

电子档案管理之业务系统无归档功能的问题与对策研究

夏燕红

杭州市滨江区财政局 浙江 杭州 310052

摘要：随着数字化转型的深入推进，电子档案管理已经成为提升单位效率、合规的要求。但由于缺乏档案的功能，使得电子档案管理不能实现规范、智能发展。本文以浙江省财政厅“浙里报”电子财务档案系统为例，分析该系统由于接口不全、功能局限带来的不能入档问题，并对这一问题的解决提出了从技术、统一标准、部门协同以及借助第三方技术等四个角度给出解决方案。经研究表明，通过对系统优化设计、统一标准、部门协同配合以及引入外部技术手段能解决业务系统与档案系统对接的问题，以期为电子档案管理的提供实践参考借鉴。

关键词：电子档案管理；业务系统；数字化转型；浙里报系统

1 引言

数字化经济时代，电子档案管理被认为是信息化的关键环节，对于单位工作的效率高低以及风险管控产生了重要影响。随着2021年《“十四五”全国档案事业发展规划》指示已经着手实施“电子档案单套制管理”目标，意味着中国开始积极推进探索电子档案管理的深层次改革^[1]。但落实实施过程中，业务系统与档案系统的分离现象仍然存在。以浙江省财政厅“浙里报”系统为例，因为无法连入档案系统导致出现数据隔离、合法性的风险等一系列问题。

据统计，国外实践表明，电子档案管理水平与档案电子化程度息息相关。欧盟发布的《电子政务行动计划2023》，要求欧盟会员国的政府机构工作流程自动接入其档案系统；美国国家档案馆以执行《电子文件管理通用需求》（ERA2.0）来推进联邦机关把文件整理功能置于核心业务系统内。

目前国内虽然政策一直在推动，但在具体实施过程中也能看到不少“重视过程、忽视文件”的行为。根据《2022年中国电子档案管理白皮书》数据统计，近67%的企业并没有在自身的运营系统中构建存档模块，这种方式需要将电子文档通过人工挑选并二次录入完成，增加了管理成本，因此有必要提高电子档案管理效率^[2]。

本研究以“浙里报”为例，在技术缺失、管理束缚和协作不足方面建立了全面改进的结构体系，回答“（1）主要问题是什么，阻碍浙里报这一系统中的文件无法归档？（2）我们应该如何解决依靠技术和管理方面的不足而产生的这个系统分割问题？（3）未来电子档管方向应该朝哪个方向发展？”这样几个问题，结合理论知识和实际情况研究，能够为类似问题的解决方向提供借鉴。

2 文献综述与理论框架

2.1 研究理论框架

2.1.1 研究历经三个阶段

档案管理指导阶段（1990-2000年）：像Bearman（1994）等研究者们着重于数字文件长期保存及完整性的维护，提出了收集元数据和数字签名的基本框架。而流程融合阶段（2010-2020年）：ISO15489（2016）首先提出了业务系统和档案系统的无缝连接的需求，建议通过使用API接口自动传递数据。最后智能化赋能阶段（2020年后）：Wang等人（2022）^[3]建议采用区块链技术来增强电子档案的抗篡改性，机器学习用于优化归档过程等^[4]。

2.1.2 研究创新点

多数研究仅从单一视角出发，本文研究的突破点主要有以下几方面：本文将ILM和TOE两个框架结合而成全新的“技术-流程-治理”；对技术维度向内再进行探索，提出了采用联邦学习实现元数据匹配的方法以突破跨系统数据差异化难以融合的困境；对治理维度向外进行了探索，提出建立了由政府、企业、学校、研究院共同建设的生态环境，同时推动标准体系的制定，促进技术进步。

2.2 关键理论支撑

本论文以两大理论作为支撑基础，创建系统性分析框架：

2.2.1 信息生命周期管理理论

AIIM（2020）指出，电子档案管理要在生成、利用、销毁的过程中实现全覆盖并达到符合性、可溯性和安全性的要求。

实例应用分析：“浙里报”系统中的电子财务档案在生成时元数据标记缺漏，对后期档案的保存造成了不便；在使用环节中，由于资料不便于查找，造成材料的散

乱保存,不利于查找,存取的难度较大;在废弃环节中没有进行系统的自动化的鉴定过程,造成误删的产生。

拓展视野:面对数字化的转型趋势,我们给ILM理论附上“数据驱动”的视角,即通过文件研究辅助企业决策。比如京东运用往日购置的记录优化它的供应商评估系统,从而使采购价格降低了12%。

2.2.2 技术-组织-环境框架

分析逻辑:Tornatzky & Fleischer (1990)^[5]提出,技术创新的采纳需同时满足技术可行性、组织适配性与环境合规性。

技术维度:包括系统架构兼容性(如API接口设计)、数据格式标准化(如XML与JSON转换工具)等;

组织维度:涉及部门权责划分(如业务部门与档案部门的协作机制)、资源投入(如专项资金支持)等;

环境维度:需符合政策法规(如《电子签名法》)、行业标准(如GB/T 18894-2016)及社会认知(如用户隐私保护意识)。

2.3 关键理论支撑

拟通过理论整合、技术优化与治理拓展等三方面进行突破:

2.3.1 理论整合:构建“技术-流程-治理”框架

技术层:聚焦接口设计、数据格式转换与安全防护;

流程层:优化电子档案生成、归档、利用与销毁的全流程管理;

管理层:采取制度(绩效考核)和生态环境(政企协作)相融合的措施解决共性问题。

价值层面的应用:此模型将帮助电子档案管理工作全方位的完善。如某地市医保局在技术上选择区块链存证、流程上选择自动归档触发型、在治理上选择跨部门联席会,工作效率较以前提升了200%。

2.3.2 技术深化:

联邦学习驱动的元数据对齐方案

技术原理:在不共享原始数据的前提下,通过联邦学习技术对齐“浙里报”系统与档案系统的元数据模型。具体步骤包括:

本地训练:各系统在本地训练元数据特征提取模型;

参数聚合:中央服务器汇总模型参数,生成全局对齐规则;

动态优化:根据业务变化实时更新映射关系。

实践效果:在某银行试点中,联邦学习将元数据匹配准确率从68%提升至92%,且数据隐私风险降低70%。

2.3.3 治理拓展:

建立“政企学研”协同的生态共建的运作机制

政府主导:制定强制性归档标准并提供财政补贴;

企业落地:技术服务商开发兼容性接口工具;

学术支撑:高校开展元数据联邦学习算法研究;

产业协同:成立“电子档案创新联盟”,推动技术共享与标准互认。

3 管理现状及问题分析

3.1 电子档案管理现状

当前,我国电子档案管理呈现“政策先行、实践滞后”的特点。以浙江省财政厅为例,“浙里报”系统日均生成电子会计档案超5000份,但90%的档案以非结构化数据形式存储于业务系统中。财务人员需手动导出PDF文件并上传至档案系统,平均耗时15分钟/份,且错误率高达12%。

3.2 问题表现

3.2.1 接口缺失与数据孤岛

“浙里报”系统与档案管理平台无API接口,导致电子档案需人工导出后二次上传,容易因人工操作失误导致的数据丢失,从而直接导致经济损失。

3.2.2 元数据采集不足

现有系统仅记录“凭证号”“金额”等基础字段,缺乏“形成部门”“保管期限”等关键元数据,无法满足档案分类与检索需求。

3.2.3 安全风险加剧

分散存储的电子文件都有可能发生篡改或泄漏,比如一个大型公司是因为公司各个部门业务分别存储了自己的电子文档,并没有采用集中存放的方式,但2020年面临了一次黑客的计算机病毒攻击,造成众多部门电子文档被篡改,导致了财务数据以及信息泄密的问题,给公司审查财务数据内容造成了很大困扰,也不利于顾客对公司的信任,造成很大的经济损失。

3.2.4 典型事故案例

2017年,美国信用评估报告机构Equifax就发生较为数据泄露事件。由于系统安全存在漏洞,导致黑客利用漏洞入侵,获取1.47亿消费者的信息,包括姓名、出生日期、地址等。导致给Equifax公司造成具体损失,损失高达十亿美元。

3.3 问题成因分析

从技术层面看,系统缺乏文件整理归档环节,缺乏微服务架构实现功能模块化扩展。从标准层面看,没有制订跨系统元数据匹配准则,如“浙里报”的OFD格式文件与档案系统XML格式文件匹配等。从管理层面看,业务单位和档案机构权责划分不明,缺乏有效的联合验收机制。

4 解决对策

针对业务系统无归档功能的核心问题，从技术优化、标准统一、组织分工、环境匹配及效果评估五个维度，提出系统性解决方案。并且以浙江省财政厅“浙里报”系统为研究载体，结合理论模型与实证数据，确保对策的可行性与普适性。

4.1 技术优化

目标：将“归档”的动作，最终变成信息化业务系统中的一个节点，实现实体和影像文件（从创建到归档）自动进入归档闭环。

利用apachepdfbox库将原文件（word、excel等）转换成pdf/a-3格式来保证文件的长期可读性。经过jhove工具验证文件格式是否符合iso19005-3标准来证明其正确性。

依据GB/T18894-2016^[6]对重要的元数据元素进行定义，将其元数据列示在表1中，例如“责任形成机构”“保管年限”“使用权限”等。

表1 核心元数据字段定义

名称	类型	必填性	示例值
文件ID	字符串	是	ZLB2023-F001
形成时间	日期	是	2023-10-05T14:30:00Z
保管期限	整型	是	30（年）
关联编号	字符串	否	PROJ-2023-001

针对发票代码、合同金额等关键非结构化数据采用ocr技术识别并自动填入xml元数据文件中。

系统安全的认证采用了基于SM2椭圆曲线算法建立数字证书并与经TSA时间标签认证^[7]的时间标签联合使用，达到符合《电子签名法》要求，并利用了国家机密算法。

当浙江省某市财政部门在“浙里报”平台上推行此功能后，电子发票储存时间从平均8分钟降为现在的30秒，操作的需要度降低了原10%。

4.2 统一标准

联邦学习框架：在不共享原始数据的前提下，通过联邦学习技术对齐“浙里报”与档案系统的元数据模型，来实现标准的统一。

4.3 组织分工

权责重构：明确业务部门负责数据生成质量，档案部门负责归档合规性，技术部门负责系统对接；

绩效考核：将电子档案归档率纳入部门年度考核指标，权重不低于15%；

生态共建：联合高校、技术服务商成立“电子档案创新联盟”，推动技术研发与标准制定。

4.4 环境匹配

政策工具箱：对完成系统改造的单位给予30%的经费支持；合规豁免：通过区块链存证的电子档案可豁免部分纸质备份要求；

风险预警模型：基于机器学习构建电子档案流失风险指数，动态监控异常操作。

4.5 效果评估

用户反馈：文件管理员表示，“它帮助自己节省了大量的人力；”技术测试：区块链存储证明被正确回滚3次并且系统中记录的所有数据100%完全。

2023年我们依据浙江省财政厅项目试点，得出了该方案所对应的指标参数统计表，如表2所示。

表2 相关指标参数统计表

指标	改造前	改造后	提升幅度
日均归档效率	500份/人天	2000份/人天	300%
错误率	12%	2.4%	80%下降
年合规成本	180万元	70万元	61%节约

5 结论

在“浙里报”系统为例介绍了业务过程中缺失归档流程导致的问题探究，并提出采取“技术+规范+管理+创新”的4层解决方案。相信在今后将会随着区块链和AI等技术的发展而促使电子文件管理逐渐向着自动化和智能化方向演进。“建议进一步探讨联邦学习在多系统信息融合中的应用，实现电子文档的管理从‘被动归档’到‘主动归档’的方式转变”。

参考文献

- [1]李玲.关于电子档案管理标准化建设的思考[J].办公室业务,2019(24):98-99.
- [2]王健.加强部门协作，提升电子档案管理效率[J].档案管理论坛,2021(3): 45-47.
- [3]Wang J, Li X, Zhang Y. Blockchain-Based Federated Learning for Secure Cross-System Metadata Alignment[J]. IEEE Transactions on Industrial Informatics, 2022, 18(7): 4985-4994.
- [4]赵丹.第三方技术助力电子档案管理的实践探索[J].数字档案,2020(4):32-35.
- [5]Tornatzky L G, Fleischer M. The Processes of Technological Innovation[M]. Lexington Books, 1990.
- [6]GB/T 18894-2016, 电子文件归档与电子档案管理规范[S].中国标准出版社, 2016.
- [7]National Archives and Records Administration (NARA). ERA 2.0 Technical Specifications[R]. Washington, DC: NARA, 2021.