流失预测、政策效果评估等价值挖掘。(4) 跨部门协作与资源共享困难: 数据分散形成“信息孤岛”, 跨部门查询需反复沟通, 公众获取信息常需线下申请, 服务体验差。

## 1.3 大数据时代档案管理面临的挑战

(1) 技术挑战: 需突破多源数据采集瓶颈, 应对海量存储压力, 如济南市需处理4100多万页纸质档案的数字化转换; 同时缺乏成熟的智能化分析工具, 难以高效挖掘数据价值。(2) 管理挑战: 各领域数据标准不一, 元数据规范混乱导致互操作性差; 电子档案长期保存受技术迭代影响, 需持续解决格式兼容问题。(3) 法律与伦理挑战: 数据主权界定模糊, 跨境传输易引发法律冲突; 隐私保护面临泄露风险, 且现有法规滞后于技术发展, 合规性判断难度大。

## 2 大数据时代档案管理模式的变化与创新

### 2.1 管理理念的变化

(1) 从“被动保管”到“主动服务”的转变: 传统模式以“存得好”为核心, 如今转向以用户需求为导向的“用得好”。例如某制造企业通过分析档案访问数据, 发现研发团队频繁查阅锂电池材料检测报告, 主动推送相关历史失效案例, 助力新品研发效率提升20%。政务领域中, 档案馆通过预判民生需求, 提前关联婚姻、不动产等档案, 支撑“一网通办”场景下的快速审批。(2) 从“单一档案”到“全生命周期管理”的延伸: 构建“采集-整理-存储-利用-销毁”闭环体系, 突破传统“阶段性管理”局限。档案宝系统实现从电子文件生成到合规销毁的全程管控, 某金融机构通过该系统完成10万份合同的全生命周期管理, 避免超期档案占用资源。中国电建更将管理延伸至业务前端, 实现电子文件形成

与归档的同步衔接<sup>[1]</sup>。(3)从“封闭系统”到“开放共享”的协作模式:打破“信息孤岛”,推动跨层级、跨领域协同。中国电建通过联盟链技术,实现设计、施工、运营等多单位档案无缝流转,构建多方可信协作环境;政务领域已建成164万条婚姻档案跨部门共享机制,支撑不动产登记等业务高效办理。

## 2.2 技术架构的升级

(1)云计算与分布式存储技术的应用:解决海量档案存储难题,实现弹性扩展。某省级档案馆采用阿里云分布式存储,将1000万份电子档案存储成本降低60%,同时支持万人并发访问。档案宝系统通过云端部署,让员工随时随地调取跨区域项目档案,检索响应时间缩短至秒级。(2)大数据分析 & 挖掘技术的引入:机器学习与自然语言处理技术赋能价值挖掘。在档案整理环节,通过BERT模型微调的NLP算法,实体抽取准确率较通用模型提升15%,可自动识别合同中的甲方、乙方、金额等关键信息;某零售企业利用数据分析识别出30%的低价值档案,优化存储策略年省百万成本。(3)区块链技术在档案真实性验证中的应用:中国电建搭建联盟链平台,将电子文件哈希值与操作记录上链存证,构建“来源可查、去向可追”的认证体系。其开发的真实性验证平台,可快速核验文件完整性,解决数字孪生环境下电子档案的信任难题,操作日志还能为监管检查提供可靠证据。(4)物联网与传感器技术对档案环境监控的支持:在实体档案库房部署温湿度、烟雾传感器,数据实时上传云端。某档案馆通过物联网系统监测到库房湿度异常,自动触发通风设备调节,避免3万份纸质档案受潮损毁;该技术还可联动视频监控,实现实体档案异动的即时告警<sup>[2]</sup>。

## 2.3 管理流程的优化

(1)档案收集:自动化采集与多源数据整合,打破人工录入瓶颈,实现“数据多跑路”。档案宝支持对接OA、ERP等10余种业务系统,某地产企业通过API接口自动归档全国项目图纸,跨部门调取时间从4.2小时缩至8分钟;移动端采集功能让物流企业实时记录运输档案,纠纷处理周期从7天降至1天。(2)档案整理:智能分类与标签化管理,替代人工编目,提升标准化水平。系统可自定义“全宗-案卷-文件”三级分类体系,某知识产权机构通过元数据强制录入,实现专利档案的全程追溯;结合NLP技术自动生成“2024Q3-华为-合作协议”等标签,分类准确率达98%。(3)档案存储:分布式存储与容灾备份策略,采用“多地多活”架构保障数据安全。某金融机构将档案数据同步存储于北京、上海、广州三地节

点,即使单一区域系统故障,也能通过备份节点快速恢复服务;加密算法对存储与传输数据双重保护,防止信息泄露<sup>[3]</sup>。(4)档案利用:个性化推荐与可视化分析,基于用户画像实现精准服务。知识图谱系统可根据研发人员查询历史,主动推荐相关专利与项目档案;通过可视化界面展示“人员-项目-文件”关联网络,某科研团队点击节点即可快速获取完整知识链条。

## 2.4 服务模式的创新

(1)基于用户需求的精准化服务:通过行为分析实现“千人千面”服务。政务档案馆根据公众查询记录,为创业者精准推送工商注册、税收优惠等关联档案;企业档案系统针对财务部门高频检索需求,优化报销凭证查询路径,审批效率提升50%。(2)跨机构档案资源共享平台建设:打破部门壁垒,构建一体化服务体系。某省建成涵盖民政、税务、不动产等12个部门的共享平台,公众办理房产过户时,系统自动调取关联档案,材料提交量减少70%;中国电建的跨单位平台实现设计与施工档案无缝流转,项目协同效率提升40%。(3)档案知识图谱构建与语义检索:突破关键词检索局限,实现“知识发现”。输入“2020年朝阳环保政策涉及企业”,图谱可解析实体关系,直接返回关联档案集合;某律所构建“案件-法规-律师”图谱,支持“张三承办的2015年A公司研发项目财务记录”等路径检索,查询效率提升70%。(4)移动端与社交化档案服务:拓展服务触达渠道。档案宝集成钉钉、企业微信等平台,员工通过手机即可完成档案借阅审批,处理时间从2天缩至2小时;政务部门推出微信小程序,公众可在线查询婚姻、学历等档案,年减少线下办理量80万人次。

## 3 大数据时代档案管理模式发展的保障机制

### 3.1 政策与法规保障

(1)完善档案数据开放与共享的法律法规:2024年施行的《中华人民共和国档案法实施条例》明确提出“促进全国档案数字资源跨区域、跨层级、跨部门共享利用”,为数据开放提供法律依据。但现存法规仍存在空白,需进一步细化数据开放范围与权限划分,例如明确民生类档案的开放边界,同时建立数据共享的责任追究制度,对擅自扩大共享范围的行为设定处罚标准,如某省档案馆因违规共享涉密档案被责令整改的案例便凸显了法规完善的紧迫性。(2)制定大数据档案管理标准与规范:国家档案主管部门已着手构建电子档案管理标准体系,《档案法实施条例》要求电子档案需满足“来源可靠、程序规范、要素合规”等条件。实践中需进一步统一元数据格式、数据分类编码等标准,如针对

医疗档案与政务档案的交叉领域,明确关联数据的对接规范;同时建立标准动态更新机制,适配AI、区块链等新技术应用场景,避免标准滞后于技术发展。

### 3.2 技术支撑体系

(1) 构建档案大数据技术中台:技术中台需整合数据采集、存储、分析等核心能力,实现多源数据的统一管控。例如某省级档案馆搭建的技术中台,通过API接口对接12个政务系统,实现档案数据的集中汇聚与标准化处理;中台还应具备弹性扩展能力,支持存储资源随数据量增长动态扩容,同时集成数据质量检测模块,自动识别重复、错误档案数据,提升数据治理效率。(2) 加强核心技术研发:AI技术研发聚焦智能分类与分析,基于BERT模型优化的档案文本识别系统,可将合同档案关键信息提取准确率提升至95%;区块链技术侧重构建可信存证体系,中国电建通过联盟链记录档案操作日志,实现“来源可查、去向可追”。未来需重点突破跨模态档案分析、长周期数据存储等技术瓶颈,推动核心技术与档案业务深度融合<sup>[4]</sup>。

### 3.3 人才队伍建设

(1) 培养复合型档案管理人才:需构建跨学科培养模式,课程体系涵盖档案学、计算机科学、法律学等多领域知识,如某高校开设“档案大数据分析”课程,融合数据挖掘与档案管理理论。实践中通过校企合作强化技能训练,组织学生参与企业档案数字化项目,培养既懂档案编目规范,又能操作大数据分析工具的复合型人才,满足“技术实操+管理规划”的双重需求。(2) 建立持续学习与培训机制:搭建“理论学习+实践演练”的培训体系,定期开展AI技术应用、数据安全法规等专题培训,如国家档案局组织的电子档案管理系统实操培训已覆盖全国2000余家档案馆。同时建立考核认证制度,将技术能力纳入档案管理人员职称评定标准,鼓励员工参与跨部门项目交流,通过实践提升跨领域协作与问题解决能力,适应档案管理模式的动态变化。

### 3.4 安全与隐私保护

(1) 数据加密与访问控制技术:采用“传输-存储-使用”全流程加密方案,传输阶段运用SSL协议,存储环节实施AES-256加密算法;访问控制通过“角色-权限”双重认证,如某企业档案系统为财务、研发等部门设置差异化权限,仅允许核心岗位访问敏感档案。同时引入行为审计技术,实时监控异常访问行为,对多次尝试越权操作的账户自动锁定,防范内部数据泄露风险<sup>[5]</sup>。(2) 档案数据脱敏与匿名化处理:针对包含个人隐私的档案,采用动态脱敏技术,在开发测试场景中屏蔽身份证号、住址等敏感信息;匿名化处理通过删除标识信息、打乱数据关联关系实现,如将医疗档案中的患者姓名替换为随机编码。需建立脱敏效果评估机制,确保处理后的数据既保护隐私,又保留统计分析价值,平衡数据利用与隐私安全的关系。

### 结束语

大数据时代为档案管理带来了前所未有的机遇与挑战,促使档案管理模式在理念、技术、方法等多方面发生深刻变革。本研究深入剖析了这些变化,并提出了适应大数据环境的新模式构建策略。然而,档案管理模式创新是一个持续发展的过程,未来仍需不断探索与完善。我们应紧跟时代步伐,积极运用新技术,持续优化管理模式,以提升档案管理的效率与质量,让档案在大数据时代更好地服务于社会发展和人类进步。

### 参考文献

- [1]荆雪飞.大数据时代提升档案管理工作质效探析[J].兰台内外,2024,(22):31-33.
- [2]孙惠媛.大数据时代国有企业档案信息资源整合与开发利用研究[J].办公室业务,2024,(11):27-29.
- [3]赵新民,秦广久,马向阳.大数据时代档案管理模式变化研究[J].山东农业工程学院学报,2020,37(5):184-186.
- [4]邢莎莎.大数据时代人事档案管理创新与模式构建[J].黑河学刊,2023,(5):99-100.
- [5]张烈芳.大数据环境下事业单位综合档案管理创新研究[J].兰台内外,2023,(18):39-40.