

数字档案馆电子文件长期保存策略研究

滕军林 周 涛 蒋中利

大唐环境产业集团股份有限公司 北京 100097

【摘 要】：数字化转型背景下电子文件载体易损、格式失效等问题加剧长期保存压力，本文梳理数字档案馆电子文件长期保存内在逻辑与现实矛盾，从多维可信存管、业务协同对接、跨系统标准统一搭建可信运行支撑体系，围绕标准、技术、制度维度创新保存实施策略，从人才、资源、风险层面完善落地保障路径，为数字档案馆电子文件长效可信保管提供可落地的实操方案。

【关键词】：数字档案馆；电子文件；长期保存

DOI:10.12417/2811-0722.26.09.037

引言

数字业务海量产出异构电子文件，传统纸质档案保管模式已难以适配数字化归档管护需求，载体衰减、标准不统一、系统割裂等现实困境制约电子文件永久可读留存。现有研究多局限于单一技术或单一标准的局部优化，缺少覆盖全流程的一体化保存体系完整构建思路。本文立足数字档案馆运行实际，逐层拆解长期保存底层逻辑、运行体系、实施策略与保障路径，形成完整可落地的长效保存方案。

1 数字档案馆电子文件长期保存的底层逻辑与现实动因

数字载体天然存在易损耗、格式迭代快、数据衰减等短板，成为电子文件长期保存的底层制约，传统静态纸质档案管护模式难以承载海量数字化业务文件的存续需求^[1]。档案管理模式从双套制并行管护转向单套制长效运维，推动档案行业重塑适配数字化载体的管护思维与工作体系。数字化业务持续生成各类异构电子文件，不同业务系统适配的元数据规范存在明显差异，海量数据持续累积引发存储扩容、真伪核验、长期可读等多项管护难题，现有保存体系尚未形成覆盖文件生成至永久留存全周期的统一运行规范。各类档案管理机构开展电子文件管护工作时普遍侧重短期资源利用，忽视长效留存建设，权责划分模糊、标准落地不到位等问题频繁出现。梳理电子文件长期保存的内在运行规律，剖析当前管护场景中的各类现实困境，能够为可信保管机制搭建、保存策略落地提供理论支撑与实践依据。

2 数字档案馆电子文件长期保存可信运行支撑体系构建

电子文件可信保管全流程闭环运行，具体流程如下图所示。



图1 电子文件全生命周期可信保存运行流程

2.1 电子文件全生命周期多维可信存管机制设计

真实性、完整性、可用性与安全性是电子文件长期可信保存的核心内核，全生命周期可信存管机制贯穿文件生成、流转、归档、存储、读取及销毁全部环节，打破传统归档工作单次核验的粗放管理模式，依托分阶段、节点化的可信核验流程，实现档案全流程可控可溯源。文件生成阶段，系统绑定操作人员专属数字身份证书，文件修改、批注、附件增减等所有操作均会生成不可篡改的哈希校验值，操作轨迹、时间戳、操作人员相关信息同步固化至元数据体系，锁定文件原始状态，规避人工后期整理引发的数据失真问题。

电子文件跨部门流转环节设置动态核验节点，流转完成后系统自动比对文件前后版本哈希值，一旦数值异常将立即阻断传输并推送预警，完整留存操作日志用于问题溯源核查。归档移交环节搭建三层闭环核验模式，先核查文件格式合规性，单独处置无法实现长期读取的非标准文件，再比对业务操作元数据与文件本体元数据，排查信息缺失、字段篡改等问题，依托多方电子签名完成归档确认，形成具备法律效力的归档凭证^[2]。

长期存储阶段开启季度自动化巡检模式，批量复核文件哈希数据，针对介质老化、磁盘故障引发的数据缺损问题，借助异地备份资源完成数据修复补齐。到期电子文件由档案、业务、法务三方团队联合开展价值鉴定，全程固化操作记录，通过专业数据粉碎技术完成文件销毁，同步归档留存销毁凭证。全流

程分段核验、元数据固化、周期性巡检的工作模式,可有效化解电子文件易篡改、易损毁的保存风险,保障长期保存工作的可信度。

2.2 档案保管与业务系统双向协同赋能保存模型搭建

现阶段电子文件管护工作普遍存在业务系统与档案馆单向移交的割裂问题,业务端办结业务并推送文件后便终止相关管理流程,档案馆接收文件后无法还原原始业务场景,大幅降低电子文件保存可信度与实际利用价值。研究以标准化数据接口对接为核心,搭建业务与档案双向互通、协同赋能的闭环保存模型,落地多场景协同实操机制。

模型第一层嵌入前置档案采集模块,在各类业务系统中增设轻量化归档端口,业务办结后系统自动甄别有效电子文件,依据预设元数据模板抓取核心业务信息,无需人工二次加工整理,每日定时将合规文件推送至档案馆预归档库,未办结的在途文件单独设置缓存分区并锁定修改权限,保障文件原生完整性。模型第二层搭建档案馆反向赋能通道,文件归档著录工作完成后,向业务系统开放受控查阅权限,业务人员可调取历史档案辅助日常办公,所有查阅行为自动生成溯源台账。档案馆按月统计业务归档完整率、元数据合规率、格式规范率等核心运维指标,生成整改报告推送至对应业务部门,配合业务系统归档弹窗提醒,从前端规范文件生成标准。

模型第三层适配批量业务场景,重大项目、专项督查等大批量文件产出工作中,档案工作人员可依托业务系统前置模块,在线划定归档范围、调整元数据采集规则,规避线下批量传输产生的各类漏洞。档案部门定期联动业务团队评估五年以上低频使用档案,分类界定永久保管、定期保管与到期清理范畴,实现档案资源动态优化更新。前置采集、反向赋能、联动评估的运行模式,打破业务与档案领域的信息壁垒,双向保障电子文件归档质量与长期保存资源的稳定性。

2.3 跨业务、跨馆系统有机衔接与保存标准规则统一

异构系统适配度低、行业标准不统一是电子文件跨业务、跨馆长期可信保存的核心阻碍,不同业务系统的接口协议、元数据字段、存储格式存在明显差异,各级档案馆管护规范互不兼容,极易引发文件移交失败、数据错乱、留存失效等各类问题。系统改造、标准落地、常态化运维核查相结合的方式,可搭建全域统一的电子文件保存运行体系。跨业务系统衔接层面,全域部署档案数据中间转换网关,各类业务系统无需改版升级,通过标准化 API 接口接入网关即可完成数据对接,系统依托内置元数据映射模板,自动转换各类业务自定义字段为国家规范标准元数据,无法自动适配的字段通过专项补充采集机制完善。网关可批量完成文件格式转码工作,统一转换为 PDF/A、TIFF 等适配长期保存的合规格式,拦截加密文件、非标失效文件并生成专项整改清单。

全域部署统一数据交换总线,明确各业务系统传输时段与单次传输上限,依托加密传输通道规避卡顿、丢包、信息泄露等安全隐患。跨馆协同层面,搭建区域档案资源共享交互平台,统一省、市、县三级档案馆数据交互协议与移交数据包封装规范,固定文件本体、元数据清单、校验凭证、移交说明四大核心模块,接收系统自动完成合规性筛查,精准退回不合规数据包并标注问题明细,缩减人工核验工作成本。

标准落地层面,全面梳理各类系统现存的规范冲突,出台统一的电子文件长期保存实施细则,明确元数据必填字段、存储格式准入清单、介质更换周期、数据校验频次、馆际移交流程等硬性规范。针对存量档案开展标准化专项整改,分批次完成老旧文件格式转码、元数据补录、哈希值补算;建立季度常态化核查机制,以转码合格率、文件合规率、移交数据包达标率为核心考核指标,督促不合规单位限期整改并复核验收,全方位破除系统壁垒与标准差异,搭建规范统一、互通高效的电子文件长期保存运行基础^[3]。

3 数字档案馆电子文件长期保存创新实施策略

3.1 标准规范优化: 保存格式与元数据体系升级完善

现行电子文件保存标准存在更新滞后、细分场景覆盖不足的短板,无法适配海量异构数字资源的长期留存需求,格式管控与元数据治理是体系升级的核心方向。长期保存格式管理工作中,建立动态更新的格式准入清单,将 PDF/A、TIFF、SIG 等无损固化格式列为永久保存优选格式,针对图片、音视频、数据库等不同类型文件划定专属标准化归档格式,配套完善格式淘汰处置流程。行每两年开展一次文件格式存续风险研判,针对逐步淘汰的商用私有格式设置三年整改过渡期,分批次完成存量文件转码迁移,留存新旧版本对照记录,规避格式迭代造成的文件不可读风险。元数据体系优化聚焦全生命周期字段补全,在国家标准基础上新增载体溯源、存储校验、智能识别标签等延伸字段,划分业务原生元数据、档案馆保管元数据、利用行为元数据三大模块,明确各类字段采集主体、填报时限与校验规范。历史存量档案开展元数据补缺专项整治,系统批量抓取业务原始数据填充空白字段,人工补齐缺失的管护处置信息,搭建统一完整、可跨系统识别调用的元数据链条,从底层标准层面夯实电子文件长期保存基础。

3.2 技术架构革新: 智能赋能与敏捷数字仓储体系构建

人工智能与云存储技术可重构档案管护技术框架,兼顾档案运维效率与存储扩容弹性,适配电子文件长期保存的核心需求^[4]。智能技术落地应用覆盖多项核心实操场景,图像识别、自然语言处理算法可自动筛查归档文件完整性,批量识别残缺附件、空白无效文档并完成风险标记。机器学习模型可预判存储介质老化、数据衰减隐患,提前推送介质更换、异地备份调度工单。智能检索与修复模块可自动修复轻微缺损的文本、图

像类电子文件,降低人工修复工作投入。数字仓储体系采用分布式敏捷云架构,划分在线高速存储、近线低频存储、离线异地灾备三级存储池,高频查阅档案适配在线高性能存储资源,五年以上低频调取文件迁移至低成本近线存储,核心永久档案留存两份离线异地备份资源。系系统内置自动化数据迁移调度规则,依据文件调取频次、保管期限自动切换存储层级,无需人工批量转移归档;配套弹性扩容资源池,可根据年度档案新增体量动态调配存储算力,解决传统固定存储架构扩容滞后、资源闲置浪费的问题,智能技术与分层仓储的融合应用可全面提升长期保存技术支撑能力。

3.3 制度协同保障:法律治理、权责界定与成本分摊机制设计

完善的配套制度是电子文件长期保存稳定运行的关键,法律合规、权责划分、成本管控三类制度协同落地,可构建全方位保障体系。法律治理工作依托电子档案相关法律法规,梳理业务归档、异地移交、开放利用等环节的合规要点,制定馆内电子文件证据效力实操细则,明确电子签名、可信校验凭证的归档留存规范。法务部门协同档案团队定期排查涉密文件管护、线上档案对外开放查阅等场景的法律隐患,建立风险台账落实动态整改。权责管理采用分层分级管控模式,明确业务生成部门的源头归档主体责任、数字档案馆的保管校验主体责任、技术运维单位的存储维护主体责任,细化各岗位巡检核查、数据修复、到期鉴定等工作权责,实现档案损毁、数据缺失等问题的精准追责。成本分摊机制结合资源产出与使用主体划定承担比例,日常存储运维基础开支纳入档案馆固定经费,业务增量文件转码、存量文件标准化整改等专项开支由对应业务单位承担,跨区域馆际共享产生的存储扩容费用由参与机构按存

储容量分摊。年度保存成本核算公示制度可规范资金使用流程,提升经费利用效率,多维度制度协同可有效化解长期保存工作中的合规、管理与经费难题。

4 数字档案馆电子文件长期保存落地保障优化路径

长效保存机制与实施策略的落地质量,依托完善的配套保障体系支撑,人才培育、资源协同、风险防控是优化工作的核心抓手。针对数字档案管理复合型人才短缺问题,搭建分层常态化培育体系,定期开展档案业务、数据运维、法律合规交叉培训,联合技术服务商开展存储架构搭建、智能校验技术实操实训,落实持证上岗与绩效考评联动机制,补齐一线管理人员数字化管护能力短板。搭建区域档案资源协同共享机制,统筹辖区各单位存储算力、备份机房等硬件设施,联合开展存量档案标准化改造、异地介质备份等重点工作,减轻单馆独立运维压力。全流程风险预警处置体系聚焦存储介质、网络传输、数据访问等核心环节,常态化排查安全隐患,针对文件损毁、数据泄露、系统宕机等突发场景制定分级应急处置方案,每半年开展数据恢复与灾备切换实战演练,持续筑牢电子文件长期保存全链条运行根基^[5]。

5 结语

电子文件长期保存是数字档案馆建设核心难题,仅依靠单一技术或制度无法实现长效管护。本文构建涵盖理论动因、可信运行体系、创新实施策略、落地保障路径的完整框架,覆盖标准、技术、法律、人才多维度实操举措。后续可结合区域档案馆实践数据持续优化分层存储、成本分摊机制,进一步完善适配不同业务场景的电子文件长期保存实施方案。

参考文献:

- [1] 张珊瑜.高校档案数字资源长期保存策略探析——以中山大学档案馆的实践为例[J].兰台世界,2025(12):43-47.
- [2] 王宇哲,宋凌云.面向电子文件长期保存的云原生架构设计与关键技术[J].山西档案,2025(12):130-132.
- [3] 于飞.电子文件归档和电子档案管理探讨[J].IT 经理世界,2025,28(12):185-187.
- [4] 兰佳丽.数字时代传统档案向电子档案融合转型的研究[J].新潮电子,2025(11):16-18.
- [5] 吴雪清.企业电子文件归档与电子档案管理策略研究[J].企业改革与管理,2024(18):35-37.