

# 科技档案数字化发展与数字化实践分析

张 华

重庆建峰工业集团有限公司 重庆 408601

**摘 要：**数字化、网络化和信息共享是现代信息科学技术进步的重要组成部分。纸质档案的数字化是将档案文件转换为传统格式的过程。换句话说，它将以纸为载体的档案文件形式转化为以磁盘、磁带、光盘等为载体的计算机文件。档案数字化也是当今电子政务、电子商务发展也至关重要。本文就科技档案的特点和科技档案数字化的发展方向，然后突出了科技档案数字化的实际问题，最后确定了科技档案数字化的实践。

**关键词：**科技档案；数字化；实践

## Analysis on development and practice of digitization of sci-tech archives

Hua Zhang

Chongqing Jianfeng Industrial Group Co., LTD., Chongqing 408601

**Abstract:** Digitization, networking and information sharing are important parts of the progress of modern information science and technology. Digitalization of paper archives is the process of converting archival documents into traditional formats. In other words, it will take the paper as the carrier of archival file form into disk, tape, CD, and so on as the carrier of computer files. Archives digitization is also the development of electronic government affairs and e-commerce is also crucial. In this paper, the characteristics of scientific and technological archives and the development direction of scientific and technological archives digitalization, and then highlight the practical problems of scientific and technological archives digitalization, finally determine the practice of scientific and technological archives digitalization.

**Keywords:** Scientific and technological archives; Digital; Practice

随着计算机技术的普及和互联网技术的飞速发展，人们获取知识和传播信息的速度越来越快。由于电子书刊的普及和数字图书馆的创建，人们已成为知识和信息的渠道。数字技术的飞速发展，使人们在信息领域的思维和认知发生了巨大的变化。数字技术可以转换、存储、传输和管理各种媒体和格式的信息记录，包括文本、照片、音频、视频内容及其使用。相反，档案信息的数字化还没有开始，业务标准和规范正在研究和完善中，在数字图书馆的创建上还远远落后。提高档案馆信息价值，发挥档案馆管理作用，需要开拓思路，积极创新，主动借鉴成熟的高新技术成果，档案馆必须对信息进行探索。数字控制不仅保护了原始文件，而且促进了文件信息的高速传输。

### 一、科技档案数字化特点

对于大多数生产企业来说，纸质持有人仍然是目前

科技档案管理的主要形式。纸质的科技档案具有原始性、完整性、可比性强等特点。此外，纸科技可以在环境保护的影响下储存。此外，它不耐用且使用不方便。科技记录是企业保存的有关生产技术、技术改进、资本积累等关系的原始记录。他们具有强烈的敬业精神、卓越意识和效率意识，具有公司有形资产和重大资产的特点。随着信息技术的飞速发展，企业产生的科技文章和技术记录的管理是科学规范的，但出于安全等原因，目前纸质地图记录仍覆盖半壁江山。就我们研究院的资档案而言，就保存着公司几乎所有的纸质底图档案，包含了业务发展过程中的所有原始结果和技术资源。

### 二、科技档案数字化存在的问题

#### 1. 工作人员整体素质不高

人是档案管理中最重要元素，档案管理信息的查找和管理离不开人。信息档案的管理对档案人员提出了

很高的要求。它不仅是要了解档案的传统职业，还要掌握自动化和网络技术，以及成为信息和知识的浏览器。现有档案设备的整体质量普遍较差，不适应档案信息化建设和发展的需要。

## 2. 科技档案纸制档案管理模式普遍

目前，人们还没有放弃传统的纸质档案管理系统。我国大部分档案馆仍采用传统的纸柜存储和目录管理方式。它不仅占用大量的存储空间，而且需要很高的管理和存储成本。更重要的是，查阅和使用非常不方便。一些设备使用计算管理，但大多数基于文件级管理。文件本身仍然以传统格式（例如纸质）存在，并且需要存储以用于管理、研究和使用的目的。换句话说，机器可读文件列表不是档案，它是一个简单的恢复工具，可以快速恢复现有档案，用户无法从档案中获取全文。因此，虽然建立可机读的档案目录是档案数字化的重要一步，但不能说这是档案间数字化的成就。

## 3. 科技档案数字化档案归档没有实现

电子政务或电子商务所产生的电子文件大多不转换成数字文件，不存储在计算机系统或磁盘、光盘上，而是直接集成到纸上，作为纸质文件存储。媒体许多档案馆都有本地计算机网络，但信息的处理和交换发生在档案馆内。但是，对于在使用工作或办公自动化生成文件的设备上创建的电子文档，文件管理没有实现，并且可以数字化转换为自动收集、鉴定、编目、标引和检索，转化为数字化档案归档。

## 4. 科技档案数字化建设存在盲目性

如果要从科学的角度对科技记录进行科技档案数字化建设，首先要从基本的角度正确理解，才能消除盲目性，避免不必要的金钱、人力、物力和时间的浪费。科技档案数字化在我国是一个新课题，还没有形成一套比较完整的原则。部分设备因缺乏实施规则、标准、指标体系和部分设备准备不足而过热。但是，提高档案作业的“现代化”“信息化”，进一步超过纸质档案的完成速度、规模和数量，可能存在隐患。从事实来看，不少单位数据库建设质量不高，检索信息的方法都是一样的，而且建立在很多设备上的数据库质量不高，因为媒体是分散的，信息的产生是独立的<sup>[1]</sup>。

科技档案数字化是一项高风险的投资，需要付出大量的努力和大量的资金。为了有意义地“使用”和“管理”档案信息，您需要关注您的贡献的影响。档案服务使用的信息资源有限且难以访问。随着信息根据当前情况演变，许多设备在计划的可行性和预期影响不足的情

况下迅速推出。没有成长空间，时间久了就会出现延迟，这实在是浪费时间。

## 5. 科技档案数字化的信息安全

信息网络正在向各个领域渗透，网络环境影响着档案的保密性和档案的安全性。一些人缺乏计算机和网络安全知识和安全意识、责任心不强、操作异常、及时备份数据等，这些人为因素是信息系统安全的最大隐患。二是技术问题，当前电脑病毒“瘟疫”横行，加上黑客的恶意攻击，使得档案信息的安全和保密受到威胁，势必引起我们的关注。

## 三、科技档案数字化实践

### 1. 科技档案数字化人才的培养

科技档案开发数字化人才可以从以下几个方面入手：一是全程支持现有人才，建立激励机制，发展档案技能，积累经验，鼓励主动档案变革管理。第二，培养才能。今天，许多档案管理部门人满为患。档案数字化的引入是提供与您的人才相关的专业培训，增加培训强度的最直接方式。第三，引进外部数字化专业人才。在招聘时，我们根据实际需要签订长期、中期和短期合同。特别是在档案数字化发展中，档案工作者要熟悉各类档案馆的数字化政策、法规、制度和标准，有效利用信息技术对档案信息进行管理和处理。您必须提供准确及时的记录。借助信息服务和馆藏，您可以利用信息技术来教育用户如何快速有效地使用信息资源。档案用户则应当有自觉利用档案信息资源的意识，并逐步提高他们通过信息技术检索和使用档案信息的能力。

### 2. 科技档案的前期处理工作

档案全文数字化实施的一个重要步骤是档案的初级处理：图形和表格的重组、打字、图形图表的扫描、光盘刻录拷贝等工作。科技档案的初步处理主要包括会计资料的编制，如科技档案的组织、文献页码的汇总、每篇文献的编号、每篇文献的描述、索引等。档案预处理是计算机文件管理的基础，在实际工作中我认为它是一个系统项目。您必须从科学家委托的档案开始，根据文件目录的细节和描述的设计和内容的，准备必要的信息并计划与索引条目对应的数据条目。至此，电脑应该可以代替说明书进行高级智能恢复使用，恢复科技档案。此外，文件预处理也很困难。为了在档案管理中发挥更好的作用，计算机可以做的就是托管高质量的、经过预处理的档案数据。在科技档案全文预处理应考虑以下问题：

(1) 对于科研类案卷，卷内中的文件和资料的组织

包括项目准备和创建阶段、研究和实验阶段、综合和评估阶段、成果申报奖励和推广应用等阶段排列。对于管理知识库，定量文档按主题、时间或重要性进行组织。此外，科技文献资料的编排量大，应该有初文，在初译后的外文解释和资料的基础上再翻译，最后是原文。打印通常放在案卷的末尾。

(2) 卷内文件的页码按照文件的序号排列顺序进行分组，科学技术地固定在文件的右下角。第1页：双面书写的科技文件材料，页码位于正面的右下角和背面的左下角，图纸中的页码写在标题栏外。

(3) 在对科技文件和档案进行描述和索引时，根据文件的内容和管理者或检索者可能的检索要求而设定的。在描述和索引档案文件时，它们应该基于档案描述和索引操作的标准，例如《档案著录规则》DA/T 18-1999。《档案分类标引规则》GB/15418-1994-《档案主题标引规则》DA/T 19-1999等规范标准。规范著录项目、统一著录格式、建立标准检索语言。

### 3. 科技档案的扫描工作

目前，最常用和最有效的文件检查方法是扫描法。扫描方法是將原始档案文件放在扫描仪平板上。使用扫描仪将文档文件转换为数字图形输入计算机·经文字识别(OCR)、校对后，在扫描过程中文本识别的准确速度直接关系到扫描的效率。通过调整扫描仪的分辨率、亮度和其他设置，您可以保持扫描仪OCR的高精度，减少返工并提高扫描研究人员和文件的效率。

**分辨率参数：**调整扫描图像的清晰度。分辨率作为一个单位是缩进(dpi)。扫描仪支持的饱和度越高，图像就越清晰。在实践中，我发现文本越小，分辨率应该越高。相反，如果文本很大，则应降低分辨率。分辨率设置是用户根据所使用的扫描仪的字体和功能定义的。一般常用文档的最小号字是小5号字或s号字，推荐分辨率分别为400.300，理论上分辨率越高，扫描图像的质量越高，但对文本的识别精度要求也越高。更高的分辨率会大大增加数据量并显著降低扫描速度。因此，应根据适用的原则来选择扫描分辨率。

**亮度设置：**这是重要的扫描仪设置之一，也是影响OCR文本识别性能的一个非常重要的因素。度设置指定扫描图像的亮度，调整到适当亮度的检测系统可以提供完美的检测速度。如果光线太亮，文字笔画残缺不全，光线昏暗，文字笔划粘在一起，形成黑色块状物。在这两种情况下，都无法完整记录输入功率。因此，亮度调节需要高分辨率。

**扫描速度：**这是决定扫描仪效率的重要扫描仪指标。尤其是在文本识别应用中，设备的速度为每秒1英寸。如果要处理大量文本字段，则可以在特定时间选择它们。扫描仪，扫描仪速度很高，在科技档案全文数字化过程中，应考虑以下因素：

(1) 扫描纸质较浅的文字包括文件色白、纸张发亮以及文字比划细的文字时，则需要降低亮度设置。

(2) 扫描小字符时，亮度设置要相应提高，但要限制在虚线不要太多。

(3) 扫描黑色文档背景或粗体文本等深色文档时，增加亮度设置。

(4) 在实际工作中，由于报纸的底色较深。渐变和斜体为粗体。因此，如果检测到这样的文档，则可以相应地增加亮度设置。而由于书刊、杂志的底色较浅，宋体字仿宋字的比划粗细适中。因此，当检测到这些文件时，您可以调整亮度设置以使它们达到适当的亮度。

### 4. 科技档案数字化信息光盘的制作

科技档案预处理完成后，即可轻松制作光盘。获取《GB/T 17678.2-1999cAD电子文档光盘文件存储、归档与管理要求：光盘信息》，同时将文件管理软件、文件数据库、文件存储文件夹及相关文件写入光盘。

### 5. 提供科技档案数字化的信息网络利用

科技档案管理系统创建了一个需要访问科技记录的用户数据库。为每个用户设置不同的权限，每个用户都有一个适当的协议来监控科学文件的使用。我们建议您为下载和打印科技文档设置适当的权限，以确保隐私和安全。也就是说，普通用户不能随意下载或打印科技资料<sup>[2]</sup>。

### 6. 探求更优的科技档案保密和档案安全解决方法

科技档案管理独特的管理方式规定档案馆应尽可能频繁、快速地交换档案资源。然而，共享和安全是难以调和的矛盾。为此，档案部门必须依据《中华人民共和国档案法》、《中华人民共和国国家秘密保护法》、《计算机病毒防治管理办法》等有关规定，制定严格的规章制度。它管理法律等，并包括设备的实际内存管理功能。组成计算机网络的系统管理文件信息安全，以确保系统正确部署和负责。识别档案信息，评估档案隐私和公开性，保持严格的安全意识，有效防止人为伤害，注意防毒杀毒，确保档案信息的安全和完整。

## 四、科技档案数字化的发展方向

科技档案的数字化通常涉及两个不同的层次。其中一个为库存档案的数字化，另一个是档案全文本的数字化。光学字符识别技术以文本格式存储文件内容，并辅

以具有全文搜索功能的数据库。然而，档案目录的数字化相对于传统的手工检索工具来说是一种极大的进步。它可以快速简洁地揭示档案的重要特征，如类别、主题、作者、创建时间等。然而，在瞬息万变的信息时代，档案目录的数字化远不是数字化的全部内容。科学技术档案管理。档案全文的数字化保证了科技档案全部内容的正确、系统呈现。通过保持源文件完整，计算机网络的优点可以快速轻松地应用于详细研究。来自各种科技档案馆的资料。利用本地网络、磁盘、驱动器等设备进行信息的远程检索、读取和交换，可以有效扩大科技档案和信息的传递空间。因此，档案全文数字化是科技档案数字化管理的主要方向。

## 五、结语

由于科技档案数字化管理实践，科技文献的使用效率提高了一倍。与传统的科学论文和技术档案相比，数字结果的使用更方便、更快捷、更准确，并且可供许多用户使用，同时，更加有利于企业档案人员对科技档案的科学管理。

## 参考文献：

[1]孙青.科技档案的数字化实践研究[J].档案管理理论与实践——浙江省基层档案工作者论文集,2014(00):237-239.

[2]王震亚,杨忠清.科研院所科技档案数字化的实践与思考[J].科技创新导报,2019,16(24):245-246.