附件2

纸质实体材料电子化扫描整理

工作流程及相关技术要求

纸质实体材料电子化扫描整理的基本流程主要包括纸质实体材料整理、纸质材料扫描、改进文档质量、文档纠错处理、数据文档重命名及存储、目录建库、数据挂接、数据验收、数据备份、装订、材料归还及登记等，并做好相关扫描注意事项。具体工作流程及相关技术要求如下：

**一、材料整理**

**（一）材料接收及登记**

由业务科室经办人移交传递的纸质材料进行逐一清点签收，纸质资料签收需做到与交接清单一一勾对，要求做到及时、准确无误。交接双方要共同核实页码及材料质量是否存在问题。在此阶段，材料提交方应明确提出电子化扫描规格和质量要求，并明确提出扫描后形成的数据文档具体命名（或编制可供相应信息管理系统检索的数据库“档号”）、数据文档存储路径的具体要求。技术服务单位可结合项目实施的情况修改完善纸质材料接收登记表。

**（二）编制目录准备**

参照《档案著录规则》（DA/T18）等的要求，配合土储中心业务科室及具体经办人进行著录，编制规范，建立并逐步完善“XX地块全链条业务办理资料台账”纸质实体材料目录及相对应的电子文档目录。

**（三）拆除装订物**

如不去除装订物会影响扫描效果的纸质材料，可将纸质材料原件的装订物拆除，按照原件排好顺序，不得出现任何漏缺页及顺序差错，更不能对纸质材料原件有任何损坏。

**（四）页面修整**

1.对破损严重、无法直接进行扫描的纸质材料，应先进行技术修复、裱糊。

2.折皱不平影响扫描质量的原件应先进行相应处理（压平或烫平等）后再进行扫描。

**（五）纸质件手动整理过程登记**

对于确需扫描，且要拆除装订物并花费较长时间处理的纸质件，由扫描人员制作并填写纸质材料手动整理过程登记表，详细记录纸质材料整理后每份材料的起始页号、页数、特殊情况及其处理等信息。

**二、纸质材料扫描**

**（一）扫描方式**

项目技术服务单位应提供扫描尺寸规格不低于A3大小且带平板的彩色高速扫描仪。

1.根据纸质材料幅面（A4、A3、AO等）的大小选择相应规格的扫描仪，尺寸规格超出A3大小的纸质件，经申请可借助机关办公楼内其它科室的扫描仪进行扫描。

2.普通A4纸质文件，采用高速扫描仪的自动进纸方式扫描。

3.纸质过薄、透明的（如信纸、便签纸）采用高速扫描仪的平板扫描。

4.纸质过厚、照片等纸质材料采用高速扫描仪的平板扫描。

5.纸质材料页面贴有附属大小页面和纸张，将大小页面单独在平板中扫描。附属大页面和纸张尺寸规格超出A3大小的纸质件，经申请可借助机关办公楼内其它科室的扫描仪进行扫描。

**（二）扫描模式**

纸质材料扫描一般采用黑白、灰度和彩色三种模式。

1.页面为黑白两色，字迹清晰的、不带图片的纸质材料材料，采用黑白方式。

2.页面为黑白两色，清晰度较差或者带有图片的纸质材料材料，以及页面多为彩色文字的纸质材料，采用彩色或灰度模式（征求科室意见，视情况而定）。

3.页面中有红头、印章或插有黑白照片、彩色照片、彩色插图的纸质材料，采用彩色模式扫描。

**（三）分辨率选择**

以下模式分辨率的设定以A4纸张为标准，其他规格纸张酌情处理。

1.采用黑白、256级灰度模式扫描的文件，其分辨率选择为不小于200DPI。

2.采用24位为彩色模式扫描的文件，其分辨率选择为不小于100DPI。

3.扫描的线数、阀值、亮度、灰度、对比度等值可根据所扫描文件材料的清晰度进行适当的调整。

4.需要时可根据原件的清晰度适当调整扫描分辨率。如原件质量较差且尺寸较小，可适当提高分辨率；反之也可相应减少分辨率，增减的多少以扫描后图像按原尺寸显示后是否清晰为准。

5.粘贴折页与表格，对于粘贴折页，可用大幅面扫描仪扫描，或先部分扫描后拼接；对部分字体很小、字迹密集的情况，可适当提高扫描分辨率，选择灰度扫描或彩色扫描，采用局部深化技术解决；对字迹与表格颜色深度不同的，采用局部淡化技术解决。

**（四）扫描登记**

填写扫描登记表，登记扫描的页数。核对每份文件的实际扫描页数与纸质材料整理时填写的文件页数是否一致，不一致时应注明具体原因和处理方法。

**三、改进扫描文档质量**

对扫描文档（图像）进行处理，以获得最好的图像质量，保证图像完整、端正、无扭区、版面无暗影、无干扰信息，主要完成图像处理包括：去黑边、去污点、纠偏等，处理完的图像保存格式为PDF。

**一是纠偏。**对扫描过程中出现的偏斜图像进行整体纠正，包括自校和手校，保证数字图像的偏斜角度小于1度（图像偏斜不超过页面内半个文字）。

**二是旋转。**按方字方向将图片旋转至正确方向，没有文字的图片，判断其方向后进行左旋、右旋、翻转、旋转等工具（前提是需与原件一致，征求具体科室意见）。

**三是去污。**对图像页面中出现的影响图像质量的杂质，去除数字图像中影响可懂度的杂质，如去除黑边、多余边、污点。

**四是裁边。**采用彩色模式扫描的图像应进行裁边处理，去除多余的白边，以有效缩小数据文档的容量，节省存储空间。

**五是删除空白页。**将扫描后页面之间的空白页进行删除。

**六是断字修补。**对部分文字不清楚可进行修补（前提是需与原件一致，征求具体科室意见）。

**七是反白字修正。**将部分反白文字可进行描述（前提是需与原件一致，征求具体科室意见）。

**八是分割。**将A3幅面的文件分割为两份A4幅面的文件。

**九是拼接。**对大幅面纸质材料进行分区扫描形成的多幅图像，应进行拼接处理，合并为一个完整的图像，以保证纸质材料电子化扫描图像的整体性。

**四、扫描文档纠错处理**

主要对扫描后形成的“图像”格式的文档进行检查，不合要求的返回上一环节重新处理，如以下问题。

**一是扫描操作不当问题。**由此造成扫描的图像文档不完整或无法清晰识别时，应重新扫描。

**二是图像偏斜度、清晰度、失真度问题。**发现不符合图像质量要求时，应重新进行图像的处理。

**三是图像处理是否得当。**问题严重的应重扫，处理不当的应重新处理。

**四是页码顺序、漏扫、多扫、倒置问题。**检查页码顺序是否与原文件一致，有无漏扫、多扫（重复）、纸张倒置等情况。

**五是顺序问题。**如果页码不连续，与原文不一致，需对页面进行排序。

**六是漏扫问题。**将漏扫的页面重新扫描，并插入到正确的位置。

**七是多扫问题。**将多扫页面删除。

**八是倒置问题。**不符合要求的需进行调整，保持与原文一致。

**五、数据文档重命名及存储**

技术服务单位和土储中心统一制定合理的命名规则，依照纸质材料名称、发文字号等属性，将所有扫描后的数据文档重新命名（或编制“档号”），将拆装多页的同一份纸质材料，合并为一个PDF文档，使实体纸质材料文档与生成的数据文档建立一一对应的关系，并将数据文档存入内部公共盘指定的文件夹中。禁止用Word文档保存扫描文档。

**六、目录建库**

参照《档案著录规则》（DA/T18）等的要求，配合土储中心业务科室经办人进行著录，规范建立并逐步完善“XX地块全链条业务办理资料台账”纸质实体材料目录及相对应的电子文档目录，并形成目录清单表。目录建库应选择通用的数据格式。所选定的数据格式应能直接或间接通过XML文档进行数据交换。可采用人工校对或软件自动校对的方式，对目录数据库的建库质量进行检查。核对著录项目是否完整、著录内容是否规范、准确，发现不合格的数据应要求进行修改或重录。

**七、数据挂接**

**（一）挂接前的数据关联检查**

以纸质材料目录数据库为依据，将每一份纸质材料文件扫描所得的一个或多个文档，合并为一个PDF文档。将数据文档存储到相应文件夹时，要认真核查每一份数据文档的名称与纸质材料目录数据库中该份文件的“档号”是否相同，数据文档的页数与纸质材料目录数据库中该份文件的页数是否一致，数据文档的总数与纸质材料目录数据库中文件的总数是否相同等。通过每一份数据文档的文件名与纸质材料目录数据库中该份文件的“档号”的一致性和唯一性，建立起一一对应的关联关系，为实现纸质材料目录数据库与数据文档的批量挂接提供条件。

**（二）汇总挂接**

纸质材料电子化扫描过程中编制的目录数据库与数据文档目录数据库，通过质检环节确认为“合格”后，通过网络及时加载到数据服务器端汇总。通过编制程序或借助相应软件，可实现目录数据对相关联的指定扫描格式的数字文档自动搜索、加入对应的电子地址信息等，实现批量、快速挂接。  **八、数据验收**

纸质材料电子化扫描转换质量抽检的合格率达到95%以上（含95%）时，给予以验收“通过”。目录数据库与数据文档挂接错误，或目录数据库、数据文档之一出现不完整、不清晰、有错误等质量问题时，抽检标记为“不合格”。验收“通过”的结论，必须经业务科室具体业务经办人核准、签字后方有效。验收结束，填写纸质材料电子化扫描数据验收登记表，随之及时更新完善相应的数据资料清单。

**九、数据备份**

经验收合格的完整数据应及时进行备份。著录条目和全文数据进行内部公共磁盘数据中建立存储备份。移交前文件读取性能，抽查文档质量，主要包括备份数据能否打开、数据信息是否完整、文件数量是否准确等。不合格的需重新整理。数据备份后应在备份文件夹编写说明文档开列清单，标注文件内容、类别、存入日期等，以便查找和管理。有保密等级要求的需标明密级，并根据保密要求存储和管理。

**十、装订**

扫描工作完成后，拆除过装订物的纸质材料应按纸质材料保管的要求重新装订。恢复装订时，应注意保持纸质材料的排列顺序不变，做到安全、准确、无遗漏。

装订不能损害纸质材料原件。装订时应按原有顺序装订，案卷不掉页、左边和底边整齐，保持拆卷前的原貌，并认真做好纸质材料页码、页数的检查对校。

**十一、纸质材料归还及登记**

纸质材料扫描完成后，及时交回业务科室经办人员并做好交接登记。由科室经办人员将材料装盒，以同一份材料尽可能不分两卷的原则进行装盒（资料盒需在盒上填写名称），督促科室经办人员及时补充完善资料盒内纸质材料清单表。

**十二、扫描注意事项**

（一）扫描前必需对文档进行拆除装订物，并检查文档内是否藏有干扰物（如钉书钉、碎纸等），以免产生卡纸、损坏扫描仪，务必保持文档干净送入扫描仪中。

（二）对破损严重、无法进行扫描的原件要先进行修复、裱糊。

（三）褶皱不平影响扫描质量的原件应先压平再进行扫描。

（四）对纸质过薄、透明的（如信纸、便签纸）、纸质过厚、照片等采用高速扫描仪的平板扫描；碳写纸、热敏纸可复印后再扫描。

（五）对文件页面贴有附属大小页面、纸张时，将大小页面单独在平板中扫描。

（六）扫描完每一份纸质文件，要对照原文仔细检查扫描是否清晰、完整。在扫描的过程中必须保证文件及资料实体件的安全不受损坏、不漏扫（如果存在不能高速扫描的需要标明，后续进行平板补扫，并利用软件将数据文档合并成一个完整整体），确保每份完整的纸质文件对应一个完整的电子化数据文档，确保扫描文档页码及质量无问题。

（七）按要求使用扫描仪、清洗，并对每次使用完扫描仪进行保洁，检查是否关闭电源。

（八）当天接收的文件及资料必须在当天完成电子化扫描及重新装订并交接入盒工作，并配合土储中心完成当日工作审核。

（九）纸质材料电子化扫描成果提供网上检索利用时，应有制作单位的电子标识，并根据具体情况分别采用可下载或不可下载的数据格式。

**纸质实体材料电子化扫描整理流程图**

接收及登记，编制目录准备，拆除装订物，页面修整，手动整理过程登记

纸质材料整理

数据验收

数据备份

接收、拆订、修整、登记

纸质材料扫描

纠偏、去污、拼接、裁边等处理

改进文档质量

文档纠错处理

目录著录、形成清单表

目录建库

数据挂接

重新装订

纸质材料归还及登记

数据文档重命名及存储

交接登记

各类问题检查及处理