

江西省纸质档案数字化管理与技术规范

为规范我省纸质档案数字化工作，保障纸质档案数字副本凭证效力，维护档案实体与档案信息安全，根据《中华人民共和国档案法》以及《文献档案资料数字化工作导则》（GB/T20530-2006）、《纸质档案数字化技术规范》（DA/T31-2005）、《信息与文献—文件档案数字化实施指南》（ISO/TR 13028: 2010（E））等相关标准，制定本规范。

1 范围

本规范规定了纸质档案源版数字副本制作的各项要求，源版数字副本的衍生品制作可参照执行。

本规范所称纸质档案数字副本是源版数字副本及其衍生品的统称。源版数字副本是指使用扫描仪等数字成像设备对纸质档案进行数字化转换而直接生成的静态图像文件。对源版数字副本进行压缩、格式转换、识别等处理而生成的静态图像文件、版式文件或纯文本文件等，为源版数字副本的衍生品。

本规范适用于各级国家综合档案馆对馆（室）藏纸质档案进行批量数字化转换，可作为编制数字化工作方案的依据。机关、团体、企业事业单位纸质档案（含纸质照片档案）的数字化可参照执行。

2 基本原则

2.1 统筹规划，分步实施

应根据馆藏档案重要程度、破损情况、利用需求等，进行统筹规划、分步实施数字化。馆藏国家重点档案、重要全宗档案和涉及民生的专业

档案优先进行数字化。

2.2 保障凭证作用，维护档案安全

为保证纸质档案数字副本是档案原件真实、完整、可靠与可用的复制件，必须在纸质档案数字化过程中实施全程安全管理。应建立安全管理措施，对数字化对象实行审批制，数字副本的制作应符合相关标准的要求，采用适合的数字化技术，执行规定的质量控制程序，准确复制档案原件，并捕获必要的数字化元数据。数字副本制作各环节应有责任人的签字，数字化成果按规范要求移交，确保档案原件与档案信息安全。

对进行数字化的档案原件，在拆装、整理、修复等过程中，应遵循最小干预、保持原貌以及适宜性、相似性与可逆性原则，最大限度减少数字化转换对档案原件可能造成的损伤或形制变化。**未解密档案不得采用委托加工方式进行数字化。**采取数字化服务与数字化系统设备整体采取委托加工方式的，必须对承包单位进行资格审查，并采取使用国产计算机产品、硬盘整体验收移交或低级格式化、加强数字化现场安全管理等措施，防止档案原件数字副本外流。

2.3 加强协调，夯实基础

纸质档案数字化是一项系统工程，应加强协调、合理设计流程，把规范档案基础业务工作与档案数字化转换相结合，在档案数字化过程中进一步完善档案原件的整理、编目、著录、修复、还原等工作，夯实档案基础业务工作。

3 纸质档案数字化项目管理

3.1 数字化项目的全程管理

应对纸质档案数字化项目的准备、前处理、数字化转换、后处理、验收移交等五个阶段实施全过程管理，对项目的安全、标准、质量、进度等进行有效控制。

3.2 数字化工作文档管理

应明确纸质档案数字化过程中形成的工作文档的整理、归档、移交等管理要求，并与全宗卷共同保管。应归档的数字化工作文档包括：数字化工作方案、档案原件数字化审批书、档案出入库交接单、数字化加工流程单、质检任务安排单、质检情况登记表、质检验收单、项目鉴定验收报告、数字化成果移交清单等凭证性文件；采取委托加工方式的数字化项目，项目招投标文件、中标通知书、项目合同等，

3.3 记录数字化元数据

为确保纸质档案数字副本的真实、完整、可靠与可用，应在数字化过程中记录有关数字化项目、数字副本生成背景及其技术结构的元数据，主要包括数字化授权信息、数字化软硬件信息、数字化日期与时间、格式、扫描分辨率、色彩空间、压缩方案等。

4 纸质档案数字化基本环节管理

纸质档案数字化全程共有 10 个基本环节，包括审批数字化对象、制定数字化工作方案、档案出库、整理分件，建元数据库、档案扫描、图像处理、数据质检、档案还原入库、数字化成果验收移交等。

4.1 数字化对象的审批

对纸质档案数字化对象实行审批制。档案原件法定持有保管单位的

法人或授权代表要签署《江西省纸质档案原件数字化审批书》(附录 A 的表 A.1)。

4.2 制定数字化工作方案

应对数字化对象基本情况进行调查摸底,以此为基础并根据本规范提出数字化项目管理与技术要求,编制数字化工作方案。数字化工作方案应包括安全保密措施、项目管理要求、技术标准、工作文档管理要求、数字化成果移交要求等内容。采取委托加工方式的数字化项目,数字化工作方案应明确数字化成果的计算标准。

4.3 档案出库

严格按照档案库房管理规定为数字化对象办理借卷手续,案卷逐卷(以件为单位的逐件)清点后出库,重点档案逐页清点后出库。

4.4 整理分件

整理分件结果与文件级目录数据、数字副本质量及其命名、存储文件夹命名等各个环节密切关联,应严格把关。从整理分件到案卷还原的各个环节,要以卷或批次为单位建立《江西省纸质档案数字化加工流程单》(附录 A 的表 A.2),记录每卷或每批档案数字化过程。

4.4.1 整理

档案整理主要包括拆除扫描对象装订线及加固件、平整页面、编制页码、修复老化和破损程度高的页面、登记特殊页面等内容。平整页面时不得破坏档案原件固有形态,不得强行撕破档案页面。没有编页码的应重新编号,并统一采用油墨打号机打号。已经编有连续并清晰可见页

码的，不再重编页码。案卷内存在页码缺漏现象时，应登记跳号情况，不再重编页码。一份文件在卷内排列位置不连续，应调整件内页面排列顺序，并就受影响部分重新打号。卷内存有钱币、邮票、照片、地图、证书等特殊材料的页面时应予以登记。

4.4.2 分件

档案分件的主要内容是将案卷拆分为自然件。关于同一事由的往来文件应分为不同的自然件，如请示与批复应分为两个自然件。分件时应保持文件的齐全、完整，文件处理单或领导（职能部门）签署意见单、正文及领导签发的定稿、附件、浮贴等不缺漏。能反映重要文件起草、修订过程的重要稿本作为扫描件单独扫描。一般性草稿、重复件及案卷封面、卷内目录、封底等原则上不扫描，用非扫描件标识区分，不脱离卷体。应在对应的文件级目录数据“附注”字段中对非扫描件加以描述，如“本件有3份草稿共计18页未作数字化”。完成分件后，卷内文件按秩序排列，按照《江西省文书类电子档案著录与数据格式规范（试行）》要求编制文件级档号。

4.5 建立文件级元数据库

档案目录数据与本规范规定的数字化元数据的集合称为元数据库。数字化元数据由数字化系统自动捕获形成。在准确分件的基础上按照《档案著录规则》（DA/T 18-1999）、《江西省文书类电子档案著录与数据格式规范（试行）》等进行著录与目录数据录入，建立案卷级、文件级目录数据库。

数字化之前已经建立了文件级目录数据库的，应在数字化过程中进

行核对、确保目录数据准确无误并与案卷、数字副本一一对应。

4.6 扫描

原则上应使用平板扫描仪进行数字化转换。可以自然展开的筒子页应作为一个页面扫描，但件内应保持一致。扫描时应摆正档案原件页面，减少图像纠偏造成的质量下降。采用分幅拼接方式对大幅面档案进行数字化时，应在对应的目录数据中记录图像拼接行为。实物档案应采用数码相机进行数字化。

扫描仪盖板内侧为黑色时，应使用与档案原件纸张色系相同、具有一定色差的纸张覆盖扫描仪盖板内侧。

4.7 图像处理

应遵循精确复制的原则实施图像处理。按照《纸质档案数字化技术规范》(DA/T 31-2005)的要求进行图像处理，保证数字图像清晰、端正。数字副本应能准确反映档案原件的原始面貌，不得出现原件没有的颜色、变形、马赛克、明显皱纹等影像。不得去除档案页面原有的纸张褪变斑点、水渍、污点、装订孔等。应在距页边最外延 2 至 3 毫米处裁剪数字图像，不得裁掉纸质档案原有页边，包括破损、歪斜页边。

4.8 数据质检

数据质量检验包括数字图像质检与目录数据质检两个部分。应按批次数据随机抽检方式进行质检，质检次数不少于两次。一次质检抽检率不少于 20%，二次质检抽检率不少于 10%，批次抽检合格率应为 100%，否则整批数据打回重新制作或不予验收。质检工作应按任务安排、质检、质检验收三个步骤进行，并据实填写《江西省纸质档案数字化数据质检

任务安排单》(附录 A 的表 A3)、《江西省纸质档案数字化数据质检情况登记表》(附录 A 的表 A4)、《江西省纸质档案数字化数据质检验收登记表》(附录 A 的表 A5), 作为数字化成果验收移交依据。

4.9 档案还原入库

完成数据质检验收后, 应将相应的案卷按原保管形态装卷还原。对已还原案卷实体进行严格的装卷质检, 依据档案原件出库登记表、《江西省纸质档案数字化加工流程单》(附录 A 的表 A2) 等检查是否有缺卷、卷内缺页、页面脱落等情况, 装卷质检合格后入库, 并履行还卷入库手续。

4.10 数字化成果验收移交

4.10.1 数字化成果内容与要求

一个数字化项目形成的数字化成果至少包括三部分内容。

(1) 符合标准化要求的纸质档案数字副本及其元数据库。纸质档案数字副本至少包括一套源版数字副本, 一套以源版数字副本为基础转换的 PDF 或双层 PDF 格式数字副本、一套 TXT 格式数字副本。源版数字副本画幅数应按实际图像数量计算, 统计结果应与扫描页面数量保持一致。元数据库应以 DBF 等通用格式移交, 验收移交前应确认数字副本与元数据在数字档案管理系统中能够一一对应并挂接成功。

(2) 数字化工作文档。包括本规范“3.2”所述数字化工作文档, 应装订成册, 封面、封底齐全, 标识规范。

(3) 存储移交介质。包括两套移动硬盘和两套光盘。数字副本存储移交介质应按《电子文件归档光盘技术要求和应用规范》(DA/T38-2008)

和《江西省电子档案与数据离线备份介质管理规范(试行)》的要求制作。

《江西省纸质档案原件数字化审批书》(附录 A 的表 A1)、数字化工作方案、数字化项目招投标文件、中标通知书、数字化项目合同书等凭证性文件的数字副本应一同存储于授权文件夹中。

4.10.2 验收移交

应依据数字化工作方案、数字化项目合同书、质检验收单等对数字化项目作整体鉴定、验收,填写《江西省纸质档案数字化项目验收移交清单》(附录 A 的表 A6),形成鉴定验收报告,移交数字化成果。

5 纸质档案数字化技术要求

5.1 版源数字图像技术参数

应采用彩色模式实施数字化扫描。原则上应采用不低于 200dpi 的分辨率、彩色 JPEG 格式或彩色 TIFF 格式、不高于 20%的压缩率等技术参数制作源版数字副本。珍贵档案源版数字副本的制作,应采用彩色无损压缩 TIFF 格式,扫描分辨率不低于 300dpi。

5.2 源版数字副本头文件技术要求

为保障源版数字副本的真实、完整、可靠与可用,数字化系统应按照《可交换图像文件格式》(Exif)标准要求,自动在数字图像头文件中写入 4 个元数据值,并确保写入信息不丢失。4 个元数据写入要求如下表所示:

序号	英文标签	中文名称	格式规范
1	Make	设备制造商	按扫描仪制造商规范名称写入
2	Model	设备型号	按扫描仪型号规范名称写入
3	DateTimeDigitized	数字化时间	按 GB/T 7408-2005 数据元和交换格式 信息交换 日期和时间表示法 规定的格式著

			录: YYYY-MM-DDThh: mm: ss
4	ColorSpace	色彩空间	按色彩空间的规范名称写入

5.3 数字化元数据捕获

数字化系统应以件为单位自动捕获数字化元数据，将元数据与文件级目录数据合并形成文件级元数据库。数字化系统不具备自动捕获功能的，应采取手工著录与系统批量赋值相结合的方式为数字化元数据赋值。元数据项及著录要求见下表。

序号	元数据名称	类型	著录细则
1	数字化时间	日期型	著录源版数字副本第一页及最后一页的数字化时间。 格式为： YYYY-MM-DDThh: mm: ss / YYYY-MM-DDThh: mm: ss
2	数字化对象描述	字符型	对需要特别说明的案卷物理特征等所作描述。
3	数字化授权描述	字符型	著录《江西省纸质档案原件数字化审批书》的名称、编号及主要内容，以及获得授权实施数字化的责任方名称等信息
4	格式名称	字符型	著录源版数字副本格式名称
5	格式版本	字符型	著录源版数字副本格式版本号
6	色彩空间	字符型	著录源版数字副本色彩的彩空间
7	压缩方案	字符型	著录制作源版数字副本采用的压缩算法
8	压缩率	字符型	著录制作源版数字副本的压缩比率
9	水平分辨率	字符型	著录源版数字副本的水平分辨率。
10	垂直分辨率	字符型	著录源版数字副本的垂直分辨率。
11	设备类型	字符型	著录制作源版数字副本的设备类型，包括扫描仪、数码相机，默认值为“扫描仪”。
12	设备制造商	字符型	著录制作源版数字副本设备的制造商名称
13	设备型号	字符型	著录制作源版数字副本设备的型号
14	设备系列号	字符型	著录制作源版数字副本设备的系列号
15	设备感光器	字符型	著录扫描仪感光部件类型名称
16	数字化软件名称	字符型	著录数字化软件名称
17	数字化软件版本	字符型	著录数字化软件版本号
18	数字化软件生产商	字符型	著录数字化软件开发商名称

5.4 数字副本命名规则

5.4.1 源版数字副本命名规则

本规范规定了两种源版数字副本命名规则。全部数字化对象只需扫

描一个稿本时，选用命名规则 1；部分数字化对象需扫描多个稿本时，则全部源版数字副本采用命名规则 2。

(1) 命名规则 1。数字图像名称由 2 部分构成：文件级档号. 件内页面流水号。

示例：一份档号为“J015-1-00016-0001”的文件有 3 页，经扫描生成的 3 个数字图像的名称分别为：

J015-1-00016-0001. 001. JPG

J015-1-00016-0001. 002. JPG

J015-1-00016-0001. 003. JPG

(2) 命名规则 2。数字图像名称由 3 部分构成：文件级档号. 稿本代码. 稿本内页面流水号。稿本代码为：01——正本，02——定稿，03——草稿。

示例：一份档号为“X038-1967Y-026-001”的文件正本为 2 页、定稿为 2 页、草稿为 4 页，经扫描生成的 8 个数字图像名称分别为：

X038-1967Y-026-001. 01. 001. JPG

X038-1967Y-026-001. 01. 002. JPG

X038-1967Y-026-001. 02. 001. JPG

X038-1967Y-026-001. 02. 002. JPG

X038-1967Y-026-001. 03. 001. JPG

X038-1967Y-026-001. 03. 002. JPG

X038-1967Y-026-001. 03. 003. JPG

X038-1967Y-026-001. 03. 004. JPG

5.4.2 衍生数字副本命名规则

(1) PDF 或双层 PDF 格式数字副本命名规则。与源版数字副本命名规则相对应，PDF 或双层 PDF 格式数字副本有两种命名规则，如以下示例。

示例 1: X038-1967Y-026-001. PDF (规则 1)

示例 2: X038-1967Y-026-001. 01. PDF (规则 2)

X038-1967Y-026-001. 02. PDF (规则 2)

X038-1967Y-026-001. 03. PDF (规则 2)

(2) TXT 格式数字副本命名规则。TXT 格式数字副本采用文件级档号命名。

示例: X038-1967Y-026-001. TXT

5.5 数字化成果存储规范

应按照《江西省数字档案与数据离线备份介质管理规范(试行)》的要求，在移交存储介质中分类有序地存储数字化成果。按档号构成项分级建立文件夹，并可根据存储的实际需要增加下级文件夹，将数字副本相对集中地存储于最低一级文件夹内。

(1) 数字副本存储基本结构

a) 源版数字副本存储结构。源版数字副本存储基本结构为：:\数字档案或数据\全宗号\目录号\案卷号\文件级档号。

纸质档案正本、定稿、草稿经数字化形成的源版数字副本的存储，应在基本存储结构下建立相应文件夹并分别存储，其结构为：:\数字档案或数据\全宗号\目录号\案卷号\文件级档号\稿本代码。

示例 1: : \数字档案或数据\X038\1967Y\026\ X038-1967Y-026-001\01

源版数字副本同时以 TIFF、JPEG 等多种格式形成时，应参照上述方式逐级建立文件夹并存储，其结构为: : \数字档案或数据\全宗号\目录号\案卷号\文件级档号\稿本代码\格式名称，或: \数字档案或数据\全宗号\目录号\案卷号\文件级档号\格式名称。

示例 2: : \数字档案或数据\X038\1967Y\026\ X038-1967Y-026-001\01\TIFF
: \数字档案或数据\X038\1967Y\026\ X038-1967Y-026-001\01\JPEG

示例 3: : \数字档案或数据\X038\1967Y\026\ X038-1967Y-026-001\TIFF
: \数字档案或数据\X038\1967Y\026\ X038-1967Y-026-001\JPEG

b) PDF 或 TXT 格式数字副本存储结构为:

: \ 数字档案或数据\全宗号\目录号\案卷号\

(2) 元数据库存储结构为:

: \著录文件\

(3) 数字化工作文档数字副本存储结构为:

: \授权文件\数字化授权文件\

附录 A
(规范性附录)
数字化管理登记表

A.1 江西省纸质档案原件数字化审批书

审批书编号:

经办人		时间	
数字化对象			
数字化责任部门意见	年 月 日		
档案保管部门意见	年 月 日		
法人意见	<p>现批准对 等 全宗共计 卷（件）纸质档案原件进行数字化，形成的数字副本作为档案原件的异质备份件，并代替档案原件提供利用。</p> <p>单 位（签 章）:</p> <p>年 月 日</p>		

A.2 江西省纸质档案数字化加工流程单

全宗号_____ 目录号_____ 案卷号_____

流程进度标识: ☐整理分件 ☐条目录入 ☐条目质检 ☐原件扫描 ☐图像处理 ☐图像质检 ☐装订还原

整理分件				条目录入				条目质检				原件扫描				图像处理				图像质检											
作业人员		完成时间		作业人员		完成时间		作业人员		完成时间		作业人员		完成时间		作业人员		完成时间		作业人员		完成时间									
装订还原				件数		实体编页数		实际总页数		(件号)件起始页号 标注																					
作业人员		完成时间																													
1		2		3		4		5		6		7		8		9		10		11		12		13		14		15		16	
17		18		19		20		21		22		23		24		25		26		27		28		29		30		31		32	
33		34		35		36		37		38		39		40		41		42		43		44		45		46		47		48	
49		50		51		52		53		54		55		56		57		58		59		60		61		62		63		64	
65		66		67		68		69		70		71		72		73		74		75		76		77		78		79		80	
81		82		83		84		85		86		87		88		89		90		91		92		93		94		95		96	
97		98		99		100		101		102		103		104		105		106		107		108		109		110		后有附加页			
实体漏编页								实体重编页								实体缺(跳)号数															
实体送裱页								拟题件页号																							
非扫描页																															
备注																															

A.3 江西省纸质档案数字化数据质检任务安排单

批次:

质 检 人		开始时间	
质 检 对 象	图像数据计 幅	目录数据计 条	
合格率			
完成时间		签 名	
备注			

A.4 江西省纸质档案数字化数据质检登记表

批次:

质检人:

质检时间: 年 月 日

[illegible]

A.5 江西省纸质档案数字化数据质检验收登记表

批次：质检验收人：质检验收时间：年 月 日

全宗号	总卷数	图像数据		目录数据			查验落实情况	验收意见
		图像数	抽检情况	条目数	拟题数	抽检情况		
合计			---			---	---	---

A.6 纸质档案数字化项目验收移交清单

全宗号	总卷数 (卷)	图像数据 (页)	目录数据 (条)	重拟题名 (条)	数字化工作文档 (件、册)	存储移交载体

移交方：

接收方：

代表人（签章）：

代表人（签章）：

日期： 年 月 日

日期： 年 月 日

