

# 机关档案信息化建设

——《机关档案管理规定》解读之十三

文/丁德胜 邹杰

《机关档案管理规定》(以下简称《规定》)明确了机关档案信息化建设的工作内容和建设思路,对传统载体档案数字化工作、电子文件归档与电子档案管理工作进行了系统规定,同时在电子档案效力、仅以电子形式归档(单套制)等方面提出了创新性要求,为机关档案信息化工作提供了全面、有效指引。

## 工作内容与建设思路

### 1. 工作内容

机关档案信息化建设是一项综合性、系统性工作。《规定》第一章至第三章及第五章有关条款对档案信息化工作的组织规划、制度机制、管理目标、工作任务、人员要求、基础设施等内容进行了规范。第五章对档案信息化建设进行专章规定,明确了数字档案室建设在机关档案信息化建设中的基础地位,同时将传统载体档案数字化工作、电子文件归档和电子档案管理工作确定为机关档案信息化工作实施的重点内容。第四章则从档案管理角度提出了电子文件归档与电子档案管理工作的工作流程控制要求。开展机关档案信息化建设,要将《规定》第一章至第五章的信息化工作内容(参见《〈规定〉档案信息化工作内容一览表》)

《规定》档案信息化工作内容一览表

主要内容	相关章节	备注
组织规划	第六条	组织领导、建设经费
	第五十五条	统一规划、工作协同
制度机制	第九条	工作机制
	第六十条	管理制度、职责分工
管理目标	第五条	“四性”管理要求
工作任务	第十条	信息化基本任务
	第十一条	社会化服务
人员要求	第十二条	信息管理知识
基础设施	第十四条	信息化用房
	第二十二條	智能库房设施设备
	第六十一条	信息化基础设施
数字档案室建设	第五十六条	工作内容、基本遵循
传统载体档案数字化	第五十七条	常态机制、真实性管理
	第五十八条	全文识别
	第五十九条	目录数据库
电子文件归档与电子档案管理	第六十二条	应用系统
	第四章	工作流程控制
	第六十三条	格式与质量
	第六十四条	存储
	第六十五条	备份、转换与迁移

表)结合起来,综合考虑,统筹推进。

### 2. 建设思路

《规定》要求:“机关应当开展数字档案室建设,统筹传统载体档案数字化、电子文件归档与电子档案管理工作。”结合《规定》第三十六条、第四十九条、第五十八条提出的仅以电子形式归档(单套制)、数据分析与文本挖掘、纸质档案数字复制件的全文识别等

要求,构成了机关档案信息化的建设思路:以数字档案室为依托,统筹开展传统载体档案数字化、电子文件归档与电子档案管理工作,并在全文识别工作和真实性管理基础上,最终实现电子档案数据化、单套制和数据挖掘3项目标,即“一个平台、两大任务、三项目标”(参见图1)。

一个平台,即数字档案室建设。数

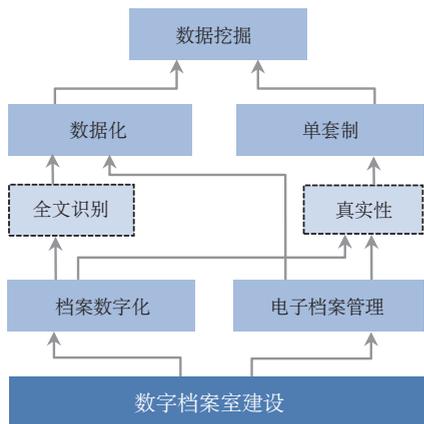


图1 机关档案信息化建设思路

字档案室建设是应对信息化挑战、解决机关档案信息化核心问题推出的一项重要工作安排。2014年，国家档案局发布《数字档案室建设指南》，正式在机关中推动开展数字档案室建设。同年，中办、国办印发的《关于加强和改进新形势下档案工作的意见》指出：“各地区各部门各单位要把数字档案馆（室）建设列入信息化建设整体规划，从人力、财力、物力上统筹安排。”《国家电子文件管理“十三五”规划》也将“推进机关数字档案室建设，完善电子文件管理相关系统”作为国家电子文件管理规划的一项重要任务。由此，数字档案室建设成为档案信息化工作的一项重要内容，也是机关开展档案信息化工作的重要抓手和主要形式。

两大任务，即传统载体档案数字化工作、电子文件归档和电子档案管理工作。2013年，在全国数字档案馆（室）建设推进会上，国家档案局提出实施“存量数字化、增量电子化”战略，确定了档案信息化的两大任务。这两大任务，一个解决传统载体档案在信息化条件下的有效管理，以及传统载体档案与信息化技术紧密联结的问题；一个解决新产生电子文件的归档和电子档案规范管理问题。两大任务兼顾档案存量和增量情况，考虑了现实需要和未来发展需求，需要同步推进实施。

三项目标，即档案数据化、单套制和数据挖掘。同传统做法中档案数字

化、双套制、档案利用相比，三大目标意味着档案工作及管理对象、管理方式、管理目标上发生了质的变化。档案数据化是将档案信息结构化、颗粒化，为数据挖掘提供标准的、通用的数据对象。单套制除节省成本、提高效率之外，还有助于倒推电子文件、电子档案管理，形成电子文件与电子档案管理范式。数据挖掘是从大量的、不完全的、有噪声的、模糊的、随机的数据中，提取潜在有用的信息和知识的过程，既能以新需求推动电子文件与电子档案管理的进一步发展，又能为档案及档案工作价值的体现创造新的机会和展现领域。

## 传统载体档案数字化

数字化是利用计算机技术将模拟信号转换为数字信号的处理过程。传统载体档案数字化工作即所谓“存量数字化”，对于提高传统载体档案共享利用水平、确保传统载体档案安全意义重大。

### 1. 基本要求

《规定》对机关档案数字化工作提出了4项要求：一是要求机关建立档案数字化常态机制，明确档案数字化是机关档案工作经常性业务，要有序开展档案数字化工作。二是档案数字化应当符合真实性管理要求，数字化过程的元数据应当收集齐全，数字复制件应当保持原貌并纳入电子档案管理系统统一管理。三是明确各类传统载体档案数字化的技术遵循，要求按照相关标准开展纸质档案和录音录像档案数字化工作，并实施纸质档案数字复制件的全文识别。四是明确档案目录数据库建设要求。

### 2. 全文识别

全文识别（OCR）是纸质档案实现全文检索、进行数据化的重要前提，可以结合档案数字化工作实施，也可以独立实施。纸质档案全文识别应按照《纸质档案数字复制件光学字符识别（OCR）工作规范》进行。

工作流程。主要包括图像导入、图像预处理、比对识别、修改校正、成果整

理输出5个业务环节。其中比对识别是关键环节，要求针对档案特征进行分析、识别和匹配。具体来说，就是要建立归档章式样库，自动识别图像中的归档章及字段位置；建立公文格式库，准确识别公文的版头、主体、版记3部分，识别公章、签章等区域及内容；建立专用表格模板定义工具，自定义文件处理单、发文稿纸、备考表等表格模板，识别表格中的字段位置；建立印章名称与印章图像的关系库，识别印章图像位置，用于版式恢复。

质量要求。一般来说，纸质档案OCR对档案中文、数字、英文印刷体的识别准确率应达到95%以上，对常用签名识别准确率达到90%以上，对手写体识别准确率达到80%以上。在抗噪能力方面，要求能够有效屏蔽较大程度的噪点干扰，准确判别纸质档案数字复制件上的污点、污线、黑边、纸张褪变斑点、水渍、污点、装订孔等。在识别速度方面，要求在主流计算机软硬件平台下A4纸幅面中文识别速度不低于1000字/秒，英文识别速度不低于2000字/秒。在版面还原方面，要求采用分栏技术，智能分析中文（简体、繁体）、英文字体，文、表、图混排文本，识别后无需人工干预，自动还原排版且还原度达到90%以上，实现复杂版面的精确还原。

### 3. 目录数据库

在机关开展档案数字化工作中，目录数据库建设是其中必不可少的一项内容。机关建立起包含各个门类档案的目录数据库，对于统筹各门类档案管理意义重大。

建设要求。一是著录项目齐全。要求档案目录数据库著录项目设计符合《档案著录规则》（DA/T 18—1999）和该门类档案管理具体要求，不缺项漏项，并且明确项目内容长度、格式、值域和必著项。二是著录条目齐全。以卷为单位整理档案的每个案卷、卷内每份文件，以件为单位整理档案的每件档案都要逐一著录。三是著录内容规范。要求按照著录标准和要求进行著录，并进行系统自动检查和

## 专题·《机关档案管理规定》解读

人工抽查相结合的核查校验。电子档案管理系统应自动控制用户著录内容的准确性,对于用户录入的非法字符、无效内容、缺漏字段等进行提示。

融合管理。《规定》提出档案目录数据应当与相应数字化元数据、电子档案管理系统相应元数据库融合管理。所谓融合管理主要指4个“融合”:传统载体档案目录数据和电子档案元数据融合管理、传统载体档案数字化过程元数据和电子档案管理过程元数据融合管理、传统载体档案数字复制件电子属性元数据与电子档案电子属性元数据融合管理、传统载体档案数字复制件与电子档案格式转换和系统迁移元数据融合管理。

### 电子文件归档与电子档案管理

电子文件归档与电子档案管理工作,即所谓“增量电子化”。需要说明的是,归档在实践中一般具有2种含义:一是在档案管理流程中作为收集、整理流程的后续流程;二是作为收集、整理以及业务部门或文书部门与档案部门移交、接收档案等工作的统称。电子文件归档和电子档案管理中的“归档”采用第二种含义。

#### 1. 主要工作内容

电子文件归档与电子档案管理工作需要统筹解决基础设施设备、应用系统平台支撑、档案资源规范、制度和人员保障等关键问题,这与数字档案室建设内容是一致的。按照《规定》要求,做好电子文件归档和电子档案管理,要重点做好以下4个方面工作:一是基础设施建设。要求配备的基础设施设备应当满足档案信息化的管理需要并适当冗余、方便扩展,其中服务器和存储设备应当专业、专用,并可在电子政务云或自建私有云上建设电子档案管理系统。二是应用系统建设。要求机关办公自动化系统和其他业务系统应当实施预归档,电子档案管理系统应当满足电子档案真实性、可靠性、完整性、可用性管理要求,等级保护不得低于二级标准,分级保护等级与电子档案最高密级相适应。三是

档案资源建设。要求电子档案应当标准规范、齐全完整,电子档案和元数据符合标准要求,满足长期保存和同级国家综合档案馆进馆要求。四是保障体系建设。要求应当建立电子文件归档与电子档案管理制度等。下面重点对应用系统建设、档案资源建设要求作出说明。

#### 2. 应用系统建设

电子文件与电子档案管理系统主要包括业务系统(BS)、电子档案管理系统(ERMS)和数字档案资源长期保存系统(TDR)。三类系统前后衔接,实现了对电子文件和电子档案的全生命周期管理(参见图2)。其中,与机关电子文件归档与电子档案管理工作相关的是业务系统和电子档案管理系统。



图2 BS (EDMS)、ERMS和TDR关系图

业务系统是指形成或管理机构活动数据的计算机系统,机关常见的有办公自动化系统及其他业务系统。这些系统内部往往设置电子文件管理系统(EDMS),对电子文件进行管理。《规定》要求“机关办公自动化系统和其他业务系统应当嵌入电子文件分类方案、归档范围与保管期限表和整理要求,在电子文件形成时自动或半自动开展鉴定、整理工作,实施预归档”,对业务系统实施预归档提出了要求。按照《电子文件归档与电子档案管理规定》(GB/T 18894—2016)要求,业务系统电子文件预归档功能包括:内置电子文件分类方案、保管期限表等工具,支持电子文件形成或办理部门完整收集、整理应归档电子文件及其元数据;应能按内置规则自动命名、存储电子文件及其组件,保持电子文件内在的有机联系,建立电子文件与元数据之间的关联关系;应能按标准生成电子文件及其元数据归档数据包,或向归档接口推送电子文件及其元数据等。

电子档案管理系统是指对电子文件、电子档案进行捕获、维护、利用和处置的计算机信息系统。《规定》要求:“电子档案管理系统应当功能完善、适度前瞻,满足电子档案真实性、可靠性、完整性、可用性管理要求。电子档案管理系统基本功能和可选功能应当参照《数字档案室建设指南》《数字档案室建设评价办法》《电子文件管理系统通用功能要求》(GB/T 29194)以及同级国家综合档案馆的相关要求执行。”按照《数字档案室建设指南》要求,电子档案管理系统主要功能包括档案管理配置、接收采集、分类编目、检索利用、鉴定统计、系统管理等。其中:档案管理配置要求具备档案门类、档案分类方案、文件材料归档范

围、档案保管期限处置表、元数据方案的配置与管理功能;接收采集要求具备通过在线接收、批量导入和手工录入等方式,自动或手动接收采集文书、照片、录音、录像等不同类型电子文件功能,辅助接收采集传统载体档案或档案数字复制件功能;分类编目要求具备电子文件分类、编目、命名、存储、归档功能,具备辅助完成传统载体档案的分类、编目功能;检索利用要求具备以权限管理为基础,进行多途径、多角度且易用的检索和利用功能;鉴定统计要求具备各门类档案资料的鉴定处置功能,具备对各种档案资料以及相关业务情况的统计管理功能;系统管理要求系统具备日志审计跟踪功能,记录、审计系统各类管理、操作行为。

#### 3. 档案资源建设

在档案资源建设方面,《规定》从工作流程控制和核心规范要素2个角度提出了要求。工作流程控制主要规范“过程”,核心规范要素用来保证“结果”。



工作流程控制。电子文件归档与电子档案管理贯穿档案工作整个流程。

《规定》第四章除一般性规定外，还对电子文件、电子档案提出了针对性要求：形成与收集阶段，对电子文件形成、电子文件和元数据收集提出要求；整理与归档阶段，对电子文件整理和电子文件随办随归、原件归档、单套归档提出要求；保管与保护阶段，对信息安全和电子档案保管提出要求；鉴定与销毁阶段，提出了电子档案和档案数字复制件销毁要求；利用与开发阶段，提出了数据分析、文本挖掘的开发要求；统计与移交阶段，对统计结果可视化、数字复制件和电子档案移交提出了要求。

实践中，各个流程还需要划分为若干工作内容或工作步骤予以实施（参见图3）。比如，形成流程主要工作内容是对电子文件格式、元数据、电子印章等进行规范 [电子公文按照《电子文件存储与交换格式》（GB/T 33190—2016）、《党政机关电子公文元数据规范》（GB/T 33480—2016）、《党政机关电子印章应用规范》（GB/T 33481—2016）分别规范文件格式、元数据、电子印章]。收集流程主要包括捕获、录入等工作步骤，要完成电子文件及其元数据的捕获工作；同时对于部分未进入业务系统、电子档案管理系统流转的文件通过扫描录入、批量导入等方式完成收集。整理流程要重点关注组件、编号、格式转换、归档数据包组织等工作步骤 [详细步骤参见《归档文件整理规则》（DA/T 22—2015）]。其中：格式转换为可选工作步骤，要求在电子文件捕获或录入后对不符合归档格式要求的电子文件进行格式转换；归档数据包组织是指将完成组件、编号的电子文件及其元数据组织为归档信息包。归档流

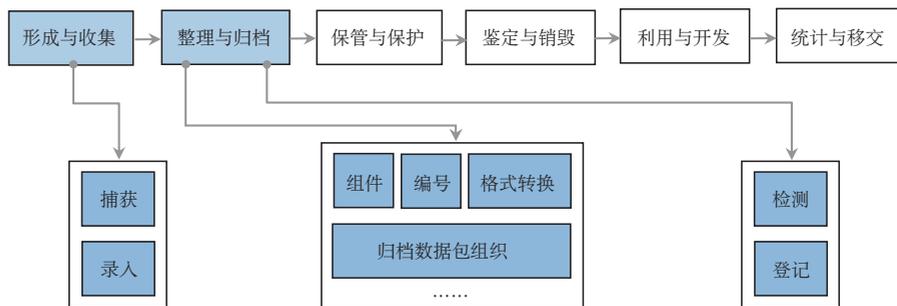


图3 收集、整理、归档流程主要工作步骤

程要办理移交与接收手续，主要包括检测、登记等工作步骤。其中：检测应按照《文书类电子档案检测一般要求》（DA/T 70—2018）等标准执行，完成“四性”检测；登记一般由系统自动完成，应按照《电子文件归档与电子档案管理规范》（GB/T 18894—2016）执行，通过计算机文件名建立电子文件与元数据的关联，在管理过程元数据中记录登记行为，并按批次或归档年度填写电子文件归档登记表，完成电子文件的归档。

核心规范要素。档案资源建设的基本要求是齐全完整、标准规范。为实现这一要求，要重点抓好电子档案的命名与存储、格式与质量、元数据要求等核心规范要素。

一是命名与存储。命名是指采用由字母和阿拉伯数字构成的一个字符串为存储于硬磁盘、光盘、磁带等载体中的电子档案命名，即计算机文件名。为方便管理，保证存储路径清晰，应按照档号依序为电子档案的计算机文件进行命名，并按照档号构成项逐级建立文件夹进行存储。

二是格式与质量。一般来说，文书档案（或者类似文书档案管理的档案）带有签批或电子印章的电子档案稿本应采用OFD等版式文件格式保存；集中记录拟

制、办理过程的稿本可采用原流式文件格式保存。在电子档案质量要求上，正本的公文版面格式、各要素编排规则等应符合《党政机关电子公文格式规范 第1部分：公文结构》（GB/T 33476.1—2016）、《党政机关电子公文格式规范 第2部分：显现》（GB/T 33476.2—2016）、《党政机关电子公文格式规范 第3部分：实施指南》（GB/T 33476.3—2016）等标准要求。照片、录音、录像档案的格式、质量要求参见相关标准。

三是元数据要求。元数据要求主要包括收集范围、存放方式等。不同门类电子档案元数据的收集范围参见《数字档案室建设指南》及相应的元数据标准。在存放方式上，归档阶段元数据一般以XML文件方式存放；管理阶段元数据一般保存在电子档案管理系统数据库中，通过档号等与电子档案内容保持关联；移交阶段元数据一般采用《电子档案移交与接收办法》（档发〔2012〕7号）规定的电子档案存储结构或《基于XML的电子文件封装规范》（DA/T 48—2009）规定的EIP封装包结构与电子档案内容统一保存。

作者单位：国家档案局、青岛市档案馆  
责任编辑：田小燕

### 中国档案杂志社领取新闻记者证人员名单公示

根据《国家新闻出版署关于2019年全国统一换发新闻记者证的通知》和《新闻记者证管理办法》要求，我社已对本年度申领记者证人员的资格进行严格审核，已收回、作废旧版新闻记者证，正式使用新版记者证。现将领取新闻记者证人员名单进行公示。

监督举报电话：中央新闻单位公布新闻记者证核发办公室电话：010-83138953  
中国档案杂志社办公室电话：010-63160071

序号	姓名	部门	职务
1	黄佳音	编辑一室	副主任

