

# 《干部人事档案数字化技术规范》 (GB/T33870-2017)

来源：中国国家标准化管理委员会

The screenshot shows the official announcement page for the national standard GB/T 33870-2017. The page is titled '中华人民共和国国家标准公告' (Announcement of National Standards of the People's Republic of China) and is dated '2017年第15号' (No. 15, 2017). The announcement is issued by the State Quality Supervision Inspection and Quarantine Administration and the Standardization Administration of China. The standard is titled '干部人事档案数字化技术规范' (Technical Specification for Digitalization of Party and Personnel Archives). The implementation date is 2017-07-01.

序号	标准号	标准名称	代替标准号	实施日期
1	GB/T 33870-2017	干部人事档案数字化技术规范		2017-07-01

## 1 范围

本标准规定了干部人事档案数字化人员建库、目录建库、档案扫描、图像处理、数据存储、数据验收、数据交换、数据备份、安全管理的技术要求。

本标准适用于各级党政机关、国有企事业单位的干部人事档案数字化工作。

## 2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件，仅注日期的版本适用于本文件。凡是不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 2261.1 个人基本信息分类与代码 第1部分：人的性别代码

GB/T 3304 中国各民族名称的罗马字母拼写法和代码

GB/T 7408 数据元和交换格式信息交换日期和时间表示法

GB 11643 公民身份号码

GB/T 18788 平板式扫描仪通用规范

SJ/T 11292 计算机用液晶显示器通用规范

全国公务员管理信息系统信息采集、报送标准组通字[2012]31号

### 3 术语和定义

下列术语和定义适用于本文件。

#### 3.1

干部人事档案数字化 digitalization of cadre personnel  
archives

采用扫描仪等设备对干部人事档案进行数字化加工，将其转化为可存储在磁盘、光盘等存储介质上，并能被计算机识别，数字方式可信、可取和可用的数字图像或数字文本的处理过程。

#### 3.2

档案信息 archival information

干部人事档案中的文字、表格、照片、印章等有效信息。

#### 3.3

数字档案管理信息系统 information system for digitalized  
archives management

具有信息录入、档案扫描、图像处理、提供利用、存储备份、导入导出等功能的信息系统。

#### 3.4

原始图像数据 the original image data

纸质档案通过扫描等方式形成数字图像后，再经过纠偏、裁边等处理后的图像数据。

### 3.5

优化图像数据 optimized image data

与原始图像数据一致. 经优化处理达到无噪点干扰、字符灰度平均值 $\leq 125.0$ . 模糊度 $\leq 200.0 \mu\text{m}$ 的图像数据。

### 3.6

图像数据压缩 image data compression

清除图像数据冗余的一种过程。

### 3.7

背景区域 background area

衬托档案信息的区域。

### 3.8

图像背景颜色 background color

背景区域所呈现的颜色。

### 3.9

背景无关标记 background extraneous mark

背景区域中与档案信息无关的噪点。

### 3.10

灰度 gray scale

图像每个像素点的颜色深度。

### 3.11

字符灰度 character gray scale

优化图像数据档案信息区域中像素点的灰度，灰度值的范围为0~255, 对应图像中的颜色为从黑到白。

注：字符灰度评价指标见附录 A。

### 3.12

模糊度 blurriness

图像中线条（或字符笔画）边缘内边界和外边界的平均距离值。

注：模糊度评价指标见附录 A。

### 3.13

持续数据保护 continuous data protection

一种在不影响主要数据运行的前提下，可以实现持续捕捉或跟踪目标数据所发生的任何改变，并且能够恢复到此前任意时间点的方法。

## 4 基本要求

### 4.1 基本原则

#### 4.1.1 真实性

应确保干部人事档案数字化后的内容与纸质档案在内容上一致。确保档案数字化过程中档案信息不被更改。原始图像应保留原纸张颜色、污损情况和文字修改痕迹等原始信息。

#### 4.1.2 完整性

应确保数字化前后纸质档案一致，档案数字图像数量与纸质档案数量相符。

#### 4.1.3 可用性

应确保数字档案可被查找、检索、呈现等，满足相关业务的要求。确保数字档案的连续性，维护其可跟踪、可回溯、可关联、可被发现和可被再用，数据链不出现断裂。

#### 4.1.4 安全性

应建立身份认证体系、加密存储体系及数据流传输方式等安全保密管理机制，确保档案信息的安全。干部人事档案数字化过程应完整记录，可查询、可追溯。档案图像数据应得到有效保护，不被非法利用、更改或销毁。档案原件不受损毁。

### 4.2 数字化对象的确定

数字化对象应是符合干部人事档案审核、整理等有关规定要求的合

表 1 人员基本信息集

项目	编码	类型	长度	必填	说明
人员唯一标识	A0100	字符	36	是	人员唯一标识,从其他系统获取或由软 统自动生成,符合全球唯一标识符要求
姓名	A0101	字符	36	是	
性别	A0104	字符	1	是	符合 GB/T 2261.1 要求的代码
民族	A0117	字符	2	是	符合 GB/T 3304 要求的代码
出生日期	A0107	字符	8	是	符合 GB/T 7408 要求的格式
公民身份号码	A0184	字符	18	是	符合 GB 11643 要求的身份代码

格档案。

### 4.3 基本环节

干部人事档案数字化基本环节包括：人员建库、目录建库、档案扫描、图像处理、数据存储、数据验收、数据交换、数据备份。其中，图像处理环节包括原始图像处理和优化图像处理,优化图像处理为可选处理环节。

## 5 人员建库

### 5.1 人员基本信息集

人员基本信息集的编码为:A01。用于描述某人自然属性和社会属性中最基本的信息。编码引自组通字[2012]31 号文件。

该信息集为单记录信息，每一记录对应一个人。至少由 6 个信息项(见表 1)组成，其他信息项可根据实际情况增加。该信息集内容可单独建立，也可从其他系统中获取。

### 5.2 人员信息录入

人员基本信息录入内容包括：

- a) “姓名”，填写户籍登记所用的姓名；
- b) “性别”，填写 GB/T 2261.1 中的性别代码；
- c) “民族”，填写 GB/T 3304 中的民族代码；
- d) “出生日期”，按照 GB/T 7408 要求的格式填写出生年月；
- e) “公民身份号码”，填写公安机关为公民编制的符合 GB 11643

要求    信息 录记	类号	RSDAML002	字符	3	是	表示为一到十,其中第四类中的1到4-4表示,第九类中的小类用9-1示,类号信息集见附录B	
	序号	RSDAML003	字符	3	是	该份档案材料在所属分类中的排列	
	材料名称	RSDAML004	字符	120	是	材料已有或代拟的标题名称	
	材料形成时间	年	RSDAML006	字符	4	否	材料形成时间的年份,格式 YYYY
		月	RSDAML007	字符	2	否	材料形成时间的月份,格式 MM
		日	RSDAML008	字符	2	否	材料形成时间的日,格式 DD
	页数	RSDAML005	数值	3	是	有文字或图表的材料页数	
	备注	RSDAML009	字符	120	否	材料的备注说明	

## 6.2 目录录入

应根据统一的干部人事档案目录格式进行档案目录录入,包括:

- a) “类号”,填写材料类号;
- b) “序号”,填写材料所属分类中的序号;
- c) “材料名称”,根据材料题目填写,无题目的材料,应拟定题目;
- d) “材料形成时间”,一般采用材料落款标明的最后时间,复制的档案材料,采用原材料形成时间;
- e) “页数”,填写每份材料的页码数;
- f) “备注”,填写需要说明的情况。

## 7 档案扫描

### 7.1 扫描仪校准

扫描仪应符合 GB/T 18788 的规定,亮度和对比度为中值,无偏移。

### 7.2 扫描方式

应根据纸质档案材料的具体情况,采用合理的扫描方式进行扫描。

扫描方式包括(但不限于):

- a) 大幅面档案宜采用大幅面扫描仪扫描,也可采用小幅面扫描后的图像拼接方式处理;
- b) 纸张状况较差、容易损坏的档案,应采用平板扫描方式;

c) 对于纸张较薄的档案, 若扫描时发生背页字迹透印而影响图像阅读的现象, 应在背页后垫白色衬底扫描。

### 7.3 扫描色彩模式

应采用真彩色 24 位 RGB 模式扫描。

### 7.4 扫描分辨率

应采用 300 dpi 分辨率扫描。

## 8 图像处理

### 8.1 原始图像处理

#### 8.1.1 纠偏处理

对偏斜度大于  $1^{\circ}$  的图像应进行纠偏处理, 纠偏后距离显示器 25 cm ~ 40 cm 观看图像应没有明显偏斜。对方向不正确的图像应旋转还原。

#### 8.1.2 裁边处理

纠偏后的图像应进行裁边处理, 去除扫描过程中产生的白边或黑边。

#### 8.1.3 图像拼接

对大幅面干部人事档案材料进行分区扫描形成的多幅图像, 应进行拼接处理, 合并为一个原始图像, 确保档案数字图像的完整性和真实性。

### 8.2 优化图像处理

#### 8.2.1 优化图像处理方式

应使用计算机软件或人工方式对原始图像数据进行处理, 得到优化图像数据。

原始图像中带有黑白、彩色照片的, 应采用人工方式处理, 以使优化图像中的照片得到更佳的视觉效果或与原始图像的视觉效果相符。

原始图像中带有印章的图像区域应采用人工方式处理。

带有身份证、学历证件、复印件及带有防伪技术的原始图像, 或内容不清晰的红色或紫色背景的早期麻纸基材等, 优化处理后仍不具有较

好辨识效果的，应直接引用原始图像数据。

### 8.2.2 优化图像质量要求

观看优化图像使用的显示器应符合 SJ/T 11292 的要求，并调整到 sRGB 状态。在保证档案信息完整的前提下应满足以下指标：

a) 图像背景颜色为 24 位真彩色，且 RGB 值为 (254, 246, 197) ；

b) 4X4 pixels (对应 300 dpi 分辨率的情况约为 328  $\mu$ mX328  $\mu$ m) 范围以下的背景无关标记为 0 个, 4X4 pixels 到 8X8 pixels (对应 300 dpi 分辨率的情况约为 656  $\mu$ mX656  $\mu$ m) 范围内的背景无关标记以不影响档案信息阅读为准, 8X8 pixels 范围以上的背景无关标记为 0 个；

注：大于 8X8 pixels 范围是能够分辨出该区域是否为有效信息，小于 4X4 pixels 范围发现有信息则认定为无效信息，4X4 pixels 到 8X8 pixels 之间范围内是无法分辨信息是否为有效信息，如标点符号或污点等。

c) 字符灰度平均值  $\leq 125.0$ ；

d) 模糊度  $\leq 200.0 \mu$ m。

## 9 数据存储

### 9.1 图像存储格式

图像数据应采用 JPEG 格式存储。原始图像数据存储时，应进行图像数据压缩，压缩率  $\geq 80\%$ 。优化图像数据不进行压缩。

原始图像数据和优化图像数据应分不同文件夹保存。

### 9.2 图像文件存储路径

干部人事档案图像文件在存储介质上的实际存储路径。

原始图像文件存储路径：系统指定路径人员唯一标识〈原始图像数据〉

优化图像文件存储路径：系统指定路径人员唯一标识〈优化图像数据〉

案材  
信息  
件的

项目	编码	类型	长度	必填	说明
目录挂接标识	RSDAWJ000	字符	36	是	该标识是图像文件挂接到对信息,从 RSDAML000 获取
页序号	RSDAWJ001	字符	3	是	该页在该份档案材料的序号
原始文件名	RSDAWJ003	字符	20	是	原始图像文件在存储介质上的规则为三位流水号
优化文件名	RSDAWJ004	字符	20	否	优化图像文件在存储介质上的规则为三位流水号

## 10 数据验收

### 10.1 验收范围

应对全部数字档案成品进行质量验收。

### 10.2 验收标准

#### 10.2.1 人员建库验收标准

人员建库时录人项目应符合 5.1 所列人员基本信息集数据格式要求;与其他信息系统共享人员信息的,所引用的姓名、性别、民族、出生日期、公民身份号码等信息项也应符合人员建库要求。

#### 10.2.2 目录建库验收标准

目录建库时录入项目应符合 6.1 所列干部人事档案目录信息集数据格式要求,类号、序号、材料名称、材料形成时间、页数、备注等录入内容项应规范、准确,档案目录应与纸质档案材料名称内容相符,档案目录条数应与纸质档案材料数量相符。

#### 10.2.3 原始图像数据验收标准

原始图像应为真彩色 24 位、JPEG 格式、300 dpi。无扫描产生的白边和黑边;偏斜度 $\leq 1^\circ$ ,方向正确;图像页码连续;图像中档案信息与纸质档案一致,图像数量与纸质档案一致,图像的排列顺序与纸质档案一致,图像与目录一一对应。

#### 10.2.4 优化图像数据验收标准

优化图像应符合 8.2.2 的要求,与原始图像尺寸一致。

### 10.3 验收登记

应根据验收标准填写相应的验收登记表（样式见附录 C），对数字档案作出合格或不合格的结论。

## 11 数据交换

### 11.1 数据交换内容

数据交换的内容应包括：

a) 人员基本信息及目录信息的描述性文件，以 XML 格式文件保存，符合 5.1、6.1 的要求，其中，人员基本信息只交换表 1 内的 5 个信息项；

b) 干部人事档案数字化形成的图像数据，存放在指定的文件夹内。

### 11.2 数据交换格式

#### 11.2.1 XML 文件格式描述

XML 根节点为〈数字档案〉。整个 XML 文件包括人员基本信息及目录信息，分别用〈人员基本信息〉、〈目录信息〉进行标记。人员基本信息中的内容分别用〈姓名〉、〈性别〉、〈民族〉、〈出生日期〉、〈公民 身份号码〉进行标记。目录信息中的内容分别用〈类号〉、〈序号〉、〈材料名称〉、〈材料形成时间〉、〈页数〉、〈备注〉、〈原始图像数据〉（即图像文件名称）、〈优化图像数据〉（即图像文件名称）进行标记。XML 文件的定义描述见附录 D。

#### 11.2.2 文件夹及文件命名规范

每个人的干部人事档案图像数据应放在一个文件夹内，该文件夹应以姓名+公民身份号码命名（如 李\* \* 11010219830606 \* \* \* \* ）。根目录下保存的是一个以姓名+公民身份号码命名的 XML 文件 和一个文件名为“图像数据”的文件夹，前者描述人员基本信息及目录信息，后者存放原始图像数据和优化图像数据。

图像文件名称的命名规则为：干部人事档案材料类号代码+ “一” +

分类  
操作  
录结

注：圆角方框表示文件夹。

图 1 目录结构图

## 12 数据备份

### 12.1 备份范围

经验收合格的人员基本信息、干部人事档案目录信息、原始图像数据、优化图像数据应及时进行备份。

### 12.2 备份方式

应采用在线、离线相结合的方式实现备份。在线备份可采用持续数据保护技术，保障数据备份的持续性。离线备份时，每个人的数字档案应形成一个数据包，数据包应采用硬盘、光盘等不同载体备份，宜异地备份。

### 12.3 备份标签

离线备份后应在相应的备份介质上做好标签，以便查找和管理。

## 13 安全管理

### 13.1 密级

13.1.1 数字档案的密级和相应纸质档案的密级相同。

13.1.2 数字档案管理信息系统（以下简称“管理系统”）和承载管理系统的网络的密级应不低于其所承载的数字档案的密级。

### 13.2 环境安全

13.2.1 数字化加工场所应设在独立、可封闭的建筑内，应符合防盗、防火、防尘、防水、防潮、防高温、防日光及紫外线照射、防有害生物、防污染等安全管理要求。

13.2.2 应配备满足安全管理需要的视频监控设备，配备符合国家标准并满足工作需要的档案装具。

13.2.3 应配备数字化工作人员存放随身物品的专用储物箱柜，并与档案装具分区放置。

### 13.3 设备安全

13.3.1 干部人事档案数字化加工设备（计算机、打印机、复印机、扫描仪等）的使用、管理、维修、报废等应符合涉密信息设备使用保密管理的有关规定。

13.3.2 与干部人事档案数字化加工相关的计算机、打印机、复印机、扫描仪等设备不得连接互联网及其他公共信息网络。

### 13.4 身份鉴别

13.4.1 应建立以数字证书为核心的身份认证系统，按照密级选择安全保密方式实现用户身份鉴别。

13.4.2 管理系统应提供专用的登录控制模块对登录用户进行身份标识和鉴别，标识范围应涵盖管理系统的用户。

13.4.3 管理系统不得设置匿名账户。

13.4.4 管理系统应具备用户身份标识和鉴别信息复杂度检查功能，确保管理系统中不存在重复身份标识、身份鉴别信息不被冒用。

13.4.5 单个管理系统客户端在同一时间仅允许单个用户登录，用户在同一时间仅允许登录单个管理系统客户端。

### 13.5 访问控制

13.5.1 管理系统应具备涉密信息和重要信息的访问控制功能。

13.5.2 涉密信息和重要信息的访问控制，主体应控制到单个用户，客体应控制到单个页面。

### 13.6 安全审计

13.6.1 干部人事档案数字化应进行全过程日志记录，修改和删除的图像数据应一并存入日志记录。

13.6.2 日志记录的内容应包括事件发生时间、发起者信息、类型、

描述和结果等。

13.6.3 管理系统应具备安全事件审计功能，审计范围应覆盖所有用户。

13.6.4 审计事件的类型应包括系统事件、业务事件、成功事件、失败事件等。

13.6.5 管理系统应不具备单独中断审计进程功能，不能非授权删除、修改或覆盖日志记录。

13.6.6 管理系统应具备对日志数据进行自动或手动备份功能。

### 13.7 权限划分

13.7.1 管理系统应设置系统管理员，系统管理员仅负责系统级的管理，不具备任何用户业务操作的权限。

13.7.2 管理系统应设置安全保密管理员，负责管理系统的日常安全保密管理工作，包括对用户账号权限管理、安全保密设备管理和管理系统所产生日志的审查分析，不具备任何用户业务操作的权限。

13.7.3 管理系统应设置安全审计员，负责对系统管理员、安全保密管理员的操作行为进行审计跟踪分析和监督检查，不具备任何用户业务操作的权限。

13.7.4 系统管理员、安全保密管理员、安全审计员应相互独立、相互制约，实行“三员分离”。

### 13.8 数据保护

13.8.1 管理系统应采用加密技术确保存储和传输过程中数据的完整性。

13.8.2 数据在网络传输和存储中应采取相应的密码保护措施，确保数据传输和存储的安全。

13.8.3 管理系统应具备涉密信息完整性检测功能，能发现信息被篡改、伪造、删除等情况，并产生审计日志。

## 13.9 资质选择

13.9.1 采用外包方式开展干部人事档案数字化工作的，涉密系统集成、系统咨询、软件开发、安防监控、运行维护、数据恢复等业务应选择具有涉密信息系统集成资质的单位，涉密档案数字化加工等业务应选择具有国家秘密载体印制资质的单位。

13.9.2 管理系统使用之前应委托有资质的第三方信息安全测评机构开展风险评估和检测。

## 附录 A

(规范性附录)

灰度、字符灰度、模糊度评价指标

### A.1 灰度评价指标

灰度的评价指标用  $Y(x, y)$  表示，按式 (A.1) 计算：

$$Y(x, y) = 0.3R(x, y) + 0.5G(x, y) + 0.2B(x, y) \dots\dots\dots$$

(A.1)

式中：

$Y(x, y)$ ——灰度；

$R(x, y)$ 、 $G(x, y)$ 、 $B(x, y)$ ——彩色图像  $(x, y)$  位置处像素的数字颜色数值。

注：改写 ISO/IEC TS 24790，定义 5.3.4。

### A.2 字符灰度评价指标

本标准使用整幅图像的字符灰度平均值评价优化图像字符的灰特征。字符灰皮平均值反映优化图像数据档案信息区域字符的整体明暗水平。各灰度关系如下：

整幅图像中档案信息灰度总和记为  $Y_{sum}$ ，按式 (A.2) 计算：

$$Y_{sum} = \sum Y(x_i, y_i) \dots\dots\dots (A.2)$$

式中：

$i$  ——档案信息像素点；

$Y(x_i, y_i)$  ——由式 (A.1) 计算的第  $i$  个档案信息像素点的灰度值。

字符灰度平均值记为  $Y$ ，按式 (A.3) 计算：

$$Y = Y_{\text{sum}} / P_{\text{sum}} \dots \dots \dots (A.3)$$

式中：

$P_{\text{sum}}$  ——整幅图像中档案信息像素的总和。

### A.3 模糊度评价指标

从优化图像中选取含有一段边缘较模糊的直笔画及其部分背景的区域，计算该区域各笔画垂线的灰度最大值  $Y_{\text{max}}$  和灰度最小值  $Y_{\text{min}}$ ，并按式 (A.4) 和式 (A.5) 求得笔画两侧各自的边缘阈值  $Y_i$  和  $Y_u$ ：

$$Y_i = Y_{\text{min}} + 70\% (Y_{\text{max}} - Y_{\text{min}}) \dots \dots \dots (A.4)$$

$$Y_u = Y_{\text{min}} + 10\% (Y_{\text{max}} + Y_{\text{min}}) \dots \dots \dots (A.5)$$

求取笔画两侧各自  $Y_i$  和  $Y_u$  对应的位置坐标值，笔画相应侧边缘的模糊度记为  $B$ ，按式 (A.6) 计算：

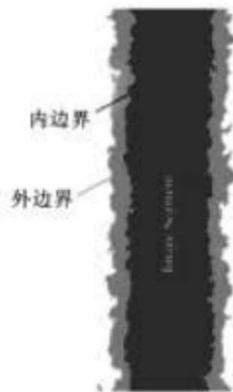
$$B = \text{DIS}_i - u(Y_{\text{max}} - Y_{\text{min}})$$

笔画的模糊度为

注 1：改写 ISO

注 2：如图 A.1

示，该线条两侧边缘即该过渡带的平均



对象) 的放大图  
过渡带，模糊度

图 A.1 线条边缘模糊特性示意图

## 附录 B

(规范性附录)

类号信息集

干部人事档案内容应分为十类：其中第四类和第九类分为 4 小类。



```
<xs:sequence>
  <xs:element name="姓名
"type="xs:string" minOccurs="0"/><xs:element name="性别
"type="xs:string" minOccurs="0"/>
  <xs:element name="民族" type="xs:string" minOccurs="0"/>
  <xs:element name="出生日期
"type="xs:string" minOccurs="0"/>
  <xs:element name="公民身份号码
"type="xs:string" minOccurs="0"/>
</xs:sequence>
</xs:complexType>
</xs:element>
<xs:element name=" 目录信息">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="档案目录条目" minOccurs="0"
maxOccurs="unbounded">
        <xs:complexType>
          <xs:sequence>
            <xs:element name="类号" type="xs:string" minOccurs="0"/>
            <xs:element name="序号" type="xs:string" minOccurs="0"/>
            <xs:element name="材料名称" type="xs:string" minOccurs="0"/>
            <xs:element name="材料形成时间
"type="xs:string" minOccurs="0"/>
            <xs:element name="页数" type="xs:string" minOccurs="0"/>
            <xs:element name="备注" type="xs:string" minOccurs="0"/>
          </xs:sequence>
        </xs:complexType>
      </xs:element>
    </xs:sequence>
  </xs:complexType>
</xs:element>
```



</xs:element>

</xs:schema>

来源：2020-05-23 兰台之家微信公众号