# 地信学院教学档案扫描加工项目技术参数和要求

本项目采购的内容包括档案资料分拆卷、页码校核、档案整理、全文扫描、图像处理、图像半成品质检及存储、档案分件、系统数据入库、机读目录上传、全文上传、全文数据质检并提供质检报告等服务。

一、项目描述

1.1纸质档案数字化需求

1.1.1项目概况

本次档案数字化加工数量35万页,为确保后期档案查询利用效果，在本次项目中，已具备电子目录的，需要对档案条目信息进行校对修正，无电子目录的需要补录，另外需要对本期项目涉及到的库存档案资料进行电子化扫描、图像处理、数据入库、挂接、机读目录上传、全文上传、全文数据质检并提供质检报告等服务。以下为具体要求：

1.1.2标准依据

DA/T 31-2005 《纸质档案数字化技术规范》；

1.1.3技术要求

1.1.3.1.属性数据录入

（1）此次招标工作中属性数据录入方面要求对已有数据的案卷进行校对、补录；对未录的案卷要按照要求录入，符合有关规范要求。

（2）其他业务档案需按照类型以不同字段进行录入，主要包括学年，学期，课程名称、任课教师、职称等类型等。

1.1.3.

2.影像扫描

（1）工作量

本次招标的工程量约为35万页左右，要求提交成品为真彩24位JPG及双层PDF两种格式。实际工作量有可能增加或者减少，增加或减少的工程量按照相关单价结算。

（2）数据格式与存储要求

a)满足国家档案行业标准《纸质档案数字化技术规范》（DA/T31-2005）。

b)投标方需提交数据存储用品两套全文成品数据，数据必须达到以下所述质量

控制技术指标：

●存储级：采用300dpi分辨率扫描，存储格式为真彩24位JPG。中标方提供光

●利用级：采用300dpi分辨率扫描，存储格式为双层PDF，硬盘保存。

（手写体及非标准字库等无法识别的可存储为PDF-1.4格式文件）。

1.1.3.3.扫描过程要求

A. 扫描时应根据纸质档案的材质选择相应的扫描设备，纸质较好的档案可采用滚筒等高速扫描方式、纸质较薄、较脆的必须采用平板扫描方式。保证档案不受损伤，禁止使用拍照方式。

B. 扫描时应认真登记扫描的页码和页数，并核对实际扫描页码、页数与前处理时填写的页码、页数是否一致，不一致时应注明具体原因和处理方法。

C. 对案卷中漏编的页面进行合页处理，确保扫描好的影像页数与原案卷页数一致。

D. 同页有两个以上文件，需分别扫描处理。（主要针对文件，试卷不用）

E. 扫描全部采用彩色扫描，以扫描后的图像清晰、完整，档案内容信息与档案原件一致为准，最大限度保持档案原貌。

F. 扫描参数如下：

彩色扫描分辨率采用300dpi或以上。

扫描时保证放纸端正、不压边、不漏扫、错扫、严格确保图像质量。扫描中漏扫、压边的要即时补全、修正。

1.1.3.4.图像处理要求

A. 图像处理后保证图像信息与原档案内容完全一致，不删除页面任何有用信息，

包括正文内容、页眉、页脚、手写注释和印鉴等。扫描的页面内容基本居中显示，不出现明显偏左或偏右现象。不准出现页面内容残缺或将其它页面信息扫入本页的现象。

B. 每页影像左右端正度控制在0.5度以内，以达到视觉上基本不感觉偏斜为准。对方向不正确的图像进行旋转还原，以符合阅读习惯。不出现图像的一部分出现倾斜或扭曲而影响阅读的现象。厚的案卷装订线较近边角的档案内容会产生扭曲现象，需保证正文能看清楚。

C. 去污：在保证文件内容完整的前提下，影像页无扫描过程中带来的污斑，无黑边，外观达到清晰、平直、干净。对图像页面中出现的影响图像质量的杂质如黑点、黑线、黑框、黑边等进行去污处理。处理过程中应遵循在不影响可读性和可理解性的前提下展现档案原貌的原则，由于档案本身因陈旧、破损或不洁等造成页面上有污点的，如果在页面视图下无法看清，或该污点嵌在文本行中无法清除，可不予清除。

D. 保证扫描图像字迹清晰、颜色恰当，不宜过浅或过深，不出现字迹笔画残缺或字迹笔画叠合而影响阅读的情况，档案原件存在斑迹变质、颜色过浅、过深或深浅不一致，保证档案原件能辨认的扫描图像也必须可以辨认。纸张太薄或字体颜色过浓，扫描时倒映反面文字或文字上的污点无法去污时，应保证正文文字可以看清楚。

1.1.3.5.数字化流程要求

（1）在数字化加工过程中，需全程在我们监控下完成；数字化流程至少应包括：出库、查卷、整理、条目著录、扫描、图像处理、质检、装订、数据备份、提交、还卷等。

（2）形成相应的员工操作记录，控制员工操作中的每一步，系统自动生成数据记录操作过程，随时检查任何工序的质量及完成情况，确保一次性达到成品质量要求。

1.1.3.6.数据安全要求

需要提供对关键数据影像的安全技术措施,采用技术手段对设备端口进行控制，对加工数据进行加密，实现档案数据拿不走、非法拷贝打不开，待移交正式成品数据时再进行解密，以实现事先防范。配合场地安全控制、设备安全控制、管理制度安全控制和人员安全控制手段，确保数字化加工过程中数据的安全。

1.1.3.7.图像数据命名方式

●归档档案实体图像命名方式

图像数据分三层文件夹存放，第一层文件夹以材料类型，第二层文件夹以年度学期，文件夹内存放该目录的所有卷，第三层文件夹以案卷名称命名，文件夹内存放图像文件，图像文件以页码命名。

1.1.3.8.时间项要求

目录数据中的时间项一律用11或8字符 位字符型（YYYY- YYYY -DH或）表示，年、月、日、学期无法确定补零充填。

例1:1992年8月2日，著录为：19920802

例2: 2021至2022第2学期，著录为：2021-2022-2

1.1.3.9.数据备份与提交

在数字化加工过程中，加工方应做好半成品数据（特别是原始扫描影像文件）和成品数据的存储备份工作。加工方自备存储设备，等项目验收数据提交完成后，中标单位收到数据提交明细验收清单后，方可将数据从备份载体中删除。

●数据备份载体要求

数据装载;移动硬盘2套。

●移交数据清单

加工方应提交硬盘数据和光盘数据移交清单，说明每张硬盘数据以及移交数据总量，包括目录数、件数、页数、数据容量和移交数据的其它说明。

1.1.3.10.数据验收

1.1.3.10.1数据检查验收

●目录数据验收质量要求

采用人工校对，核对著录项目是否完整、著录内容是否规范、准确，对不合格的数据进行修改。详细要求见下表：

●图像数据验收质量要求

图像数据中不能出现扫描格式错误、命名错误、歪斜、黑边、多扫、漏扫、扫描质量模糊等问题。具体要求见下表：



1.1.3.10.2验收方式

●以抽检的方式检查已完成数字化转换的所有数据，包括目录数据库、图像文件及

数据挂接的总体质量。

●一个全宗的档案，数据验收时抽检的比率不得低于30%。

●目录数据库与图像文件挂接错误，或目录数据库、图像文件之一出现不完整、不

清晰、有错误等质量问题时，抽检标记为“不合格”。

●一个全宗的档案，数字化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”。但部分重要档案的一次抽检合格率应达到98%以上（含98%）.

●合格率=抽检合格的文件数/抽检文件总数×100%

二、项目完成时间：30 (工作日)

三、项目完成数量：35万页左右