

档案数字化加工软件 使用说明（通用版）

依据 DA/T 31—2017《纸质档案数字化规范》编制
适用于纸质档案扫描、图像处理、著录挂接与成果移交全流程

编制单位：_____

文档版本：V1.0 编制日期：_____年__月

目 录

一、软件概述	3
1.1 核心功能	3
1.2 适用对象	3
二、运行环境（通用要求）	3
三、总体加工流程	4
四、详细操作流程	5
4.1 建立项目与参数设置	5
4.2 数字化前处理	5
4.3 建立目录数据库（著录）	6
4.4 档案扫描	6
4.5 图像处理	6
4.6 数据挂接	7
4.7 数据校验与质量检查	7
4.8 成果验收与移交	7
五、关键技术参数（速查）	9
5.1 扫描色彩模式	9
5.2 扫描分辨率	9
5.3 存储格式	9
六、文件命名与目录结构规范	10
七、数据安全性与保密事项	10
八、常见问题与注意事项	10
附录：常用工作文件清单	11

一、软件概述

档案数字化加工软件是档案数字化项目的核心生产工具，用于将纸质（及部分非纸质）档案转化为规范、可检索、可长期保存的数字成果。软件覆盖从档案出库登记、扫描采集、图像处理、目录著录、数据挂接，到质量检查与成果移交的完整流程，并对全过程进行任务管理与质量管控。

本说明为通用版本，描述档案数字化加工软件的标准功能与作业流程，不绑定具体软件品牌与版本。各机构可结合实际使用的软件功能菜单进行对照执行。

1.1 核心功能

- 项目与任务管理：建立数字化项目，配置扫描参数、著录字段、命名规则、人员与权限。
- 扫描采集：对接平板扫描仪、高速馈纸扫描仪、书刊扫描仪、缩微/古籍专用设备，支持批量采集。
- 图像处理：自动纠偏、裁边、去黑边、去污点、方向校正、空白页剔除、图像拼接等。
- 目录著录：按 DA/T 18《档案著录规则》建立目录数据库，支持批量录入、导入与校验。
- 数据挂接：将目录条目与对应数字图像建立关联关系，支持自动与手动挂接。
- 质量检查：分辨率、色彩模式、完整性、清晰度、命名规范等多维度自动 + 人工抽检。
- 成果输出与移交：按规范格式导出图像、目录、元数据及过程文件，生成验收与移交清单。

1.2 适用对象

各级各类档案馆、机关企事业单位档案室，以及承接档案数字化业务的加工服务机构。适用于文书档案、科技档案、专门档案等各类纸质档案的数字化加工。

二、运行环境（通用要求）

以下为档案数字化加工软件的一般性运行环境要求，具体以所用软件的安装手册为准。

项目	推荐配置
操作系统	Windows 10/11（64 位）；服务器端可采用 Windows Server 或主流 Linux 发行版
处理器	Intel i5 及以上 / 同等性能处理器；批量图像处理建议 i7 或多核服务器
内存	8 GB 起步，批量扫描与图像处理建议 16 GB 及以上
存储	本地高速磁盘用于在制数据；成果数据建议接入网络存储（NAS）或服务器阵列，并预留备份空间
扫描设备	支持 TWAIN/ISIS 接口的扫描仪；幅面、速度、色彩按档案类型选型
网络	加工区域建议物理隔离或专网；涉密档案须断网作业，符合保密要求
数据库	单机版可用内置数据库；多人协作版需部署 MySQL / SQL Server 等数据库服务

三、总体加工流程

档案数字化加工遵循 DA/T 31—2017《纸质档案数字化规范》规定的基本环节，软件按下列流程顺序组织各项作业：

序	环节	主要内容
1	档案出库	履行出库登记手续，清点核对，记录档案数量与状态。
2	数字化前处理	确定扫描页、编制页号、目录数据准备、拆除装订、技术修复。
3	目录数据库建立	按著录规则建立目录数据库，确定字段与内容要求。
4	档案扫描	按色彩模式、分辨率、存储格式要求采集图像。
5	图像处理	拼接、纠偏、裁边、去污，并进行图像质量检查。
6	数据挂接	建立目录数据与数字图像的关联关系。
7	成果验收与移交	对图像、目录、元数据、过程文件等进行验收并移交。
8	档案归还入库	恢复装订，清点核对，办理归还入库手续。

说明：各环节环环相扣，前道工序质量直接影响后道。软件对每一环节设置状态标记与质检节点，未通过质检的批次不得流转至下一环节。

四、详细操作流程

4.1 建立项目与参数设置

1. 新建项目：填写项目名称、立项单位、加工范围、起止时间等基本信息。
2. 配置扫描参数：设定默认色彩模式、分辨率、存储格式（详见第五章技术参数）。
3. 设置命名规则与目录结构：按档号规则配置图像文件命名模板（详见第六章）。
4. 配置著录字段：依据 DA/T 18 设定目录字段、字段长度与录入校验规则。
5. 分配人员与权限：按扫描、图像处理、著录、质检等岗位分配账号与操作权限。

注意事项：

- 同一批档案应采用统一的色彩模式、分辨率与存储格式，避免标准不一致。
- 命名规则一经确定，加工过程中不应随意更改，以免造成图像与目录错乱。

4.2 数字化前处理

前处理的核心原则：一是建议全部扫描、不挑扫，确保信息完整、避免二次扫描；二是以保护档案原件为前提。

(1) 确定扫描页：原则上对档案进行全部扫描，确保档案信息完整。

(2) 编制页号：使用阿拉伯数字从“1”开始连续编写；同一案卷或文件内页号应保持连续。

(3) 目录数据准备：核对案卷题名、文件题名、责任者、起止页号、页数等，发现错误或不规范的予以更正。

(4) 拆除装订：以保护档案为原则确定是否拆除；拆订时避免损伤档案。特殊装订且需恢复的，应先拍照记录原貌。

(5) 技术修复：破损严重无法直接扫描的先行修复；折皱不平影响扫描质量的，先压平或烫平后再扫描。

软件操作：

- 在“前处理登记”模块填写并打印《数字化加工交接登记表》，记录每份文件的起始页号与页数。

- 对需特殊处理（修复、特殊装订）的档案在系统中标记，便于后续追溯与恢复。

4.3 建立目录数据库（著录）

按 DA/T 18《档案著录规则》规范目录内容，确定著录项、字段长度与内容要求。

1. 录入或导入目录：支持手工逐条录入，或从已有电子目录（Excel、原有系统）批量导入。
2. 字段校验：系统按预设规则校验必填项、字段长度、日期格式、档号唯一性等。
3. 查重与纠错：对档号、题名等进行查重，对不规范条目集中修改。

注意事项：

- 目录数据应与档案实体内在联系一致，保持案卷—文件—页的层级关系。
- 涉及学生、人事等个人信息的档案，著录与存储须符合《个人信息保护法》相关要求，严格控制访问权限。

4.4 档案扫描

根据档案幅面与状态选择合适设备，按既定参数采集图像。

1. 选择设备与批次：建立扫描批次，关联对应案卷/文件。
2. 设置参数：确认色彩模式、分辨率、存储格式（详见第五章）。
3. 扫描采集：批量扫描，实时预览，发现漏扫、重扫、倾斜的及时补扫或重扫。
4. 即时初检：检查页数与实体是否一致、有无缺页、图像是否清晰完整。

注意事项：

- 严格做到“一份档案只动一次”，一次扫描满足多重利用需求，减少对原件的反复翻动。
- 扫描过程随时核对页号与图像数量，确保不缺页、不串页、不重页。

4.5 图像处理

在保持档案原貌的前提下，对图像进行规范化处理，提升清晰度与可用性。

处理项	操作要求
图像拼接	对超出扫描幅面、需分次扫描的大幅档案进行拼接还原。

处理项	操作要求
旋转及纠偏	校正图像方向，纠正扫描产生的倾斜，使文字水平端正。
裁边	去除图像周边多余黑边、白边，保留档案完整内容，不得裁切到正文。
去污	去除扫描产生的污点、噪点，但不得损失档案原有信息。
质量检查	检查清晰度、完整性、偏斜度、失真等，剔除空白页、纠正方向错误。

注意事项：图像处理应以真实反映档案内容、最大程度保持原貌为原则，不得过度处理而改变档案原始信息。

4.6 数据挂接

数据挂接是将目录条目与对应数字图像建立准确关联的关键环节，直接决定成果能否被正确检索与调阅。

1. 自动挂接：依据统一的命名规则，由软件按档号自动匹配目录与图像。
2. 人工核对：对自动挂接结果逐条或抽样核对，确认目录与图像一一对应。
3. 异常处理：对挂接失败、多挂、漏挂的条目进行人工修正。

注意事项：

- 挂接前须确保命名规范统一，否则会出现批量错挂。
- 挂接完成后应随机抽查，验证“点目录能调出正确图像、图像页数与目录页数一致”。

4.7 数据校验与质量检查

软件提供自动检测与人工抽检相结合的质检手段。能够通过程序自动检测的指标，应实现 100% 检测；自动检测项目要求 100% 合格。

- 自动检测：分辨率、色彩模式、存储格式、文件命名、图像数量、是否空白页等。
- 人工抽检：图像清晰度、完整性、纠偏裁边效果、著录准确性、挂接正确性等。
- 问题反馈：建立问题反馈机制，对各环节问题及时退回相应工序修正，并记录处理过程。

4.8 成果验收与移交

验收内容主要包括：数字图像、档案目录数据、元数据、过程中产生的工作文件、存储

载体等。

1. 成果导出：按规范格式导出数字图像、目录数据、元数据及验收/移交清单。
2. 验收检查：按验收指标逐项核验，形成验收结论；不合格批次退回整改后复验。
3. 成果移交：填写《数字化成果移交清单》，与接收方清点核对、签字确认。
4. 归还入库：拆订档案恢复装订，清点核对后办理归还入库手续。

五、关键技术参数（速查）

下列参数依据 DA/T 31—2017 整理，为通用最低要求。具体项目可结合档案珍贵程度、利用需求等适当提高。

5.1 扫描色彩模式

色彩模式	适用情形
彩色	页面含红头、印章、照片、彩色插图、多色彩文字等的档案——应采用彩色模式（强制）。批量数字化时建议尽量采用彩色，最大程度保留原件信息。
灰度	页面为黑白两色，但字迹清晰度差或带有插图的档案，可采用灰度模式。
黑白二值	页面为黑白两色、字迹清晰、不带插图的档案，可采用黑白二值模式。

5.2 扫描分辨率

利用方式	分辨率要求
一般扫描	不小于 200 dpi
文字偏小、密集、清晰度较差	建议不小于 300 dpi
COM 缩微输出	建议不小于 300 dpi
高精度仿真复制	建议不小于 600 dpi
印刷出版	结合幅面与印刷精度选择合适分辨率

提示：分辨率并非越高越好；超过一定值后清晰度提升不明显，却会大幅增加文件体积与存储成本。应在清晰度与文件大小间取得平衡。

5.3 存储格式

用途	推荐格式
长期保存	TIFF、JPEG、JPEG2000 等通用格式；压缩率按实际需求选择。
利用 / 网络浏览	可从浏览速度、易操作性、存储占用考虑，转换为 OFD、PDF 等格式。

要求：同一批档案应采用相同的存储格式；遵循“对原件只动一次”原则，优先满足长期安全保存需要。

六、文件命名与目录结构规范

规范的命名与目录结构是数据挂接、检索与长期管理的基础。命名规则应在项目开始前确定并固化。

- 以档号为基础命名：图像文件名应与档号、页号建立明确对应关系，确保唯一、可追溯。
- 连续与补零：同一文件内页号连续编排，建议统一位数（如 0001、0002），便于排序。
- 分级建夹：按全宗号—目录号—案卷号—文件号逐级建立文件夹，保持与实体层级一致。
- 避免特殊字符：文件名与文件夹名避免使用空格、斜杠、引号等特殊字符。

示例（仅供参考，以本机构档号规则为准）：

全宗号-目录号-案卷号-页号.tif 例：J001-01-0005-0012.tif

七、数据安全与保密事项

- 场所管理：加工场所应配备门禁、监控等安防设施，人员出入登记，非工作人员来访登记。
- 设备管理：扫描与存储设备专机专用；加工区计算机原则上断网或专网作业。
- 涉密档案：涉密档案数字化须符合涉密档案保密要求，在符合保密条件的环境中作业。
- 个人信息：涉及学生、人事等个人信息的数据，严格遵守《个人信息保护法》，落实最小必要、权限分级与访问留痕。
- 数据备份：成果数据应及时备份，建议采用多副本/异地备份策略，防止单点丢失。
- 外包管理：如采用外包方式，应审查加工企业资质与保密能力，签订保密协议，并按相关安全管理规范执行全程管控。
- 过程文件管理：工作方案、流程单、验收单、移交清单等过程文件应规范整理、归档、移交，做到全程可追溯。

八、常见问题与注意事项

常见问题	处理建议
扫描后图像偏色、发灰？	检查色彩模式与设备校色；含红头印章的必须用彩色模式，避免信息丢失。
文字模糊、笔画断裂？	提高分辨率（文字密集/偏小建议 ≥ 300 dpi）；检查原件是否需先行压平、修复。
图像与目录对不上？	核对命名规则是否统一；重新执行数据挂接并抽样验证页数一致性。
出现缺页 / 串页 / 重页？	回到扫描批次核对实体页数与图像数量，补扫或删除多余页后重新质检。
文件体积过大、存储紧张？	在满足清晰度前提下合理设置分辨率与压缩率；利用副本可转 PDF/OFD。
质检不通过如何处理？	按问题反馈机制退回对应工序整改，记录处理过程，复检合格后方可流转。

附录：常用工作文件清单

数字化加工过程中通常需形成并归档以下工作文件（采用外包方式的还应包括招标文件、合同、保密协议等）：

- 数字化工作方案 / 实施方案
- 数字化加工交接登记表（含起始页号、页数记录）
- 数字化流程单 / 工序流转单
- 数据验收单 / 质量检查记录
- 项目验收报告
- 数字化成果移交清单

本说明依据 DA/T 31—2017《纸质档案数字化规范》编制，供数字化加工作业参考。

具体技术指标与操作以现行有效标准及所用软件实际功能为准。