

大数据环境下档案管理效率评价

文 | 刘岩

大数据技术在档案管理领域的广泛运用，推动档案工作从传统人工模式向信息化、智能化模式转变，当前档案管理效率评价缺少系统化的指标体系，限制了档案管理工作的持续改进，建立科学的档案管理效率评价指标体系并通过量化测评识别效率短板，对提升事业单位档案管理水平具有重要的实践意义。

大数据环境下档案管理效率评价概述

档案管理效率的内涵与评价维度

档案管理效率指在档案收集、整理、存储、检索以及利用等全流程中，单位时间完成的工作量和所消耗资源之间的比值关系，其核心在于衡量档案管理活动的投入产出水平。从评价的维度方面来看，档案管理效率包含时间维度、质量维度和成本维度三个层面，时间维度更侧重于档案从生成到归档、从检索到调取的时效性指标，质量维度关注档案数据的完整性、准确性和可用的水平情况，成本维度则聚焦于人力、物力以及技术资源配置方面的合理程度。在大数据环境下，海量档案数据的涌入让传统的效率评价标准面临着挑战，数据处理能力、系统响应速度和信息挖掘深度成为评价档案管理效率的新兴指标，这些指标的量化测度需要建立在对档案业务流程深度理解和对大数据技术特征准确把握的基础上。

大数据技术对档案管理效率的影响机制

大数据技术通过改变档案数据的采集方式、存储模式及处理流程，对档案管理效率产生了系统性影响。在数据采集环节，大数据技术借助自动化采集设备和智能识别算法，将传统的人工录入模式转变为批量化实时化数据捕获机制，显著降低档案收集时间成本并拓展数据来源广度。在存储环节，分布式存储技术和云计算平台突破传统档案库房的物理空间限制，让海量档案数据能在多节点间分散存储，增强存储容量弹性扩展能力和数据冗余备份机制以提升档案保存安全性。在检索利用环节，基于机器学习的智能检索算法能理解用户语义需求，通过关联分析和相似度计算实现精准推送，将档案查询响应时间从传统模式下的分钟级缩短至秒级，大幅提升用户获取所需档案的效率。大数据技术还通过数据挖掘和可视化分析，提取隐藏在档案资料中的价值信息为决策提供数据支撑，显著改善档案利用的深度和广度。

大数据环境下档案管理效率评价指标体系构建

档案管理效率评价指标的设计原则

构建科学合理的档案管理效率评价指标体系需遵循系统性、可测性和适用性三大原则。系统性原则要求指标体系覆盖档案管理全业务流程，在数据采集、分类归档、存储备份到检索利用各环节设置相应的效率测度指标，各指标相互独立且形成有机关联以全面反映档案管理整体效率水平。可测性原则强调指标需具备可量化、可操作的特征，指标数据应通过档案管理系统日志记录、统计报表或专项测试获取，避免使用难以量化的模糊性描述。适用性原则关注指标体系与实际业务场景的契合度，不同类型和不同规模的事业单位档案管理模式存在差异，指标体系设计应充分考虑这一差异性，通过设置核心指标和可选指标的组合方式，让评价体系既具通用性又保持灵活性，确保指标能真实反映各单位档案管理的实际效率。

档案数据采集与处理效率评价指标

档案数据采集与处理效率的评价需要从时效性和质量两个维度建立指标，如表 1 所示，时效性指标通过测量数据采集响应时间和处理吞吐量来反映系统的处理速度，质量指标则通过数据完整率和元数据准确率来衡量采集处理的精准程度，这些指标的设置覆盖了档案数据采集与处理的关键控制点，能够全面反映该环节的效率状况。

档案信息检索与利用效率评价指标

档案信息检索与利用效率的评价应当围绕检索性能和服务效果两个层面构建指标体系，检索性能层面侧重于测量

表 1 档案数据采集与处理效率评价指标

指标类别	指标名称	计算方法
时效性	数据采集响应时间	从档案生成到系统采集完成的平均时长
	数据处理吞吐量	单位时间内完成分类归档的档案数量
质量性	数据采集完整率	实际采集档案项数 / 应采集档案项数 × 100%
	元数据标引准确率	准确标引的元数据字段数 / 总元数据字段数 × 100%

来源：宽甸满族自治县中心医院

系统对用户查询请求的响应能力，检索响应时间直接影响用户的查询体验，检索结果准确率则反映系统是否能够精准理解用户需求并返回相关档案资料。服务效果层面关注档案资源的实际利用情况，档案调取成功率体现系统运行的稳定性，档案利用频次则揭示档案资源的活跃程度，如表 2 所示。

表 2 档案信息检索与利用效率评价指标

指标类别	指标名称	计算方法
检索性能	检索响应时间	用户提交检索请求到返回结果的平均时长
	检索结果准确率	相关档案数量 / 返回档案总数 × 100%
服务效果	档案调取成功率	成功调取档案次数 / 总调取请求次数 × 100%
	档案利用频次	月度档案调阅总次数 / 馆藏档案总量 × 100%

来源：宽甸满族自治县中心医院

大数据环境下档案管理效率评价的应用分析

案例概况

选择某市级事业单位为评价对象，该单位档案馆藏数量大约有 12 万卷，平均每年新增档案 8000 余卷，档案类型涉及行政管理、项目档案、人事档案等多种类型。该单位在 2022 年开始进行档案管理信息化建设，引入大数据平台和智能档案管理系统，完成了存量档案的数字化转换工作，数字化率达到 85% 以上。档案管理部门配备了 5 名专职人员，其中 3 名已经接受大数据技术应用培训，系统平均每天处理档案查询请求约 200 次，每月平均档案调阅量超过 3000 卷次。该单位档案管理工作基本实现从传统人工模式向信息化模式的转型，但在数据采集自动化程度、检索响应速度和档案价值挖掘深度等方面还有提升空间，运用构建的评价指标体系对其档案管理效率进行系统测评，能够准确识别效率短板，为后续改进工作提供方向指引。

效率评价指标的测算与分析

基于前文建立的评价指标体系，对案例单位档案数据采集与处理环节以及信息检索与利用环节的效率水平进行量化测评，测评数据来源于该单位档案管理系统的运行日志、统计报表和现场测试记录，测评结果如表 3 所示。

测评结果表明，案例单位档案管理效率整体处于良好水平，档案调取成功率达到优秀标准，表明系统稳定性较好，但档案利用频次偏低，仅为 2.1%，说明档案资源的开发利用深度不足，大量档案处于闲置状态，未能充分发挥其决策支持和信息服务功能。

档案管理效率优化建议

针对案例单位档案管理效率评价中发现问题，现提出如下优化建议。在数据采集环节，应缩短采集响应时间，

表 3 案例单位档案管理效率评价结果

评价环节	指标名称	测算值	评价等级
数据采集与处理	数据采集响应时间	48 小时	良好
	数据处理吞吐量	650 条 / 小时	良好
	数据采集完整率	96.5%	良好
	元数据标引准确率	97.8%	良好
信息检索与利用	检索响应时间	4.2 秒	良好
	检索结果准确率	92.3%	良好
	档案调取成功率	98.6%	优秀
	档案利用频次	2.1%	待改进

来源：宽甸满族自治县中心医院

借助与业务系统建立实时数据接口，将档案采集响应时间从 48 小时压缩到 24 小时以内，以此实现档案数据的即时捕获。在检索性能方面，需对检索算法进行优化，引入倒排索引和缓存机制，将检索响应时间控制在 3 秒以内，进而提升用户的查询体验。在档案利用环节，应加强档案价值挖掘工作，利用数据分析技术对档案内容做主题提取和关联分析，建立档案专题数据库，通过主动推送服务提高档案利用频次，将档案利用率提升至 5% 以上。管理人员需定期开展效率评价工作，将评价结果纳入绩效考核体系，形成效率改进的长效机制，通过持续的监测 - 反馈 - 改进循环推动档案管理效率的稳步提升。

结束语

构建档案管理效率评价指标体系应遵循系统性、可测性和适用性原则，从数据采集与处理、信息检索与利用这两个核心环节设置时效性、质量性和服务效果等维度的量化指标。对某市级事业单位开展实证评价后发现，该单位档案管理效率整体处于良好水平，不过档案利用频次偏低，档案资源的价值挖掘深度有待提升。评价指标体系为识别档案管理效率的薄弱环节提供了工具支撑，为制定针对性的改进措施提供了数据依据，对推动档案管理工作的科学化发展具有参考价值。

作者简介：刘岩 宽甸满族自治县中心医院

责任编辑：杨佳宇 投稿邮箱：zhouhl@staff.ccidnet.com