附件：

纸质档案数字化标准要求

一、技术标准

按照北京市地方标准《档案数字化规范》（DB11/T 765——2010）执行。

二、关键指标要求

色彩模式：彩色

分辨率：300dpi

存储格式：TIFF、JPEG、PDF

图像文件命名：与对应纸质档案的档号保持一致

三、加工设备要求

纸质档案数字化加工设备应包含但不限于计算机、扫描仪、打印机、刻录机、外存储设备等。应确保加工设备的安全和先进，能够满足数字化加工的安全要求和技术需要。

四、数字化加工流程

纸质档案数字化加工的基本环节主要有：档案整理、目录建库、档案扫描、图像处理、数据质检、数据挂接、数据验收和档案出入库等。

具体流程见“图1：纸质档案数字化加工流程图”。

图1：纸质档案数字化加工流程图

五、档案整理环节的技术要求

**（一）确定扫描页**

原则上应将确定为数字化对象的纸质档案全部扫描，不宜进行挑扫。如有特殊情况不能扫描的页面应在备考表中注明并进行登记。

扫描内容应能反映档案的完整性和连续性，对于维持档案完整性和连续性的辅助性页面也应同步扫描。

**（二）编制页号**

1．宜使用2B铅笔进行编页。

2．应对没有页号或页号不正确的档案重新编制页号；并对排列顺序不准确的档案进行重排。

3．重新编制页号时，应在页面符合阅读方向的底端中部编写页号，且不能压盖档案内容。底端中部已有页号，且页号错误时，应将原错误页号用铅笔划去以示区分。

4．使用阿拉伯数字从“1”开始依次编写页号，应保证每卷或每件页号编写的连续性。

5．应将破损页面、缺页等特殊情况在备考表中注明并进行登记。

**（三）拆除装订**

编页之后，应以对纸质档案的保护为原则确定是否拆除装订，以保证扫描的质量和效率。对需要拆除装订的档案，在拆除装订物时，应注意保护档案不受损害，并对排列顺序不准确的档案进行重排；对确实不宜拆除装订的档案可以采取其他技术手段处理。

特殊装订且拆除装订后需恢复的档案，在拆除装订物时可采用拍照等方式记录档案原貌，以便于恢复。

**（四）页面修整**

1．破损严重或其他无法直接进行扫描的纸质档案，应先由专业技术人员进行技术修复。

2．折皱不平影响扫描质量的纸质档案应先进行压平等页面修整后再扫描。

**（五）填写备考表**

在整理纸质档案过程中，如对档案实体进行了调整需要记录，应在备考表中注明。

六、目录建库环节的技术要求

**（一）建立目录数据库**

按照《北京数字档案馆目录数据库建库标准》确定著录项，建立目录数据库。

**（二）目录数据质量要求**

可采用软件自动校对与人工校对相结合的方式，对目录数据库的建库质量进行检查，包括著录项目的完整性、著录内容的规范性和准确性等。发现不合格的数据应及时进行修改或重新著录。

七、档案扫描环节的技术要求

**（一）基本要求**

档案扫描应根据纸质档案原件的实际情况、数字化目的、数字化规模、计算机网络和存储条件等选择相应的扫描设备，和进行相关参数的设置和调整。参数的设置和调整应保证扫描后数字图像清晰、完整、不失真，图像效果最接近档案原件。

**（二）扫描设备**

1．扫描设备的选择应特别注意对档案实体的保护，尽量采用对档案实体破坏性小的扫描设备进行扫描。纸张状况良好的档案可以采用连续扫描方式，纸张状况较差（过薄、过软、有破损等）的档案应采用平板扫描方式。

2．大幅面档案宜采用更大幅面扫描仪进行扫描，也可以采用小幅面扫描仪分幅扫描后进行图像拼接的方式处理。分幅扫描时，相邻图像之间应留有足够的重叠。

**（三）色彩模式**

为最大限度保留档案原件信息，便于多种方式的利用，宜全部采用彩色模式进行扫描。

**（四）分辨率**

1．扫描分辨率的选择，应保证扫描后图像清晰、完整，并综合考虑数字图像后期利用方式等因素。

2．彩色模式扫描，建议分辨率不小于300dpi。含有照片或有展陈需求的纸质档案，建议分辨率不小于600dpi。

**（五）存储格式**

1．纸质档案数字图像长期保存格式为TIFF、JPEG或PDF。

2．纸质档案数字图像在利用时，考虑到网络浏览速度、易操作性、存储空间占用等综合因素，可以将TIFF图像转换为JPEG或PDF等其他格式。

**（六）图像命名**

1．应以实体档案的档号为基础对数字图像命名。图像命名方式的选择应确保图像命名的唯一性。

2．建议将数字图像存储为单页文件，并按实体档案的档号与图像流水号的组合对图像命名。

例如：全宗号-目录号-年度-件号-页号（不立卷）

全宗号-目录号-案卷号-件号-页号（立卷）

3.数字图像确需存储为多页文件时，可采用该实体档案的档号对图像命名。

例如：全宗号-目录号-年度-件号（不立卷）

全宗号-目录号-案卷号-件号（立卷）

八、图像处理环节的技术要求

**（一）图像拼接**

对分幅扫描形成的多幅数字图像，应进行拼接处理，合并为一个完整的图像，以保证纸质档案数字图像的整体性。拼接时应确保拼接处平滑地融合，拼接后整幅图像无明显拼接痕迹。

**（二）旋转及纠偏**

对不符合阅读方向的数字图像应进行旋转还原。对出现偏斜的图像应进行纠偏处理，以达到视觉上基本不感觉偏斜为准，偏斜不宜超过3度。

**（三）去污**

扫描过程中产生的污点、污线、黑边等影响图像质量的杂质，应遵循展现档案原貌的原则，予以去除。在去污过程中不得去除档案页面原有的内容信息。

九、数据质检环节的技术要求

**（一）目录数据质量检查**

目录数据库著录项目不完整，著录内容不规范，目录数量不对应的，应修改或重新著录。

**（二）图像质量检查**

1．数字图像不完整、无法清晰识别或图像失真度较大时，应重新扫描。

2．对数字图像拼接、旋转及纠偏、裁边、去污等处理情况进行检查，发现不符合图像质量要求时，应重新进行图像处理。

3．数字图像的排列顺序与档案原件不一致时，应及时进行调整。

4．对于漏扫、重扫、多扫等情况，应及时改正。

十、数据挂接环节的技术要求

**（一）数据关联**

应借助相关软件对目录数据库中的目录数据与其对应

的纸质档案数字图像进行挂接，以实现目录数据与数字图像的关联。

**（二）关联质量检查**

逐条对挂接结果进行检查，包括目录数据与纸质档案数字图像对应的准确性、已挂接数字图像与实际扫描数量的一致性、数字图像是否能正常打开等，发现错误及时纠正。

十一、数据验收环节的技术要求

**（一）验收内容及方式**

以抽检方式检查目录数据库、图像文件、数据挂接的质量以及数字化工作记录是否完整等，抽检比率应不低于数字图像总页数的10%。可采用计算机自动检验与人工检验相结合的方式对纸质档案数字化成果进行验收。

**（二）验收指标**

当目录数据录入有错误，或图像文件不完整、不清晰，或目录数据与图像文件挂接错误，以上问题出现任意一种确定为“不合格”，统计“不合格”文件数。

抽检合格的文件数= 抽检文件总数-抽检不合格的文件数

抽检合格率= 抽检合格的文件数/ 抽检文件总数×100%抽检合格率应不低于99.7%。

**（三）验收结论**

1．抽检合格率达到99.7%（含99.7%）以上，判定数据验收“通过”。

2．抽检合格率低于99.7%时，判定数据验收“不通过”，该批次数据全部退回返工或修改后，重新进行验收。

**（四）移交**

验收“通过”的数据应及时移交，并履行交接手续。

十二、档案出入库环节的技术要求

**（一）档案出库**

1．档案管理人员应严格按照档案库房相关管理规定对待数字化档案办理出库手续，与档案数字化加工人员共同清点无误后，对档案进行交接出库。

2．纸质档案数字化过程中，应设置距离数字化加工场地较近的保管库用以临时存放纸质档案，并对纸质档案的领取与归还进行严格管理，认真做好检查、清点、登记等工作，确保档案实体的安全。

**（二）档案装订**

纸质档案数字化工作完成后，拆除装订的档案应重新装订，装订时应注意基本恢复档案原貌，做到安全、准确、无遗漏。

**（三）档案归还入库**

档案管理人员应按照档案库房相关管理规定对档案实体进行检查和清点，并履行档案入库手续。