 DOI: https://doi.org/10.12349/iser.v6i4.5303

Quality Control and Security Confidentiality System for Digitalization of Archives

Hua Li

Natural Resources and Planning Bureau of Changyuan City, Changyuan, Henan, 453400, China

Abstract

Archives are important carriers of a country's cultural heritage and historical memory, and preserving and inheriting archives is of great significance for social development and cultural inheritance. In today's rapidly developing information technology, digitizing archives is a development trend. This system can not only greatly improve the efficiency of using literature, but also provide convenient references for users to extend the storage period of literature. However, in the process of digitizing electronic documents, ensuring the quality and security of files is a key issue. If the quality of these data is problematic, then these data cannot truly reflect their original appearance and there may be security risks. This article provides a comprehensive analysis of the quality control and security confidentiality system of archive digitization, hoping to provide suggestions or inspiration for relevant workers.

Keywords

digitalization of archives; Quality Control; Security and confidentiality system

档案数字化的质量控制与安全保密体系

李华

长垣市自然资源和规划局，中国 · 河南 长垣 453400

摘　要

档案是一国文化遗产和历史记忆的重要载体，保存和传承档案对于社会发展和文化传承具有重要意义。在信息化快速发展的今天，对档案进行数字化是一种发展趋势。该系统不但可以大大地提升文献的使用效能，而且可以方便地为使用者提供方便的参考从而达到延长文献保管年限的目的。但是，在实现电子文档的数字化过程中，如何保证文件的质量和安全性是一个关键问题。如果这些数据的品质有问题，那么这些数据就不能真实地体现这些数据的本来面貌，还会存在安全隐患。本文就档案数字化的质量控制与安全保密体系进行综合分析，希望以此可以给广大相关工作者以建议或启发。

关键词

档案数字化；质量控制；安全保密体系

1 引言

针对目前随着时间的推移，由于各种原因，传统的纸质文件容易受到外界的影响而出现老化、破损等情况，从而加大了文件档案遗失的风险。随着社会经济的发展，社会对档案资料的使用越来越多，而传统的检索方法由于其自身的局限性，已不能很好地适应现代社会对信息的迅速获取。正是在这样的环境下，才有了档案的数字化。通过对纸质文件进行数字化处理，可以有效地缓解文件的保管困难，增强文件的安全和稳定，同时也可以通过互联网的方式进行方便地分享，从而提高文件的使用效益。

【作者简介】李华（1980–），女，中国河南长垣人，本科，档案馆员，从事档案信息收集研究。

2 档案数字化质量控制与安全保密的现状

随着大数据对信息的处理，通过对大量文档资料的深入分析，可以在杂乱无章的文档中提炼出有价值的内容。比如对公司的存档，通过对大数据的分析，可以准确地找到与行业发展方向有关的重要资讯，帮助企业制订准确的发展策略。不仅如此，云计算的发展，可以建立一个高效率、低开销的文档数据储存和运营平台成为可能。它具有可伸缩的储存与运算资源，可以随着档案资料的数量而进行灵活地调节，在保证系统运作效率的同时，不会造成任何资源的浪费。由于传统的档案管理，面对海量的文件不但占据了珍贵的物理空间，比如对于一个中等规模的公司来说，长期积累下来的文件往往要租用一栋写字楼才能保存下来，这就增加了企业的管理与维修费用。那么采用手工查阅的方法，既费时又低效，还存在一定的人为遗漏。就文件使用而言，由于地域、时间等因素的影响，使用者很难在所需的任何时候、任何地点都能方便地获得自己想要的文件资料，这极大地阻碍了档

案的价值。同时，各行各业都在加速进行数字化变革，不管是企业出于支持战略决策、保证合规运作，又或者是研究单位出于对知识的继承和创新发展的需要，都迫切需要对文档数据进行深入的发掘和广泛的共享 [1]。

其中腾讯是一个很好的例子，它的产业覆盖了社交，游戏，金融科技，数字内容等各个行业，并且表现出了快速增长和多样化发展的势头。在公司的成长历程中，各种文件，如合同，文档，邮件，工程文件等，都是公司的重要组成部分。根据权威的专家调查，腾讯网站一年增加了好几个亿的用户记录。面对这样的海量信息，若仍沿用传统的纸张文件管理模式，单就其所占的面积而言，就腾讯目前所拥有的办公面积而言，其所产生的纸张文件将会占用上千平方米的面积，这是一种极大的浪费。此外，文件查询的效率非常低，每一份文件的查询都需要几个小时或者几个星期，这一低下的工作效率已经成为公司内各个部门之间协同工作的瓶颈，妨碍了公司的正常运作，从而大大降低了公司的运作效率和决策的反应能力。腾讯在这一背景下，积极推进“数字档案馆”的发展，把它作为提高企业核心竞争能力的一项重大战略措施 [2]。

3 档案数字化的质量控制与安全保密体系构建

3.1 数字化前档案质量的控制与风险评估

在对电子文档进行数字化之前，要对其进行品质管理和风险评价，制定一套专业化和精细化的管理战略是非常必要的。根据不同的纸质材料，利用酸碱滴定法和脆度仪精确测定不同种类的纸质材料的酸碱性和脆性。比如，酸性纸在贮存过程中容易发生黄色和脆性的现象，因此，在 pH 值小于 6 的情况下，应进行适当的预处理。在此基础上，建立相应的分类方法，针对较易碎的造纸材料，将其进行脱酸强化，即使用二乙基锌等药剂与造纸原料中的酸作用，使其在造纸过程中的化学结构得到有效的改善。如果是重度脆裂的纸，在进行数字化处理之前，需要进行修补。损坏的文件修补验收程序一定要严格，使用修补前、修补后的对比记录表，通过影像解析软件对修补前和修补后的文件图片进行像素级别的比对，从而实现对修补效果的定量评价。如果一份历史文件在修补之前损坏面积达到 20%，修补后减少到 5%，则判定通过 [3]。

针对特种载体的数字预处理技术标准对微型电影的制作提出了苛刻的标准，在微型电影中，需要使用专门的微型电影扫描仪，根据不同的银盐和染料的成像机理来调节扫描的参数，比如银盐电影，可以使用高分辨率、低噪点的扫描方式；蜡印利用 3D 扫描技术，建立一个完整的立体模型，精确地捕获了每一个细微之处。在选择数字器件时，对于扫描装置来说，需要有一个分辨率的动态调整规范。在文档文件中，300 dpi 的分辨率通常就能达到要求，图案型文件由于其线条和尺寸标注等详细信息，要求提高到 600 dpi 以上，图片级文件为了减少颜色和细节，可以设定 1200 dpi. 采用 ICC 模板创建颜色管理闭环体系，采用光谱仪标定扫描装置的颜色，对比标准颜色空间，产生 ICC 模板并进行校验，保证颜色还原精度。还可以利用密度仪测定其密度，并与参考值进行对比，在误差大于 ±0.05 的情况下，进行仪器的维修和标定，保证仪器的正常工作，为高品质的档案数字化工作打下良好的基础。

3.2 数字化过程安全的控制链构建：物理环境、数据流加密

在建立数字文件的质量管理和安全保障系统中，对数字流程的安全监控环节是至关重要的。实体环境的安保系统采用三区两路的物理隔离方式，将预处理区、扫描区和存储区分隔开来，阻止非许可的人与装置自由出入，避免了可能出现的安全风险。以一家大型企业的电子文档管理系统为例，应用本系统后，系统的安全性事故发生率下降 80% 以上。通过人脸识别和指纹识别等高精度的人脸识别，实现了对进出的严格控制，准确率在 0.01% 以下。另外，该跟踪技术还可以将行人的运动轨迹全部记录下来，为以后的跟踪提供了方便。通过对 USB 接口的白名单控制，只有在白名单上的 USB 设备才能访问，才能保护硬件和软件的安全。

对网络中的数据流进行加密传送是网络的一个关键步骤。热点文件夹的自动密码传送是基于 AES-256 的即时密码算法，它是当前业内公认的一种高密度密码，能够在传送过程中实现对 256 个比特密钥的即时加密，非常难以通过蛮力破解。针对这一问题，本项目拟采用基于时间标记和操作者 ID 的隐藏性方法，在不破坏用户隐私的前提下，实现对用户隐私的有效保护。例如，在一个电子政务文档的电子文档中，就出现了一次信息泄露的情况，并利用该方法对用户进行源数据水印的定位。采用 CRC32 检验代码的断点续传检验机构，实现了在数据传送中断后，对已经传送的与要传送的数据进行检测，以保证数据的完整性。CRC32 检验代码能够对数据传递过程中出现的差错进行检测，差错检测率达到 99.99%[4]。

3.3 后处理质量检测体系，数字化成果的三级检验与安全审计

在电子文档数字化后的质量检验系统中，以“三级”为中心，对其进行质量控制。利用先进的图像倾斜、黑边和噪声自动检测方法，为电子文档的数字化工作奠定了基础。以霍夫变换为例子，利用笛卡尔坐标与 Hough 空间之间的关系，利用霍夫空间中的波峰探测信息，实现对影像倾角的精确辨识。在实践中，通过对多个样品的实验表明，所提出的方法可以使测量结果的精度保持在 ±0.5 摄氏度以内，从而保证了图像的垂直度和清晰度。利用传统 Canny 算子等边界探测算子，对图像进行平滑处理、梯度幅度和方向计算、非最大值压制边界、双门限识别和连通边界，从而使图像中的黑色边界保持在 1 个像素以下，防止黑色边界对文件的

显示产生影响。采用高斯滤波器和小波分析相结合的方法，以减少高频噪声的干扰，可以保证像素的浓度小于 0.01 个 / cm，从而提高成像的分辨率。

手工复验实行双盲法抽样，对抽样的样品进行仔细检验。根据统计原则，选取了 500 个样品作为研究对象，在满足测试速度的前提下，确保测试的精度。根据抽样结果，将各种疵病进行了分类和统计，包括图像模糊、文字变形和内容缺失等几种常用的疵病类型。根据图像的灰度信息及边界清楚指数对图像进行定量评价，汉字变形是由汉字辨识算法的精度起伏和汉字特性的改变来判定的，其中“专家终检”采用像 X-Riti1Pro 等颜色还原性的专业评定仪器，按照 CIELAB 的颜色空间规范进行工作。CIELAB 彩色空间是一个统一的彩色空间，它可以更精确地刻画人类对彩色图像的感受。X-Riti1Pro 型光学仪器，将样品的光谱反射率测定后，转化成 CIELAB 颜色空间上的坐标，并与标准色数值进行比较，使颜色偏差小于 2，保证了文件影像的颜色与原件文件的颜色保持高度的一致性。

3.4 安全生命周期的安全防护，储存介质与访问控制模型

在整个数字档案的整个寿命过程中，储存媒体的安全性是保障其安全性的基础。三个拷贝的脱机战略，利用异构媒体相结合的方法，构筑了一道强有力的备份防御体系。由于磁盘具有快速的读写速度和方便的存取数据的特点，可以完成对数据的多次复制的存储工作，特别是 LTO 型带，由于其高容量、低廉的存贮费用，已被广泛应用于大容量信息的存贮中。以 LTO-8 为例，其单片可达到 12 TB 的容量，在大型文件的数字化储存领域具有明显的优越性。LTO（一写多读）带的 WORM 将硬件级的写保护和与之相匹配的软件的固定逻辑相结合，只要记录下来，就不能用一般方法更改，从根本上避免了对文件的恶意篡改，保证了文件的原始与真实。在储存媒体即将被淘汰的情况下，必须有一套严谨的退磁检验程序。利用高斯计精确探测介质残余磁场，根据国家对磁记录媒体的严苛国家标准，要求其在退役后残余场强小于 320 A/m。达到这个要求，可以保证由于电磁干扰导致的退役媒体资料不能被还原，防止由于传输媒体而引起的信息泄露，从而保证文件的安全 [5]。

针对动态访问控制是文档数据存取的保卫者，通过对用户、资源和环境的多维特性进行深入分析，实现动态精准的权限管理。其中，使用者的特性包含了部门归属、职位等级、安全等级等重要的资讯，而资源的特性包含了档案的机密程度、所属类别等重要的元素。例如，在一家大型企业中，根据其所在的行业特性，该公司的工作人员可以根据其岗位职权与资源类型的匹配情况，在工作期间，由该系统对该部门的有关研发文件进行存取，但是在非上班或在外地存取时，该软件会根据不同的情况，启动附加的审核程序，以保证数据存取的安全性。该算法利用了传统的规则对特定的内容进行快速的特征匹配，从而可以对特定的内容进行有效地调整。可以对档案读取者的身份证号码进行精确的识别，该方法在保障企业运行过程中对数据的形式及局部信息的要求的同时，也能巧妙地掩盖用户的隐私。通过对不同类型的数据进行长时间的监控和分析，建立符合使用者日常使用习惯的行为基准，从而为整个档案的整个寿命周期提供一张严格的安全网。

4 结语

为了保证档案信息的完整性、可靠性和安全性，必须建立起一套完整的档案管理系统。在品质管理上，从数据收集的初期，到中间的后期品质检验，每一个步骤都是环环相扣，以保证数据的准确性和适用性。在安全与保密系统中，从存储媒体的管理、访问控制到运行过程的监控等方面，建立起一道完整的安全防线。这两个系统是相互补充的，在对电子文档进行数字化之前，要对其进行品质管理和风险评价，制定一套专业化和精细化的管理战略是非常必要的，并且数字化过程安全的控制链构建：物理环境、数据流加密，后处理质量检测体系，数字化成果的三级检验与安全审计更可以保证档案的质量。

参考文献

[1] 欧阳盛麟. 纸质档案数字化加工管理存在的风险及防范对策[J]. 参考, 2025, (06): 113-115.

[2] 邹新根. 档案数字化进程中的纸质材料保护策略创新[J]. 造纸信息, 2024, (11): 130-132.

[3] 何畏, 左词平. 传统载体声像档案数字化关键技术运用和质量控制体系研究[J]. 中国档案, 2024, (08): 64-65.

[4] 郭永强. 纸质档案数字化加工管理存在的风险和防范对策[J]. 华东纸业, 2024, 54 (08): 84-86.

[5] 安美娥. 浅谈乡镇档案数字化扫描质量控制——以曹县曹城街道办事处为例[J]. 办公室业务, 2024, (15): 151-153.