

## 关于档案数字化的实践与思考

当前，我国信息化发展战略的实施，电子政务的推进，为我国档案事业发展提供了难得的机遇。加强档案信息化建设是档案事业适应时代和社会发展的必然选择，是加速档案管理现代化的客观要求，是提高档案服务水平的必由之路，势在必行。对此，作为管理和拥有大量产权产籍信息资源的房产档案部门，为更好地适应形势发展的要求，我们着眼未来，运用数字化、信息化技术，对档案数字化进行了有益的实践与探索。

### 一、档案数字化的意义

#### 1、档案数字化的含义

档案数字化是指利用计算机技术、扫描技术、数字成像技术、数据库技术、多媒体技术、存储技术等高新技术把各种载体的档案资源转化为数字化的档案信息，以数字化的形式存储、网络化的形式互相联结，利用计算机系统进行管理，形成一个有序结构的档案信息库，及时提供利用，实现资源共享，是档案信息化建设的重要内容。目前，大多数档案馆保存的档案信息形态主要以纸质、缩微胶片和底片等载体形式存在，只适应于传统的管理与传播方式，即实施手工管理和以点到点的传播，即便是经过编研等初步加工，编辑印发档案资料，实现以点到面的传播，其覆盖面也还是太小，很难达到社会化程度。档案信息资源的经济价值和社会价值也难以充分实现，更难适应当代“数字化生存”的形势。档案原件数字化，对档案信息的现代化管理、使用和传播等都具有重要的作用。

#### 2、档案数字化的主要内容

①档案目录信息的数字化，建立档案目录数据库。

②载体档案的数字化，如纸质档案、照片档案及录音录像档案等的数字化，建立档案影像数据库或多媒体数据库。

#### 3、档案数字化的主要作用

①档案数字化是档案信息化建设重要内容。档案信息化建设的核心是资源建设。资源建设包括两大方面任务：一是现有馆藏档案的目录数据库建设和馆藏重要纸质档案和照片、录音、录像档案的数字化；二是现行电子文件归档与电子档案管理。

②档案数字化能有效地保护档案原件。(1)、代替原件使用，保护档案原件。(2)、数字化副本异地保存，输出磁盘、磁带供永久保存。这样可使这些档案资料在出现天灾人祸的情况下不致于遭到毁灭性的破坏。(3)、恢复档案材料模糊褪变的字迹及对污损残缺照片档案的修复。

③档案数字化能改善档案的利用方式。(1)不再受“孤本”的限制，一份文件可以同时提供给所有需要它的人共享。(2)数字影像文件可以通过计算机局域网或者广域网进行异地传输，使异地调阅利用成为可能。扩大了档案的利用空间，让更多的人来了解档案，利用档案。(3)“时间”不再受限制，利用者可以随时使用自己需要的文件。档案数字化以后，将已开放的档案上网，这样所有利用者就可以在任何时间上网利用档案信息资源。

④档案数字化是传统档案馆走向数字档案馆的必经之路。数字档案馆无疑是 21 世纪各级各类档案馆的发展方向。尽管当前它的具体组成结构、组织管理模式还处于探索之中，可其“馆藏”的数字化特征是肯定的，传统档案的数字化将是其“馆藏”的重要组成部分。

## 二、档案数字化的实践

### 1、档案数字化工作回顾

为了加快实现档案的现代化管理，方便检索和提供利用，我馆多年来花大力气致力于档案信息库建设和档案的数字化工作。1993 年，我馆购置微机，独立开发人事档案管理系统，开始档案数字化的工作，使得全局职工基本情况在电脑中一览无遗，工资情况也能被系统所显示。1998 年开发文书档案管理软件，利用电脑管理文书档案目录。经过这些年的努力工作，我们将馆藏所有文书档案、人事档案全部录入电脑，并实现了电脑查询，极大方便了档案管理。2002 年，我馆与市房地产产权处合并，而产权处在 1995 年，开发出产权产籍管理软件，房产交易办证实现电脑化办公，并在实践中不断完善程序，将绝大部分产权产籍档案录入电脑，为全市房产发证加快了办事流程，减少了出现错误的几率。

### 2、实施档案数字化工程的必要性

首先，大力推进国民经济和社会信息化，以信息化带动工业化，实现社会生产力的跨越式发展，是党中央抓住信息化这一机遇，在党的十五届五中全会作出的战略决策。为了适应国家信息化建设和档案事业发展要求，2002 年 11 月 25 日，国家档案局制定下发了《全国档案信息化建设实施纲要》，指出了“十五”期间档案信息化建设的指导思想、目标与主要任务。

其次，随着城市建设的快速发展，老百姓购房办证数量的激增，为保障尽可能缩短老百姓办理权证时间，不出现发证错误和重复发证，传统管理根本不能满足要求。

最后，2002 年 10 月随着市房地产档案馆与市产权监督管理处的合并，档案管理工作出现了新的局面。湘潭市产权监督管理处开始建设湘潭市首个 GIS 系统——湘潭市房地产 GIS 系统，该系统最基础的工作也是最重要的工作是档案数字化工作（或称档案数据导入工作），只有保证档案数字化才能保障房地产 GIS 系统的基础的牢靠。

### 3、档案数字化工程的定位

产权产籍档案数字化工程是湘潭市房产局房地产 GIS 系统的子项目，旨在将海量产权产籍档案信息化、数字化，实现房地产 GIS 系统底层数据的全数字化，并通过建立后的系统实现“以图管房”，减少办证错误，提高办证时效，并通过 GIS 系统数据统计，向各职能管理部门和政府宏观调控部门提供数字化的档案信息及较为精确的统计报告，更好地为城市建设和房地产市场管理服务。

### 4、档案数字化工程需求分析

需处理量：**A、产权档案：**馆藏产权档案共计 35 万多卷。**B、房产抵押、担保、司法查封档案：**房产抵押档案 12900 余卷，担保档案 2050 余卷，司法查封档案 1500 余卷。**C、正常业务：**2003 年 9 月 1 日之后的所有正常业务档案。

处理措施：划分时间线 2003 年 9 月 1 日，区分处理方式。时间线之前的档案将其档案实体整理与装订，各类档案详细录入相关页面数据，档案原件通过数字照相技术扫描；时间线之后的档案，按照新开发的正常办证流程来实现数字化。

## 5、档案数字化工程的内容

将所有各类档案进行规范整理，然后，分成档案鉴定、档案扫描、档案导入三个大的环节，11 个小环节分类录入服务器数据库。不仅仅将档案原件扫描生成图片存储，并将档案原件上记载的 55 项相关重要业务数据真实、准确的录入数据库。

## 6、档案数字化工程实施情况

需求分析与设备考察阶段：调查分析馆藏情况，包括档案的类型、载体形态与状态、馆藏数量等基本情况。根据馆藏情况制订档案数字化的科学规划，确定项目需求，根据需求考察当前各种高速扫描系统的集成解决方案。在设备考察阶段我们特别注意这么几个问题：**A、档案数字化设备配置计划的成套性。****B、档案数字化设备的配套性，**也就是文档扫描仪和数字照相机优缺点对比，通过不断的实践，我们觉得数字照相机更加适应我们的工作。**C、档案数字化项目硬件与软件的集成性。****D、档案数字化方案与我馆需求的适应性。**总之，考虑到档案信息数字化要经过一个较长的过程，因此我们从实际需要和长远打算出发，力求确保电脑、照相机、操作人员等设施人员数量的成龙配套，使其具有较强的支撑能力和扩展能力。

系统实验阶段：2002 年底，项目开始立项。组织业务、技术骨干远赴深圳、上海、黑龙江佳木斯市等地实践考察，2003 年 6 月，正式成立房地产 GIS 系统领导小组、数据导入小组，调集相关业务、技术精英加入团队，正式开始实验阶段的工作，新购进了 12 台电脑微机，4 台佳能数码相机，4 台 EPSON 激光打印机，一台工作站服务器，设备安装调试完毕。通过前期和软件公司合作，共同开发我们所需要的软件。探索档案数字化工作在快速系统下，档案前期整理、档案扫描、质量检查、扫描文件管理、后期整理等全过程的协调配合问题，并探讨在高速系统下如何解决某些技术问题。经过近一年的测试，我们对整体系统流程和配套软件、数据库的性能有更进一步的了解，在项目的运作模式和系统管理等方面积累了一定的经验，同时，也锻炼了一支技术过硬的队伍，为系统的正式运行打下了基础。

项目实施阶段：本项目是一个长期、流程化运作的系统，从建立开始就必须确保能长期高速、稳定运行。在系统的开发与研究方面有一个时间周期，现把工作计划分为以下几个阶段：

### ①、系统总体方案框架设计：

2002 年 12 月—2003 年 6 月，完成系统总体框架，生成详细可行性分析报告，及相关配套业务标准、各类管理制度。

## ②、系统建立：

2003年6月—2003年10月，系统建立，包括：硬件配置及软件的开发。

## ③、系统试运行：

2003年10月—2004年2月。

## ④、系统正式运行：

2004年2月起，系统正式运行，同学校签订合作框架协议，调集员工和学生共100多人，经过一段时间的业务及技能培训，参与系统的正式运行，在运行中不断完善升级。

## 三、对档案数字化的思考

### 1、提高认识，统筹规划

首先必须在思想上充分认识档案信息资源的重要意义，在当今信息公开程度越来越高的形势下，档案的文化性质和社会性质逐步强化，利用的范围和对象将逐步扩大。在这个信息时代，只有当档案信息资源在社会主义现代化建设中发挥重要作用时，档案和档案工作的意义和价值才能充分地全面地展现出来。

其次，作好统筹规划。将大量的档案数字化，是一个庞大的系统工程。首先要做好馆藏情况的调查，包括档案的类型、载体形态与状态、馆藏数量、档案信息利用等基本情况。根据馆藏情况制订档案数字化的科学规划，包括服务器、电脑、扫描设备等硬件的购置计划和数字化处理规划。档案数字化实施处理规划应贯彻“突出重点，分步进行”的原则，对馆藏重点和利用率高的档案先行数字化，早日满足大多数利用者的需求。

最后，档案数字化是一项费时较长的工程，在大规模、流程化的数字化过程中，所有档案原件都需从档案库房分批大量取出，一定要保障档案原件的完整，不允许出现损毁和遗失的现象。

### 2、加强基础工作，认真做好档案数字化

①、档案数字化必须有所选择。针对馆藏“浩瀚”的档案资源，不加选择地全部数字化是不可能的、也是不必要的。我们应该根据一定的原则确定哪些馆藏档案原件应该数字化、哪些档案原件中哪些页不需要数字化。

②、做好数字化时扫描参数的优化设置。扫描系统扫描参数的选择和确定对扫描数字影像质量有着较大影响，其中扫描分辨率直接关系到扫描文件的清晰度和还原效果。我们在选择分辨率时应根据实际需要综合考虑，包括扫描文件的可阅读性、存储空间、输出打印质量等。

③、做好档案数字化的全程控制。在档案数字化过程中，要注重全面的质量检查，加强数据的质量控制。质量检查应包括以下几方面内容：**A**、数字化前的档案整理检查。要进行数字化的档案原件必须完整、排序正确，对折损严重影响扫描质量的原件应预先修整。**B**、

数字化中的人员业务培训。数字化工作是需要人来完成的，人员业务培训不过关，直接导致返工，间接导致成本的增加，时间段的延长。C、数字化后的数据质量、图像质量检查。录入后的档案信息要做到非常高的正确率，扫描后的文件图像应保持档案的原貌，字迹清楚不失真，无错扫、漏扫，对文件图像质量有问题的应重新扫描。

### 3、开展档案数字化问题的研究，及时解决数字化进程中的疑难问题。

档案数字化是档案工作中的一项新内容，不可避免地会遇到一些新情况、新问题。比如，数字化档案信息压缩及存储的关键技术问题；业务形态于现有技术手段相抵触；数字化档案信息安全、保密技术问题；数字化档案信息的访问形式与安全控制问题；数字化档案信息知识产权问题；数字化档案信息的法律地位问题。针对这些新情况、新问题，应在抓档案原件数字化工作的同时，组织力量开展档案数字化问题的研究，及时解决工作中的一些业务问题和相关技术性问题，同时澄清和统一思想认识，确保档案数字化工作的顺利开展。

另外，必须加强档案数字化和数字化档案信息管理人才的培养。只懂得计算机技术或只懂得纸质等载体档案管理知识的人无法胜任此项工作，所以应加强这方面技术人员的培养。

第16届国际档案大会刚刚闭幕，此次会议的主题是：“档案、管理与发展：描述未来社会”。我们应不断学习最先进的档案管理模式，不断走出去看看，开拓眼界，才能跟上时代发展的主旋律。

## 国土房管档案馆档案影像管理系统

### 一、需求分析

——国土房管档案影像管理系统，其核心是完成纸面档案的影像化管理工作，将传统的纸面文档管理转变为对影像文档的管理，将传统的对纸面文档的数据索引方式转变为面对影像文档进行索引的方式，从而为业务的开展提供足够的、快速的信息保障，进而为提高工作效率打下坚实基础。

——系统要求：分别成立专职的扫描工作点、索引工作点、查询工作点。扫描工作点主要进行集中的扫描工作，开展索引前纸面文件的预处理；索引工作点主要完成索引数据的录入，数据入库、生成相关业务报表、建立影像文档库等工作；查询工作点可以是本单位内局域网上的任意多个工作站。

——以下对已掌握的需求要点分别进行分析，以及提出解决建议：

#### 1、档案科目分类

需求：档案业务单式种类较多、分类查询方式也较多样

系统要求：软件在分类、录入上应当实现简便、快捷。

#### 2、档案规格、尺寸

需求：虽然档案的尺寸、规格基本相同，但考虑到可能要涉及到的附件纸张规格有不同规格和纸张质量（薄厚），这对扫描仪应当有着较高的要求。

系统要求：采用能够实现多种规格尺寸的档案、档案可以实现同时扫描的设备，不必对规格、尺寸做任何限制和要求。

#### 3、档案处理手续

在处理电子文档过程中，所有的档案都要按照类别、顺序号排列，但是如有缺漏要立即发验，

保证文档的完整。原始档案和影印件要准确无误的保存起来。

系统要求：系统工作流程定义时，避免流程的逆转，要求图像扫描和处理时要准确无误，不能存在纠偏变形、遗失档案和要素等问题。

#### 4、档案、档案要素

需求：档案的数量和格式不具有绝对固定性，不定期的会增加一些档案，档案的格式也存在变更的可能性。

系统要求：软件可根据档案数量、种类和要素的变化，可以自定义档案和要素。

#### 5、档案纸张

需求：档案的纸张的质量差别虽然不是很大，但要保证原始档案的完整性；另外，有的手写体和印刷体字迹不很清晰。

系统要求：不能撕坏纸张；必须能够对字迹进行增强；扫描是能够自动识别纸张的薄厚，对于一部分年久的纸面文档，应采用平板扫描的方式进行处理。

#### 6、档案的信息处理：

系统要求：历史档案处理，需要采用两台高端的扫描仪、2台中端自动进纸扫描仪，5台平板扫描，多台数码照相机，扫描时需人手做前期整理工作，以及后期的档案归档工作。

完成所有的历史档案的处理需要准备不低于 1.3T 的存贮空间。

#### 7、扫描方式

需求：档案的扫描不进行分散扫描，分散扫描会造成各口工作人员的工作效率的极大下降，并且扫描后的文档较为混乱，不利于后续的索引工作。

系统要求：在减少硬件投入的情况下，有利于提高工作效率，减少扫描的工作环节和更有效地进行质量控制，扫描之前应对被扫描的档案进行归类整理，以方便扫描，扫描过后由 QA 进行质量检查。

#### 8、正确索引

需求：要求写入数据库的索引数据要确保正确。

系统要求：为了确保索引相关信息被正确入库，可灵活采用一次录入，二次较对，或操作人员两次录入的工作方式，系统自动校对。

#### 9、影像文档及数据安全

需求：档案管理部门所涉及的档案，对安全性及其保密性有很高的要求。

系统要求：可对影像化的档案资料进行加密。

#### 10、影像文档的查询

需求：按一定的权限要求进行系统设置，完成查询的需求。

系统要求：采用安装客户端的方式，按一定的权限要求进行系统设置，完成查询的需求。

#### 11、备份归档

需求：国土房管档案影像管理系统要求提供可靠的数据存储与备份。

系统要求：按现有的数据量以及计划中的投资，可以采取大容量存储设备磁盘阵列和磁带机作近线备份的方式保存数据。

#### 12、查询响应速度

需求：系统查询速度需要考虑多用户、多数据时系统要能达到较佳的响应速度。

系统要求：影响系统查询速度的因素主要有三个：网络带宽、影像文件大小及并发用户的数量。所以影像文件的调阅速度基本上由周边网络设备决定。

## 二、解决方案

### 1、目标

——国土房管档案影像管理系统担负着档案的管理的任务，是一个专职处理档案的职能部门，通过为国土房管档案影像管理系统建立档案影像管理系统，引入新一代的文件管理技术，为



国土房管档案影像管理系统建立原始档案的影像档案库，以达到加强原始档案的管理、提高管理效率、降低管理成本、实现资料的快速查询及共享的目的。

## 2、概述

——信息行业作为知识经济时代的核心产业之一，近年来一直以惊人的速度在不断发展，而以纸面文件作为媒介的各种信息资料也呈爆炸性增长。如何有效管理这些信息资料，不断提高信息的使用效率，已经成为一个重要课题。

## 3、现状

——国土房管档案影像管理系统的在日常业务中，会产生大量档案，档案多以纸面文档的方式产生，以传统的数据处理方式进行处理，纸面档案每天都在大量地产生，再加上历史遗留文档，数量是非常惊人的，其中所涉及的信息数量也非常庞大。对于这些档案及信息，如何进行分类、保存和归档，录入、高效的查询，一直是一件非常棘手的事情。在传统模式下，总是被以下几个难题所困扰：

- 检索、查询和日常管理麻烦，工作效率不高；
- 纸面文档不好保存，很容易被损坏或弄污；
- 管理效率低下，管理成本居高不下；
- 档案调阅及信息利用效率较低；
- 工作低效率、高成本；

——从以上的介绍，不难看出仅仅用手工进行纸面档案的管理有很大局限性，在信息技术飞速发展的今天，也越来越不合时宜了。如何利用最新科技，完成纸面档案的电子化管理工作已经成为当务之急。

## 4、影像技术简介

电子影像是指各类文档的电子照片。

——影像技术是指将各种文档制作成电子照片，使这些文档的采集、分发、使用、管理、存储完全电子化的相关技术，而传统的电子化技术是把文档上的信息制作成电子数据，把文档的流转和处理转化为数据的流转和处理。

——影像系统是指对经影像化处理的文档完成数据录入、查询、调阅、批改、存储等功能的一整套解决方案，通过方案的实施能有效的解决传统手工纸面文档管理所无法解决的难题，具体有以下几方面：

- 原始文档电子影像化备份；
- 可靠安全的数据管理；
- 文档“影像原件”的高效查询；
- 在减少工作难度的同时，进一步推进管理工作的规范化。

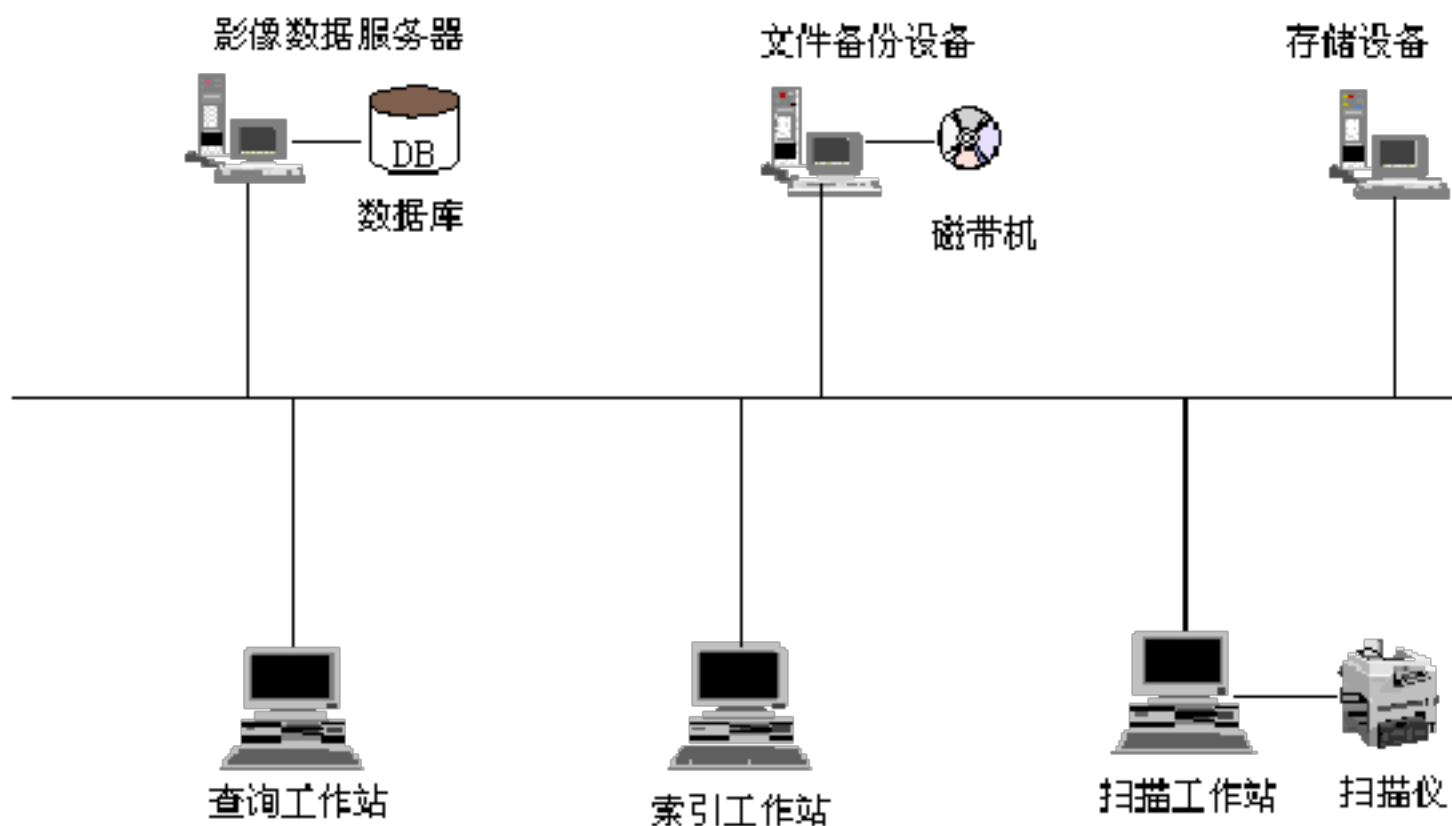
——文档影像化作为一种先进的信息管理手段，一经采用，必将大大提高纸面文档及数据的管理效率，降低成本，改进服务，提高信息利用率。

影像技术应用的可以分不同阶段开展

- 初级阶段：影像化存档：  
完成文档影像化工作及影像文件查询  
不涉及 workflow 技术  
不涉及业务过程
- 深化阶段：影像化业务处理：  
workflow 技术与影像技术结合  
融入业务处理过程  
从后台走向前台

### 三、详细的技术方案设计

#### 1、工作流程设计



注：当有以下情况发生时，工作流程按以下的规定进行：

- A. 扫描 QA 时发现漏扫、误扫、扫描质量太差：重新扫描。
- B. 索引时发现漏扫、误扫、扫描质量太差：重新扫描。
- C. 自动校对时发现错误：取错误的输入数据对照文件重新修改，计入日志，生成相应报表。

#### 接收档案

对所接收的纸质文件，根据一定的规则按类别、按份数完成文档的第一次整理工作。

#### 确立批次

以加入标识纸的方式，对每一份档案进行分隔；以一定的份数为一批，系统将自动生成批次号（注：也可以采用条码自动识别的方式进行分页及生成流水号，但采用条码方式需要加大软、硬件的投入。）

#### 2、扫描

##### 中高速扫描

系统采用 PANASONIC 高速扫描仪 KV-SS905CCN 2 台、中速扫描仪 KV-S7065CCN 2 台，以每个批次为单位，放入扫描仪的自动进纸进行扫描。扫描时，自动生成档案的分类、分页。

#### 3、影像处理

##### 扫描质检

检查扫描的电子影像文件是否有漏扫、误扫和扫描质量欠佳等，以保证扫描输出高质量的电子影像文件。

进行扫描质检时可以采取扫描完后，在扫描工作站上进行扫描的电子影像文件进行质检。

##### 扫描补漏

对于有质量问题的影像文件建议全批重扫以减少操作复杂性，个别情况下可以用平板扫描仪、数码相机对那些漏扫、误扫和扫描质量欠佳的文件进行追加扫描。

#### 4、原始档案装订入库

在对扫描后的电子影像文件检查无误，将原始档案采用常用的档案管理方式装订及入库。

#### 5、索引

在索引时可采用一人两次录入，系统自动校对，对系统校对不一致的数据，进行校验及录入。

#### 6、数据归档



接收数据及相应的报表，数据分别供业务数据库使用及影像文档库使用。

#### 7、归档备份

本系统采用一台服务器、一个专用磁盘阵列对索引文件、影像文件进行备份；在进行索引文件备份可以采用定时备份或定容量备份，还可以采用不定时备份，这样影像文件服务器遭到灾难性破坏，可以直接将所备份的索引文件恢复。

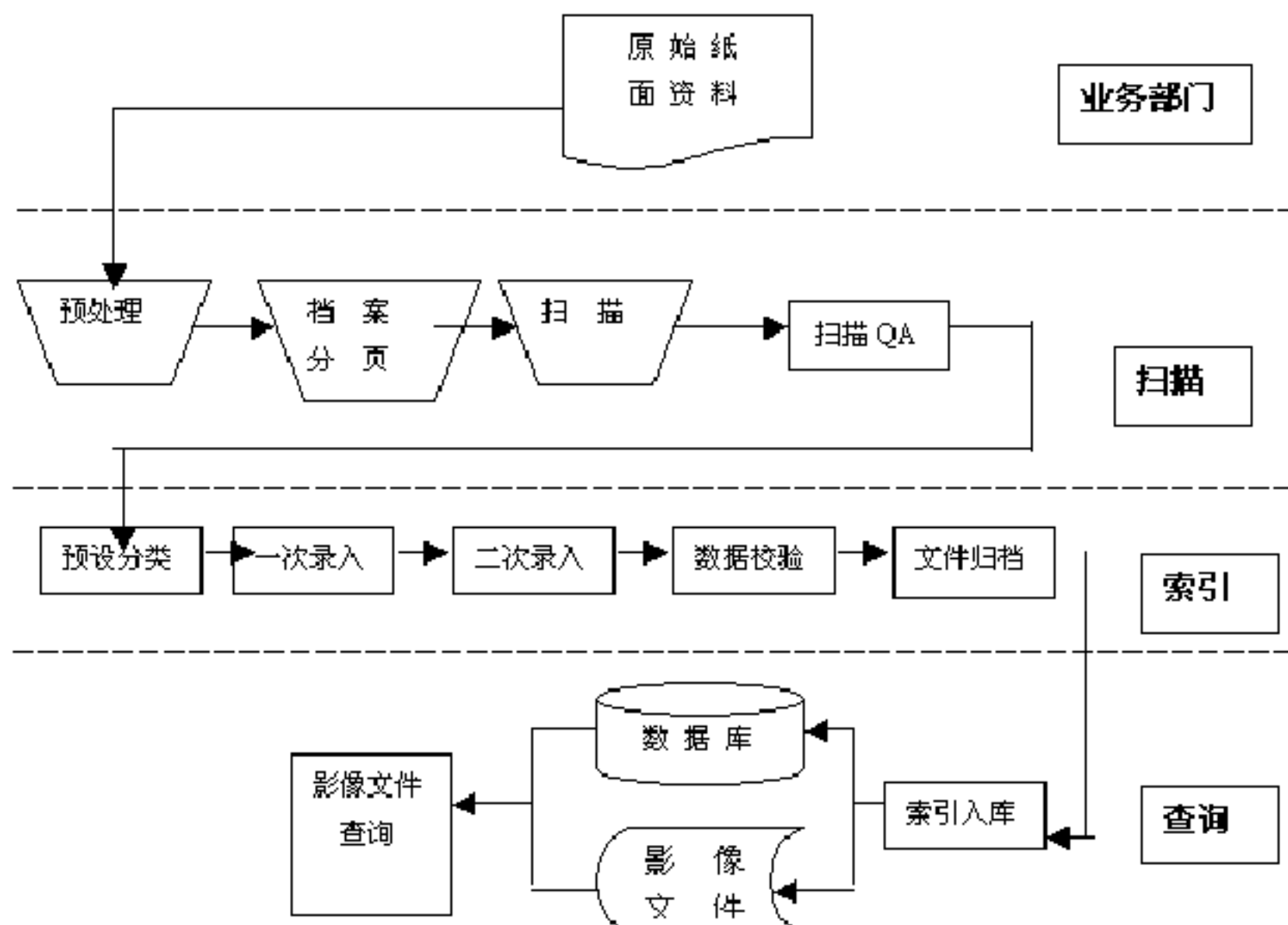
#### 8、检索、查询、管理影像文件

在对影像文件进行查询时，基于本系统的在局域网里的查询，采用安装工作站软件的方式，生成查询日志。

### 5 系统结构及配置

#### 四、系统结构

系统结构如下图所示：



#### 系统功能

我们设计的解决方案，由以下几个子系统功能组成：

##### 高速扫描

由高速扫描软件提供高速扫描图像资料的功能，用户的原始资料转换为电子影像资料。由高速扫描员使用。把保存的数据送到质量检验工序。

##### 质量检验及数据录入

影像质量检验与部分索引录入是把影像质量有问题和数据不正确的影像资料找出来，由质检员使用。索引的资料将会分别放在影像库中及数据库之中。

##### 系统管理

本系统提供的维护和生产情况查询功能，由系统管理员使用。

##### 安全管理

本系统分系统管理员、高速扫描员、质检员和查询客户。其中系统管理员负责整个系统的管理，拥有系统的所有权限。高速扫描员、质检员有相对应的工序使用权。而查询客户只拥有

#### 生产日志查询

可统计和查询在各工序间的未处理资料情况；

可统计和查询资料保存到影像数据库的情况。

#### 出错处理

如果在使用过程中发现索引信息有误，系统管理员可以删除错误纪录，调出相应存放地中的资料重新扫描并建立索引。

#### 影像查询

本系统提供资料影像查询功能，使业务人员可以通过 EZDOC 影像管理系统随时查找需要的影像数据，查询相应的影像文件；影像文件在查看时可以有放大、缩小、旋转、打印的功能；记录查询日志。

#### 校验功能

系统可以与其它系统的有机联接，可以自动与业务系统提供之数据进行对比。

#### 数据备份

数据备份分为两个部分：影像文件索引信息数据库备份、影像文件备份。

对于影像文件索引信息数据库的备份，电子档案管理系统不再另外提供备份功能，而是直接使用数据库管理系统的备份、恢复功能，保证数据库的完整性、有效性、可恢复性。

有关影像文件的备份：电子档案影像索引之后，首先存储于影像文件服务器，然后，系统自动完成备份机的备份任务。

### 5.3.2 系统特点

#### 无业务流程的逆转

国土房管档案影像管理系统要求：“准确”、“迅速”、“安全”、“方便”。自然，在业务流程中尽可能避免业务流程逆转，也就是说：在影像扫描的过程中尽可能避免档案重新扫描、档案漏扫等问题。流程的逆转会增加数据错误的可能，流程逆转的环节越多错误可能性越大。

#### 用户可以自定义档案分类

档案种类繁多，而且档案种类、档案要素并不固定，而且可能发生变化，如果系统将档案类别做成固定方式，则无法适合业务的发展。于是，我们在系统中为用户提供了一个开放式接口，可以让客户按照自己的方式增加、修改一些档案分类和档案要素。——不用编程，全部通过鼠标即可完成。

#### 用户可选择当前扫描、输入的文档类型

虽然，每天处理的档案的种类和数量较大，但并不是每种档案都有——通常数量大的主要分布在大业务上，有些档案通常几天出现一次。如果每天都要建立所有的分类，这将消耗不必要的存储空间。因此，我们提供了对当前要扫描档案类别的选择功能，客户可以根据当前业务情况进行选择。

#### 用户可以自己确定档案的要素

据上所述，我们的档案类别差异很大，而且不定期在发生着变化，那么，档案的数据要素也在发生着变化，为了让客户更好的适应这种变化，我们提供一个开放式档案要素设定。可以根据具体的档案设置、更改某些要素。

#### 多种安全保护模式

软件登录是采用密码方式。

数据库系统及影像文件也做了加密工作，防止他人修改影像数据和数据库数据。

#### 质量控制

在扫描过程中若出现漏扫、误扫或影像质量不好的问题，通过有效的工作日志的管理，解决以上问题。

### 课题组

《档案数字化方案研究》项目对档案数字化优化策略与原则、技术路线与支持、流程重组与再造、操作方法与步骤进行了全面的研究，达到了档案数字化全过程优化和管理系统整体优化的目的，形成了完整的档案数字化方案和成功的实践案例。国家档案局组织的专家鉴定委员会结论是：“该项目基于长春市档案馆以档案利用为导向、渐进性的档案数字化实践，提出了档案数字化工作的指导思想、优化原则和优化策略，符合国情，符合现代管理理念，对全国各档案馆的相关工作具有较大的参考价值；注重档案数字化工作与档案数据库建设、档案信息网络开发的有机衔接，提出了系统优化的技术路线，规划设计了档案鉴定与数字化同步、档案数字化与上网利用同步、多种档案数据库建设同步的流程，并成功地付诸实践，实现了档案数字化与上网利用的低成本、高效率、高效益。同时，运用 CA 认证、数字水印、电子签章等技术，为网上档案信息资源安全、准确利用提供了技术保证，也为档案管理部门提供了很好的实践模式”。该项目获国家档案局 2006 年优秀科技成果二等奖的同时，长春市档案馆被国务院信息化工作办公室和国家档案局确定为传统载体档案数字化试点单位。

#### 一、档案数字化的指导思想

以优化理论为指导，以用户需求为导向，以利用档案为目的，充分应用计算机软硬件功能，最大限度地发挥人力资源和数字化加工设备能力，兼顾保护档案原件的要求，保证数字化档案的真实可靠，更好地发挥档案信息资源的作用。

#### 二、档案数字化的优化原则

实用性原则：针对馆藏档案具备重复利用和紧急利用两大特点，急用、常用档案优先数字化是档案数字化的首选。

开放性原则：互联网给档案信息资源共享提供了超越时空的条件，国家民主政治建设和政府信息公开催促档案解密和开放，开放档案及时数字化是实现其社会共享的前提。

价值性原则：重要档案优先数字化具有档案利用和档案保护双重意义，是网络时代更好实现档案价值的客观需要。

抢救性原则：计算机存储给档案抢救开辟了新的途径，档案数字化是保护濒危档案原件最经济和便捷的方法。

时效性原则：档案数字化加工与数字化档案利用同步实现，是充分利用计算机网络技术，更好、更快地实现档案数字化效益的最佳途径。

#### 三、档案数字化的优化策略

需求导向策略：社会需求是制定和调整档案数字化工作策略的“风向标”。汇集多年档案利用情况统计，把利用频繁档案的进行综合分析，科学地圈定优先档案数字化的范围。并建立快速反应机制，执行常用档案继续做与急用档案突击做的计划，达到数字化档案利用即时见效的目的。

本级为主策略：地方档案馆保存的本级永久档案是数字化的重点。进入信息时代，馆际间重复的上级档案文件，与其分头进行数字化及开放鉴定，不如发挥中国特色档案管理体制的优势，由上级档案行政管理部门组织，向馆际共建、共享数字化、开放鉴定成果的方向发展，尽量减少档案馆多层的重复劳动。

分步实施策略：馆藏档案数字化不可能一步到位，而且随着时间的推移，档案室进馆范围的档案将陆续移交到档案馆，这些档案采取馆室共建、共赢、共享的数字化策略非常有意义。

剔除无用策略：馆藏现行档案是否进馆往往由立卷人员决定，加之保管期限表比较笼统，馆

年以后，市以下档案馆长期保管的档案逐年超过保管期的上限。因此，开展档案价值鉴定，剔除无用，以免再造数字化垃圾，避免资源浪费应是长久的策略。

#### 四、技术路线的优化选择

档案鉴定与数字化同步：从利用上讲，解决制约档案信息资源共享的主要瓶颈——档案开放鉴定问题是当务之急；从成本上讲，档案价值鉴定是避免失效档案数字化的最佳方法。只有这样，才能控制档案数字化投入与产出比例失调、数量很大共享很少的问题。

多种档案数据库建设同步：一是首先建立文件级目录数据库。满足一次检索需要，建好文件级目录数据库。目前，绝大多数档案馆，因没实现全部档案的文件级目录机检，查档不快、不准、不全的问题不可避免。所以，把文件级目录数据库建设作为基础性工作，是提高档案查全率、查准率必经之路。二是适时建立专题数据库。随着社会发展，对某一专题档案集中需求时有发生，突击建立专题数据库，满足社会需求是发挥档案作用的最佳时机。三是优选建设全文数据库。满足档案信息资源开发利用需要，优选建设全文数据库。按照国家档案局提出的利用频繁档案首先数字化的要求，遵循“二八”定律（即以 20% 的努力达到 80% 的效果），同步开展档案全文数据库建设。

档案数字化与上网利用同步：档案数字化的根本目的是充分发挥其作用。通过计算机网络技术，实现档案数字化与局域网上利用同步，开放档案数字化与互联网上利用同步。

#### 五、重组和再造的优化流程

为发挥计算机网络的优点，重组和再造优化流程的实质，就是对档案数字化整个工作流程进行系统分析，不但要打破原有业务规程，还要改变行政管理分工，进行全方位的变革。档案数字化的任何一个细节都会对整体产生影响，轻者降低效率，重者决定成败。长春市档案馆高度重视每一个细节，再造和优化了上百个流程和内部标准，保证优化技术路线的顺利执行。

#### 六、档案数字化示范案例——长春模式

从 2004 年开始，长春市档案馆以科研课题为先导，以项目管理为手段，以效益实现为检验标准，组建并成功运行档案数字化工作组织架构、资源配置、数字化加工、上网利用“一站式”流水作业线，探索档案数字化和网上利用服务工程优化模式，终于闯出一条独特的档案数字化之路。

##### （一）实施三个工程，建立一个团队

长春市档案馆把信息化带动作为首位战略，视为一把手工程、班子工程、全员工程突出抓，并创新管理机制，打破行政分工，在计算机网络环境下，建立了行政领导、指导专家（技术业务总负责）、专业人员、外聘技工 20 人的长期流程团队，并与现有计算机和扫描设备进行合理搭配。该团队由局（馆）党组直接领导，馆技术部门提供支持，管理部门提供档案，专家提供指导，创立了高效的组织体系，搭建了投入小产出大机制框架。

##### （二）集中四个环节，坚持流水作业

文件级目录数据录入是档案数字化的基础，档案鉴定是档案信息资源共享的前提，全文数字化是提升档案服务水平的途径，整理是档案文件实体管理主要环节，把上述四个环节集中，是提高工作效率和质量的最佳选择。长春市档案馆四集中流水线上，目录数据录入岗位安排业务和技工 7 名、档案鉴定岗位安排 5 人、档案数字化加工 4 人、档案实体整理 4 人，在网络环境下实行集中办公和流水作业。

##### （三）数字化成本降低，数字化速度加快

由于采用高效的组织体系和优化的技术路线，已有 20 万件档案走下流水线。

低成本主要体现在：20 万件档案缓销剔出占 24.7%，留存档案占 75.3%；留存的 75.3% 档案，全文数字化占 25.6%，现用现扫占 49.7%，比全部扫描节省费用 40% 以上。同时，三年时间建成数据库 50 个，文件级数据库 120 万条（包括 20 万件全文）；专题数据库 120 万条；照

张；资料数据库 5000 条，案卷级目录数据库 23 万条；现行文件和政府信息公开信息 1 万余件（条）。局域网运行档案目录数据 260 万余条，20 万件全文；互联网运行档案目录数据 120 万条，8 万件全文。完成这样规模的数据库建设，全部人工费只有 58.4 万元（包括流程团队档案干部的工资），经济成本很低，时间成本很少。

高效率主要体现在：价值鉴定、开放鉴定、数字化筛选、原件扫描、实体整理同步完成，比各项分步做效率高数倍，同时有现用现扫的方法配套，投入少量设备，相当于将 20 万件（200 万页）档案全部扫描的效果。

#### （四）利用方式变革，社会效益显著

为充分发挥网络优势，长春市档案馆采取了直接与间接相结合的方法提供利用。

1、直接提供利用：对数字化档案全文，在局域网内，接待人员利用网络检索后直接打印，为用户提供数字化档案打印件。在互联网上提供数字化档案全文下载、打印服务。并通过 CA 认证、电子签名等安全技术，对特定的用户直接提供加盖电子签章的数字化档案打印件，已有案例。

2. 间接提供利用：对目录数字化档案，在局域网内，接待人员机检到档案后通知档案库管人员，档案库管人员调出原件全文扫描后，通过网络打印机直接传输打印件，由接待人员提供给用户。

目前，长春市档案馆基本实现了三个替代：网络检索替代手工目录检索，查看数字化档案替代查看纸质档案，网络打印替代手工复印。

高效益体现在：档案全文数字化占馆藏档案的 10%，提供数字化档案占利用率的 70%。为上万人提供了档案证明，公众成为档案数字化的最大受益者。档案数字化成果得到社会公众，上级档案和信息化部门，当地党委、政府，国内外同行的普遍赞誉，数以百计的留言、锦旗、感谢信、奖励证书、媒体报道就是最好的证明。

#### 档案数字化解决方案

档案数字化是指采用扫描仪或数码相机等数码设备对纸质档案进行数字加工，将其转化为存储在磁带、磁盘、光盘等载体上并能被计算机识别的数字图像或数字文本的处理过程。公司装备多种扫描设备：高速扫描仪、零边距扫描仪、宽幅扫描仪、书刊扫描仪，可处理各类纸质文档。可根据不同文档的材质、幅面大小及内容，提供不同的电子化扫描及索引（著录）解决方案，并可根据客户需求量身定制扫描加工方案。

经验丰富的实施生产队伍

自 2002 年开展数字化扫描加工业务以来，掌握了较全面的专业知识及积累了丰富的工作经验。

成熟的管理机制、有效的管理控制手段。

严谨的质量控制手段，可实现扫描加工过程中各个环节的质量控制和自动化的量化记录，可确保扫描加工的影像、索引的正确性。

严格的客户财产保全制度及控制手段。可确保客户文档资料、数据在扫描加工服务中的安全性、保密性。

先进、灵活的扫描加工生产系统

公司自主研发的扫描加工生产软件系统，基于工作流的任务机制，实现流水线式的生产模式，可根据客户的需求灵活的配置、调整，以适应不同客户、不同类型文档的加工任务。

包括：文件、档案、文献、资料、报纸、书刊、杂志、合订本、汇编

扫描模式：黑白二值、灰度、彩色

扫描分辨率：150DPI~2400 DPI



、JPG、CEB、PDF 等。

扫描加工生产软件系统针对文件、档案管理工作的特点，通过集合图像处理、OCR 等多项自动化智能技术，实现处理扫描图像自动纠偏、除黑边、去污等处理，实现扫描文件自动命名与装订，生成多页 TIF 文件，通过 OCR 识别技术实现电子文件元数据半自动获取，在大幅度提高扫描加工的工作效率的同时，可以完成电子文件元数据的获取工作，并在保证扫描图像文件质量前提下，以最低的容量存储文件。

主要功能介绍

- ①自动扫描与目录图象属性加工设置；
- ②自动倾斜校正/去黑边/去白页/加粗/A3 切割为 A4；
- ③前置处理设置：去斑、亮度/对比度、颜色通道设置；
- ④自动文件命名；
- ⑤自动装订（将多个单页 TIF 文件合并为 1 个多页 TIF 文件），支持固定页分割、白页分割、标志按钮分割；
- ⑥自动 JPG 输出质量设置；
- ⑦自动扫描页计数（见主界面扫描计数器），目录中文件与页数自动统计。

根据各档案馆（室）在处理历史档案问题的经验，结合我公司图像处理技术上的优势，特提出了档案数字化制作流程及相应的技术参考。

#### 1、档案分类整理

档案分类整理是指将案卷按扫描要求进行分解整理，包括原件的拆卷、起钉、粘贴页撕开。并按扫描后的电子文件组织形成进行分类，以便扫描时的批次处理。

#### 2、档案交接

分类整理后的档案按时间先后或其他分类项顺序提出，并按实际卷（页）数填写详细的档案交接清单，并由双方管理人员签字。

#### 3、标引著录

标引著录为了方便查阅与管理，打印制作档案目录及档案封面，对文字录入的准确性要求较高，但著录的文本域并不是完全统一的在某一页面上，一些文本域要在多个文件中选择，所以录入时需要相关档案管理人员进行指导，制作相应的《文字录入工作说明书》，确保文本域录入的准确率。

#### 4、档案的扫描

原件的扫描与存储格式完全按照国家档案管理的统一标准要求扫描。扫描图像使用先进的图像扫描处理软件，扫描人员可根据原稿质量，对系统进行定义，如倾斜校正、去污等批处理功能。在由系统自动处理功能的同时，扫描人员根据原件的实际情况做相应的调整，如超大页面的处理、纸张颜色深浅及薄厚的处理、扫描时可根据不同原件的情况，调整图像的分辨率、明暗度以及扫描方式和扫描速度，确保在扫描图像质量清晰的情况下，使原件完好无损。图像的分辨率正常情况下为 150dpi，如遇字间距和行间距过密、原件本身是复印件等字迹不清楚的情况，可适当增加扫描的分辨率，但要保证图像清晰的同时，又不影响远程查询和浏览的速度。

#### 5、图像处理

档案数字化制作模块中的图像处理采用自动化处理和人工处理相结合的方式，确保图像质量的完美。通过对图像进行校对，确保图像顺序正确，在校对时，发现不合格图像及时返回前一工序进行改正。

- 1) 污渍、黑白、偏斜处理
- 2) 全文字面的扫描密度

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/575301032301011312>