

纸质档案数字化验收标准体系研究

项目研究报告

云南省档案局（馆）

科研项目课题组

2012. 11

目 录

引言	1
一、国内外同类项目研究现状.....	2
1.1 国外研究现状	2
1.2 国内研究现状	3
二、项目工作基本情况.....	4
2.1 项目名称	4
2.2 项目承担单位	4
2.3 项目课题组成员	4
2.4 项目研究的目的意义及立项背景	4
2.5 项目研究的主要内容	5
2.5.1 数字化加工安全保密验收	5
2.5.2 档案图像扫描验收.....	5
2.5.3 图像处理验收.....	6
2.5.4 图像存储验收.....	6
2.5.5 目录建库验收.....	6
2.5.6 数据挂接验收.....	6
2.5.7 档案文件级目录及题名验收	6
2.5.8 档案整理验收.....	6
2.5.9 数据备份验收	6
三、项目组织实施情况.....	7
3.1 项目组织	7

3.1.1 项目研究	7
3.1.2 项目基础调研	8
3.1.3 项目完成情况	10
3.2. 项目研究人员、经费投入及使用情况	11
3.2.1 项目研究人员情况	11
3.2.2 经费投入及使用情况	11
四、项目研究成果	11
4.1 数字化加工安全保密验收	11
4.2 档案扫描验收	13
4.3 图像处理验收	17
4.4 图像存储验收	20
4.5 目录建库验收	21
4.6 数据挂接验收	21
4.7 档案文件级目录及题名验收	29
4.8 档案整理验收	37
4.9 数据备份验收	37
五、项目研究成果推广应用	43
六、项目研究报告及相关验收规定、办法、教材及表单	43
6.1 档案数字化验收规定、办法及教材	43
6.2 档案数字化验收申请、报告及表单	44

纸质档案数字化验收标准体系研究

引言

根据国家档案局制定的《中华人民共和国档案法》、《数字档案馆建设指南》、《电子文件归档与管理规范》、《纸质档案数字化技术规范》及《云南省数字化档案馆建设规划》《云南省档案馆数字化加工管理办法》等相关规定，为稳步推进我馆数字化档案馆建设步伐，规范我馆档案数字化加工和安全管理，确保档案数字化过程中档案实体和信息安全，确保档案数字化成果有效提供利用，实现档案数字化验收工作更为科学、合理、有效、规范的管理，经云南省档案局（馆）研究决定，成立数字化管理中心，采用对外招投标外包的方式，组织4家数字化加工公司开展档案数字化加工工作。

随着馆藏档案数字化工作的有序开展，从2009年7月至2012年11月，云南省档案馆已圆满实现馆藏档案62个全宗188293卷的原文数字化工作（包括文件级目录2604552条及27208424画幅的图像数据）。同时，为更好地有效保护数字化完结全宗档案原卷，我馆将扫描后顺利通过数字化验收的档案原件及时进行封存。

通过三年多的数字化工作，我们认识到：数字化扫描影像（图像）能否与档案原件信息一致，数字化成果能否有效提供利用，是我们数字化工作成败的关键。因此，验收工作成为档案数字化建设不可或缺的重

要环节，在现有档案数字化工作经验基础上，有计划地开展《纸质档案数字化验收标准体系研究》项目研究显得尤其必要。

项目组以确保档案原件与图像数据一致，并能保证信息安全和方便快捷提供利用为课题研究目标，以数字化工作实践经验为基础，在《纸质档案数字化验收标准体系研究》项目研究过程中，紧密结合数字化工作职责、工作任务、岗位特点及档案资源建设体系、档案利用服务体系、档案安全体系建设情况和档案数字化成果验收工作实际，围绕加工公司自检、数字化管理中心初验、组织专家验收三个步骤，从档案数字化加工安全保密验收等九个环节，有针对性地制定了一系列操作性较强的档案数字化成果测试验收标准，系统、认真、细致的对《纸质档案数字化验收标准体系研究》科研项目开展全方位的研究。现就《纸质档案数字化验收标准体系研究》科研项目研究成果报告如下：

一、国内外同类项目研究现状

1.1 国外研究现状

美国和英国是国外档案数字化工作起步较早的国家。1994年，美国国家档案文件局与美国研究图书馆团体合作成立了一个数字档案特别工作组，并出台了一份关于数字档案的最终报告和建议。在这份报告中提出了国家数字档案馆方案。同时，美国国家档案文件局宣布将与美国国家科学基金会共同创建一个电子文件档案馆；英国公共档案馆也和伦敦大学合作推出了联合王国数字档案馆计划。档案界学者也纷纷对虚拟档案项目建设开展了认真细致的研究工作，自1996年戴维·比尔曼在

十三届国际档案大会上发表《虚拟档案》以来，数字化档案馆问题开始成为档案界研究的热点。但是，截至今日，针对纸质档案数字化验收标准体系的项目研究，国外档案学者却很少涉及。

1.2 国内研究现状

在国家档案局一系列政策标准的引领和规范下，2000 年国家档案局在深圳召开了全国第一次数字档案馆建设研讨会，拉开了全国数字档案馆建设的序幕。北京、上海、天津、浙江、福建、江苏、重庆、深圳、青岛等省市先后开展数字档案馆建设。山东省已建成青岛、淄博、威海、烟台等 8 个市级数字档案馆和 29 个区县级数字档案馆；浙江省 2009 年全省开展数字档案馆建设达 68 个；深圳数字档案馆是公认的拓荒者。各地档案部门在充分利用数字档案资源，加快信息服务平台建设中，辽宁省档案信息网构建了四个共享平台，青岛市档案信息网提供全市 13 个综合档案馆网页及开放档案查询平台，这些研究与探索标志着我国档案数字化研究与实践的开始。目前，《纸质档案数字化技术规范》（中华人民共和国档案行业标准 DA/T 31—2005）是我国纸质档案数字化工作较为系统全面的规范指导。但国内档案界对数字化成果质量检验尚缺乏行之有效的技术手段和统一标准。

浙江省档案局积极开发研制档案数字化成果质检系统，并在浙江省推广应用。其特点是：通过计算机技术手段辅助检测，对存储在光介质或磁介质上的海量数据自动形成各类统计报表、检测报告和移交清单，大大提高数字化加工和入馆数据的准确性与可靠性，填补了国内档案数

据质量检验实用工具的空白。

我馆目前开展的《纸质档案数字化验收标准体系研究》项目研究，与浙江省档案局研制开发的档案数字化成果质检系统有所不同，侧重于纸质档案数字化验收标准体系的研究，依据一系列可操作性较强的标准将数字化成果检测指标和流程逐一细化，逐项开展档案数字化成果测试验收工作。

二、项目工作基本情况

2.1 项目名称

《纸质档案数字化验收标准体系研究》

2.2 项目承担单位

云南省档案局（馆）

2.3 项目课题组成员

项目组组长：黄凤平

项目组成员：张文芝、孙立徽、陈虹、罗备针、邵海燕、史家菊、尚丽萍

2.4 项目研究的目的意义及立项背景

为确保档案原件与图像数据一致，实现馆藏档案目录及原文图像数据的有效提供利用，我局（馆）坚持统一领导、统筹规划、突出重点、

分步实施、稳步推进的数字化原则，积极组织开展数字化档案验收课题研究。验收方式：集中云南全省档案部门数字化专家，建立专家库，以全宗为单位，从专家库随机抽调数字化专家进行成果验收。经过十八批次的数字化成果测试与验收，积累了大量的科研数据及实践操作经验，初步形成一套行之有效的数字化验收的标准规范。为规范档案数字化项目验收的管理，确保数字化加工后的数据真实、准确、安全、实用奠定了坚实的基础。

2.5 项目研究的主要内容

依据国家局和云南省局制定的档案法律法规和数字化加工标准规范，全省数字化验收专家组以全宗为单位，围绕加工公司自检、数字化管理中心初验、组织专家验收三个步骤，从档案数字化加工安全保密验收等九个环节，对纸质档案数字化条目及图像信息逐卷逐条逐页开展一系列规范的测试和验收以及审核，使档案条目和图像数据质量达到相关要求和标准。

纸质档案数字化验收标准体系研究项目主要包括以下内容：

2.5.1 档案数字化加工安全保密验收

从档案数字化加工的场所、加工流程、现场管理、存储介质管理和处理等环节检查加工公司是否按照相关的安全管理制度和规范进行数字化加工和管理。

2.5.2 档案图像扫描验收

检验图像扫描色彩模式、扫描分辨率的设定参数是否符合规范。

2.5.3 图像处理验收

对图像数据质量、纠偏、去污三个方面进行检查。

2.5.4 图像存储验收

检查档案图像存储格式是否按照 TIFF 和 PDF 格式分别存储和备份。

2.5.5 目录建库验收

对目录数据转换格式及建库质量进行检查。

2.5.6 数据挂接验收

检查图像文件与档案目录数据库中的字段对应关系是否正确。

2.5.7 档案文件级目录及题名验收

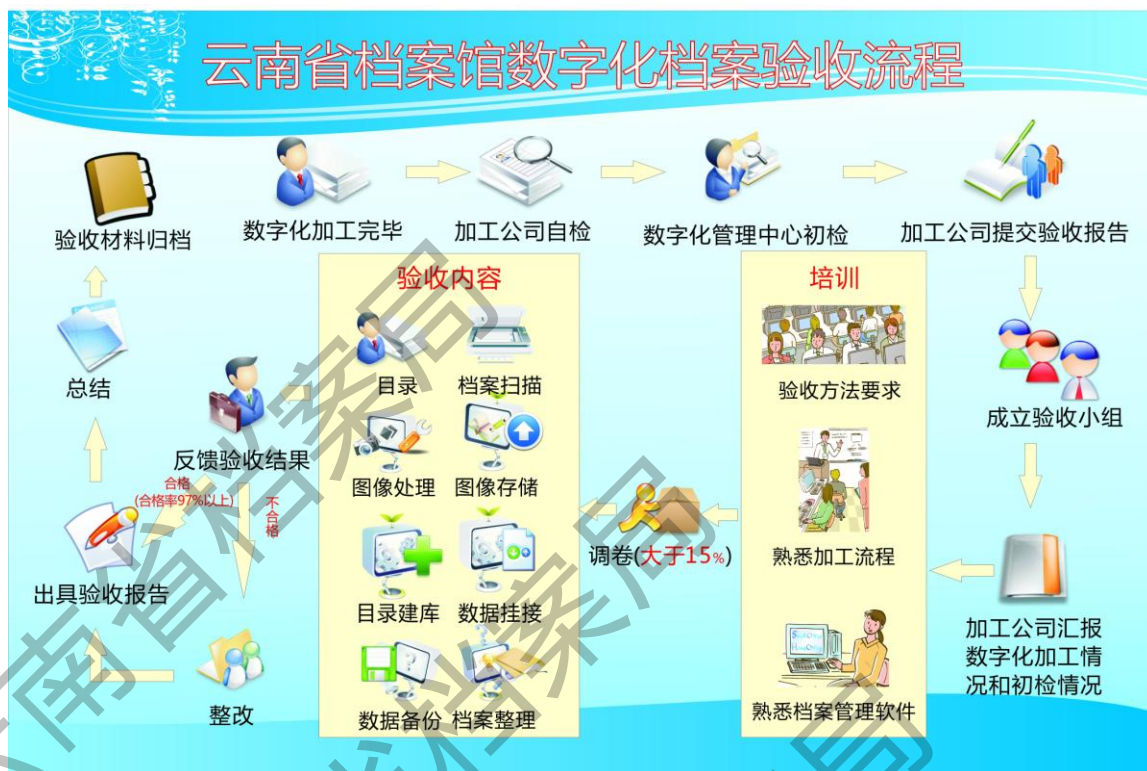
根据国家档案局《DA/T18-1999 国家档案著录标准》，对档号、责任者、题名、文号、文件日期、起止页号等项目著录质量分项验收。

2.5.8 档案整理验收

检查数字化工作完成之后是否按规定将档案复原装订等。

2.5.9 数据备份验收

检查数字化加工公司是否按照合同要求备份三套所加工档案数据，分别为两套 TIFF 格式和一套 PDF 格式，移交本单位。



注：以上流程图为我馆刚开始档案数字化加工时制作，随着工作的深入和相关制度规范的完善，部分类目及指标已作调整。

三、项目组织实施情况

3.1 项目组织

3.1.1 项目研究

云南省档案局（馆）承担《纸质档案数字化验收标准体系研究》科研项目研究工作。主要负责项目计划、项目研究、项目实施、落实项目责任，保证项目进度与项目成果质量等。

项目研究人员组成情况：研究生学历 2 人、本科学历 6 人；研究馆员 1 人、副研究馆员 3 人、馆员 3 人；其中，4 名人员曾获国家、省级科研奖励。

3.1.2 项目基础调研

《纸质档案数字化验收标准体系研究》科研项目以我局（馆）十八期馆藏档案数字化验收情况为基础开展分析研究工作。首先由数字化加工公司自查档案目录数据和图像文件总体完成质量，再经数字化管理中心进行初验，最后提交全省数字化验收专家组验收。全省数字化验收专家组和数字化管理中心均采用抽检的方式，以全宗为单位，检查目录数据、图像文件总体质量。抽检合格率达到98%以上（含98%），给予验收“通过”。（注：加工公司质检员逐条逐幅自检，数字化管理中心抽检比例不低于30%，全省数字化验收专家组抽检比例不得低于20%。合格率=抽查合格的文件数/抽查文件总数×100%）。

验收情况详见：云南省档案局馆藏档案数字化完结全宗验收情况汇总表

云南省档案局馆藏档案数字化完结全宗验收情况汇总表

验收期数	加工公司	验收全宗	目录数	案卷数	文件级目录数	画幅数	文件级目录合格率	图像扫描合格率	验收总评
第一期	***公司	****	12	3182卷	65868条	507355幅	99.90%	99.80%	合格
		****	1	1279卷	24813条	342665幅	99.90%	99.70%	
第二期	***公司	****	18	6858卷 /7338件	74462条 /7338条	1151406幅 /170836幅	98.80%	99.90%	合格
第三期	***公司	****	25	7802卷	148332条	1356230幅	98%	99.90%	
第四期	***公司	****	7	3729卷 /3577件	87115条 /3577条	728517幅 /29394幅	99.30%	99.80%	合格
第五期	***公司	****	21	9609卷 /10499件	123749条 /10499条	1439817幅 /268110幅	98.83%	99.74%	

第六期	***公司	****	39	5484卷	110849条	899513幅	98.90%	99.60%	合格
	公司	*		382卷	32093条	82949幅	99.50%	99.50%	99.50%
		****		6盒	415条	1206幅			
		****		270卷	846条	58720幅			
第七期	***公司	****	3	344卷	6382条	60870幅	99.20%	99.80%	99.37%
		****	10	7666卷	139759条	1163442	98.50%	99.40%	98.47%
第八期	***公司	****	36	8660卷	77590条	1042163幅	99.60%	99.97%	99.75%
第九期	***公司	****	4	1402卷 /7615件	20525条 /7615条	375736幅	98.80%	99.18%	98.99%
		****	2	517卷 /1108件	8280条 /1108条	126184幅	99.14%	99.26%	99.20%
		****	9	5053卷	65497条	809257幅	99.30%	99%	99.15%
第十期	***公司	****	18	11996卷	457787条	4009909幅	98.33%	99.99%	99.16%
第十一期	***公司	****	16	12868卷	141857条	1634965幅	99.81%	99.99%	99.93%
第十二期	***公司	****	22	9020卷	128941条	1281367幅	99.10%	99.90%	合格
第十三期	***公司	****	40	9810卷 /4362件	133651条	1682793幅	99.52%	99.90%	合格
第十四期	***公司	****		8677卷	145725条	1261596幅	99.93%	99.99%	合格
第十五期	***公司	****		7328卷	113987条	1161063幅	99.58%	99.97%	合格
第十六期	***公司	****		8329卷	79787条	1108261幅	99.95%	99.99%	99.99%
第十七期	***公司	****		10116卷	134380条	1319148幅	99.43%	99.99%	100.00%
第十八期	***公司	省****		13407卷	272250条	3134952幅		99.98%	100.00%

3.1.3 项目完成情况

《纸质档案数字化验收标准体系研究》科研项目完成情况：

第一阶段：2011年3月至10月，调查、分析、研究及总结云南省档案局馆数字化成果验收的经验；

第二阶段：2011年11月至12月，对收集的材料进行筛选并进行深入的研究，凡调研结果均在各数字化加工公司进行论证、检验并拟写出项目研究成果有关书面材料；

第三阶段：2012年1至10月，以文字、报表、表单等形式完成整个项目的研究工作，提交国家档案局组织验收并结题。

3.2 项目研究人员、经费投入及使用情况

3.2.1 项目研究人员情况

略

3.2.2 经费投入及使用情况

按照项目计划任务书预计所需项目咨询、评审经费为**万元。

项目组人员赴省外*****、省内*****等地档案部门开展数字化工作调研，实际经费具体投入使用*****元。

四、项目研究成果

《纸质档案数字化验收标准体系研究》科研项目现已获得以下几方面的研究成果：

4.1 数字化加工安全保密验收

安全保密包含两方面的内容：实体安全保密和信息安全保密。通过软硬件及人员管理，实现全方位的安全管理。包括网络平台搭建、监控设备的安装（16个摄像头）等。

本单位的数字化管理中心对加工公司实行流程管理，通过对加工人员进行岗前培训，增强人员安全意识、保密意识、责任意识等，确保档案在数字化加工过程的安全完整。

首先，验收小组检查加工公司是否与省档案局签订“信息保密协议”、公司是否办理《中华人民共和国国家秘密载体复制许可证》、工作人员是否持有云南省保密局颁发的《涉密人员上岗培训证》。

第二、档案原件是否严格按照档案管理要求进行管理。在未拆卷以前，如发现有缺失文件现象，必须填写《馆藏档案卷内文件材料登记备查表》，经签字认可后再进行拆卷处理。

第三、下班之前，所有档案是否进箱、柜，档案箱、柜钥匙要由加工公司专人保管，并且要由专人检查水电、窗户是否已关闭完好后，方可离去。

第四、在加工过程中，如发现文件属密级以上，是否进行登记，文件所涉内容。同时，不能随便在文件上添写任何字及符号，只能用 2b 铅笔编写页码。

第五、严禁将手机、相机、摄像机、私有移动存储介质等设备带入加工现场。由于工作需要，每个公司只允许两名管理人员可以在加工现场使用手机，其余员工只能在加工现场外或下班之后才能使用手机，如

有违反上述规定，一律严惩。

第六、验收小组查看现场时，注意下列事项：数字化加工过程中所使用的扫描仪、声频（视频）采集卡等设备，必须符合有关保密技术标准和保密要求，主机不得有U盘插口，数字化加工所使用的计算机及其存储介质，必须严格遵守计算机有关保密管理规定。使用的移动硬盘、U盘、光盘等存储介质，按涉密载体进行管理，不得在非涉密计算机及其网络上使用，并且每天公司拷贝数据所用硬盘等，必须由专人在管理中心进行登记后，方可领取使用，同时履行归还登记手续。加工公司在加工合同履行结束后撤离现场时，所使用的电脑硬盘、移动硬盘等数据载体需移交业主单位。

4.2 档案扫描验收

档案扫描是档案数字化加工的中心环节，加工公司必须根据所扫描档案幅面的大小选择相应规格的扫描仪或专业扫描仪进行扫描。档案扫描验收，主要检查扫描色彩模式是否按照合同的分辨率要求和模式情况进行，是否反映档案的原貌等。刚开始数字化加工工作的时候，我们也有过一些争议，主要就是在档案是提供一般性利用，还是可以提供特殊利用，比如原文出版等，因为两者所要求的存储空间差别很大。根据这几年来数字化扫描加工工作的进一步发展，存储介质的飞跃发展，我们也逐渐统一了思路，规定了档案扫描的一些基本标准。比如，扫描分辨率：文件均采用彩色色彩模式，分辨率 $\geq 300\text{dpi}$ ；照片档案： < 5 寸、 $600 \times N$ ， 5 寸、 600dpi ， > 5 寸、 $600 / N$ 。（注： 5 寸照片采用 600dpi

分辨率；原照片小于 5 寸照片 N 倍的，则分辨率为 $600 \times N$ ；原照片大于 5 寸照片 N 倍的，分辨率为 $600/N$ 。但在数字化扫描及验收的过程中，发现了很多意想不到的特殊情况。例如：

4.2.1 在扫描中发现档案纸张存在油性纸、复写纸、铅笔字、字迹模糊、纸张较差及较薄等情况。具体做法是：在保持档案原貌的前提下进行扫描，并提高分辨率或采用垫白纸再进行扫描，尽量反映档案的原始原貌。

4.2.2 在扫描中发现部分档案原件纸张的底色较深，所标页码扫描后看不清楚。具体做法是：用一张白色小纸条写上该页页码，放在该页右上角，然后再进行扫描。

4.2.3 在扫描中发现有重份档案、重份照片、重份表格等情况。具体做法是：只扫描一份。

4.2.4 在扫描中存在多张零散照片。具体做法是：将照片平铺在白纸上，并在白纸右上角标上页码后再进行扫描。

4.2.5 档案中存在部分超大、超长、图纸及表格，无法进行扫描。具体做法是：用白纸打上“此页为特大图纸”字样，扫描此页来代替特大图纸并在备注栏注明。如：



4.2.6 在扫描中发现档案上有多张粘贴的小纸条或部分档案纸张粘贴原文且破损严重等情况。具体做法是：为保持档案原貌，用平版扫描仪进行扫描，尽量将档案可视部分扫描出来，如果有遮挡部分，重新进行标页后再另扫一页。

4.2.7 档案中有部分大画幅（大于 A3 幅面）无法一次性扫描。具体做法是：分幅扫描后进行完整拼接或用相机拍摄。如：



4.2.8 档案数字化过程中存在照片底片、锦旗、奖状等实物档案、音像档案。具体做法是：把实物档案放于白纸上并在右上角标注相应的页码进行拍照或只对音像封面进行扫描。如：





4.2.9 在扫描民国档案时，很多是桐子页，由于纸张较破较薄，正反两面都有内容，如果正常扫描，就会把背面的内容印到正面。具体做法是：扫描时在桐子页中间夹放一张白纸。

4.2.10 在扫描民国档案时，档案原件存在纸张较差、破损严重，虫蛀、霉变、老化、缺字、字迹模糊等情况。具体做法是：尽量照档案原样进行扫描，字迹模糊的在扫描时垫上“原件字迹不清晰”进行扫描，有的老化已碎的，用专用信封将已碎的档案装袋不进行扫描。

4.2.11 在扫描民国档案中发现部分档案有严重粘贴或无法展平等情况。具体做法是：扫描成多幅图像，再进行图像拼接。

另外，我们要求各公司在扫描过程中，凡是发现或遇到档案存在的各种情况，都必须填写数字化加工备查表。如：

数字化加工备查表

全宗号	XX全宗	目录号		案卷号	
数字化加工公司	云南数字化	法人	张三	加工时间	2012年1月
拆卷人	田晓	扫描人	张呈文	装订人	毛艳平
数字化加工情况说明： 第1、159页原件破损/第1-5页原件严重破损/整本破损/整本霉烂/整本虫蛀/此卷有实物档案/第25、169页严重霉烂未扫描、装订/邮票12张					
填写要求：凡卷内纸质目录信息与卷内档案文件信息不一致、有重大改动和错漏字以及其他需要说明的。					

以上做法，是我们在数字化扫描过程中以及在进行多次验收的基础上，总结出来的常见特殊情况及处理方法，我们也对各公司进行了规范和统一。在数字化档案的验收过程中，我们也按照以上的标准，对抽查的每一份档案进行逐页的检查验收，让每一份、每一页档案都能做到有来源、有记录，最大限度的保证档案的齐全完整。

4.3 图像处理验收

档案扫描图像处理结束后加工公司必须对扫描图像的偏斜度、清晰度，失真度等进行质检，对存在偏斜、污渍、不完整等问题的扫描图像进行纠偏、去污、拼接等技术处理，确保图像完整无误。加工公司质检完成后须认真填写“档案数字化流程工作单”，对图像处理结果进行相应说明。

档案数字化加工流程单

本卷	全宗号	目录号	案卷号	总页数	页	备注
前 处 理	实际整理:		条目录	张纸	签名:	
	存在问题	A 纸张破损 B 跳号 C 内容缺失 D 内容模糊 E 原件污渍 F 原件粘贴 G 原件缺失 H 超大幅 J 缺字 K 内容重复 L 漏编页码 M 重条 N 跳条 O 缺条 P 题名不正确 Q 多起止页号 W 页码重编 R 重条 S 虫蛀				
	整理结果					
前 处 理 检 查	实际检查:		页		签名:	
	存在问题	A 漏编页码 B 页码重编				
	整理结果					
著 录	实际扫描:		条		签名:	
	存在问题					
	整理结果					
质 量 检 查 (条 目)	实际处理:		条		签名:	
	存在问题					
	整理结果					
文 件 扫 描	实际处理:		幅		签名:	
	存在问题	A 大于 A3 B 大于四张 A3 C 铅笔字 D 灰度 E 真彩 F 原件内容模糊 G 原件污渍 H 原件粘贴				
	整理结果					
图 像 处 理	实际核查:		幅		签名:	
	存在问题	A 纸张破损 B 字迹不清 C 原件模糊 D 印章不清 E 原件内容缺失 F 拼接大图 G 杂点多 H 重页 J 背面 K 重扫				
	整理结果					
质 量 检 查 (图 像)	实际核查:		幅		签名:	
	存在问题	A 字迹不清 B 背面 C 字被遮 E 字被裁 E 空白页 F 文件名不对 G 图像处理不干净 H 纠偏 J 去污 K 裁边				
	整理结果					
装订还原	<input type="checkbox"/> 贴边处理		<input type="checkbox"/> 无遗漏		签名:	
装订检查	<input type="checkbox"/> 页码颠倒		<input type="checkbox"/> 订漏		签名:	

初验和验收，重点从图像数据完成质量、是否纠偏、是否去污三个方面进行审核检查。在图像处理过程中，要求保持档案信息真实原貌，不得擅自修改档案信息。

图像数据质量：对图像偏斜度、清晰度、失真度、完整性及漏扫、排列顺序与原件的一致性进行检查。

纠偏：对出现偏斜的图像必须进行纠偏处理，以达到视觉上基本不感觉偏斜为准，允许的倾斜度不得大于1度；对方向不正确的图像必须进行旋转还原，以符合阅读习惯。以达到视觉基本不感觉偏斜为准，并符合阅读习惯。

去污：对图像页面中出现的会影响图像质量的杂质，如黑点、黑线、黑框、黑边等必须进行去污处理。处理过程中必须遵循在不影响可懂度的前提下展现档案原貌的原则。

档案扫描图像处理结束后加工公司必须对扫描图像的偏斜度、清晰度，失真度等进行检查。

对不符合图像质量要求的，重新进行图像处理。对图像文件不完整或无法清晰识别的，重新扫描。对漏扫文件必须及时补扫并正确插入图像。扫描图像的排列顺序与档案原件不一致时，必须及时进行调整。图像质量力求图像完整、清晰，无影响图像美观的黑边和污点，图像不偏斜或倒置，符合正常阅读习惯。对初验和验收过程中图像存在的质量问题，数字化管理中心和全省专家验收委员会认真填写“数字化档案质量验收记录表”图像处理情况，督促加工公司对有问题的图像数据进行相应的整改。

格式进行图像数据备份，数据挂接备份则采用 PDF 格式进行存储。档案原文图像数据审核验收主要检查加工公司在档案图像处理完成后是否按图像格式进行数据存储，凡存储格式不符合以上规定而采用 JPG 等格式保存的数据，一律重新处理图像格式。图像存储的验收主要是检查数字化加工公司所提交的光盘或者移动硬盘等其它存储介质，其图像存储的格式是否以 TIFF（LZW）格式和 PDF 格式存储；存储介质中的每一份文件，是否都有一个与之对应的唯一档号，并以该档号为扫描后的该份图像文件命名。

首先检查数据格式是否标准，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；其次检查档案的著录，采用人工校对或自动校对的方式，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、内容是否规范和准确。

4.5 目录建库验收

目录建库工作是此研究项目体系中一项很重要的组成部分，它包括数据格式选择、档案目录数据库的建立和档号唯一性选择，在档案目录数据库中每一份文件赋予一个与之相对应的唯一的档号，该档号又作为这份文件扫描后的图像文件命名的唯一性，在目录数据库建库工作中显得十分重要。

4.5.1 数据格式选择

我们在纸质档案数字化验收标准体系研究项目中，对目录建库主要是选择通用数据格式（DBF、TXT、MDB），各数字化加工公司的数据移交

时统一转换为·DBF格式，向云南省档案局（馆）的局域网数据服务器进行导入，其所选定的数据格式都能直接或间接通过XML文档进行数据交换。

4.5.2 档案目录数据库的建立

按照国家档案局《档案著录规则》（DA/T 18）和《云南省档案馆档案文件级目录著录要求》等规定项目建立规范化的档案目录数据库，其结构如下：

现行档案文件级目录数据库结构

字段	字段名	类型	宽度
1、	全宗号	C（字符型）	4
2、	目录号	C（字符型）	3
3、	案卷号	N（数值型）	4
4、	卷内顺序号	N（数值型）	4
5、	档号	C（字符型）	40
6、	案卷档号	C（字符型）	40
7、	责任者	C（字符型）	100
8、	题名	C（字符型）	250
9、	人名索引	C（字符型）	250
10、	文件编号	C（字符型）	80
11、	文件日期	T（日期型）	8
12、	日期备注	C（字符型）	12
13、	起页号	N（数值型）	4
14、	止页号	N（数值型）	4
15、	页数	N（数值型）	4
16、	保管期限	C（字符型）	4

17、	录入时间	T (日期型)	8
18、	录入员	C (字符型)	10
19、	录入部门	C (字符型)	16
20、	扫描部门	C (字符型)	16
21、	密级	C (字符型)	4
22、	备注	C (字符型)	40

历史档案文件级目录数据库结构

字段	字段名	类 型	宽度
1、	全宗号	C (字符型)	4
2、	目录号	C (字符型)	3
3、	案卷号	N (数值型)	5
4、	卷内顺序号	N (数值型)	3
5、	档号	C (字符型)	40
6、	案卷档号	C (字符型)	40
7、	责任者	C (字符型)	100
8、	题名	C (字符型)	250
9、	文件日期	T (日期型)	8
10、	日期备注	C (字符型)	12
11、	起页号	N (数值型)	4
12、	止页号	N (数值型)	4
13、	页数	N (数值型)	4
14、	保管期限	C (字符型)	4
15、	录入时间	T (日期型)	8
16、	录入员	C (字符型)	10
17、	录入部门	C (字符型)	16
18、	扫描部门	C (字符型)	16

19、	密级	C (字符型)	4
20、	备注	C (字符型)	40

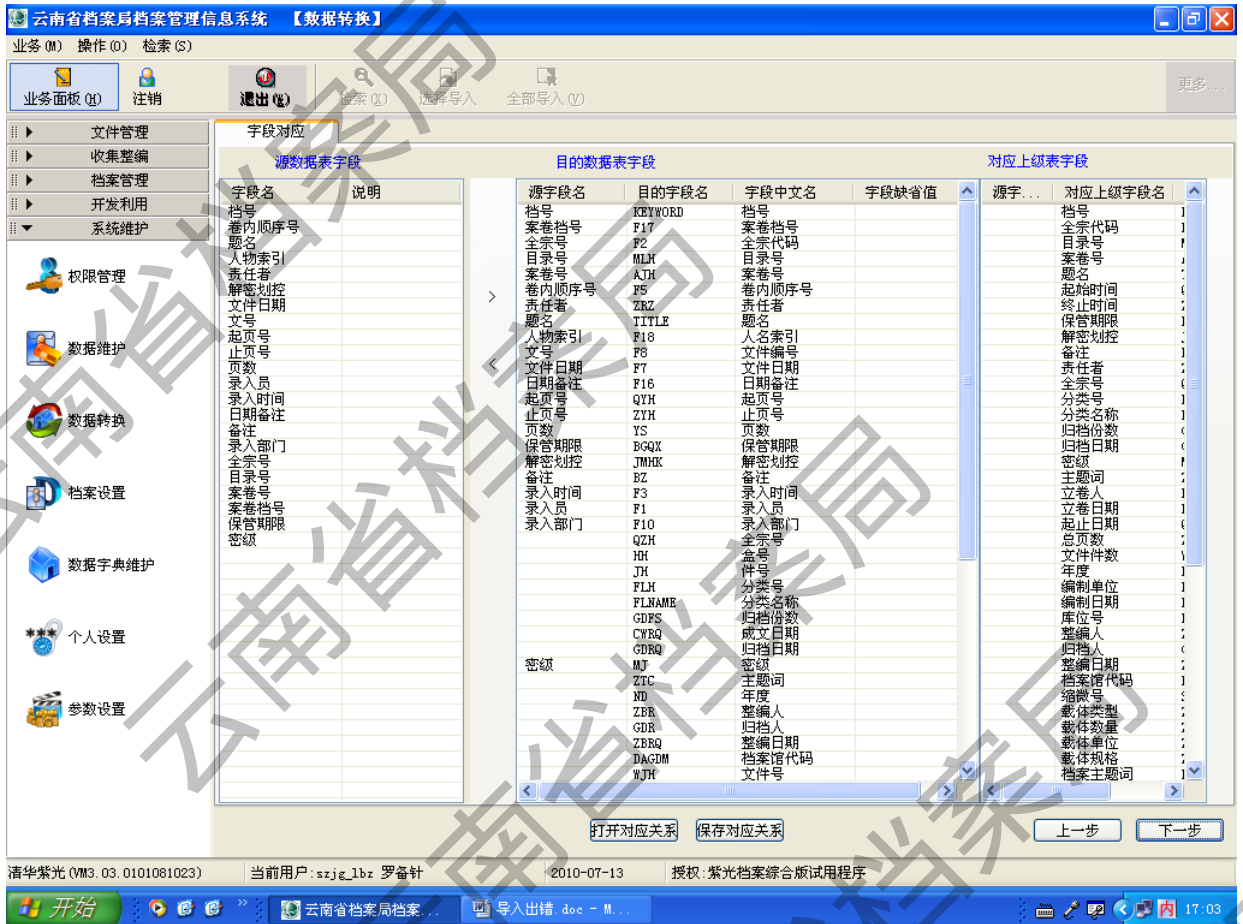
归档文件（新规则）目录数据库结构

序号	字段名称	字段类型	字段长度
1、	档案馆代码	字符型	6
2、	编制单位	字符型	30
3、	全宗号	字符型	4
4、	年度	字符型	4
5、	保管期限	字符型	4
6、	机构（问题）	字符型	20
7、	件号	数字型	5
8、	档号	字符型	60
9、	责任者	字符型	100
10、	题名	字符型	250
11、	人名索引	字符型	250
12、	文件编号	字符型	100
13、	时间	日期型	8
14、	日期备注	字符型	12
15、	页数	数值型	4
16、	密级	字符型	4
17、	备注	字符型	30
18、	录入时间	日期型	8
19、	录入员	字符型	20
20、	录入部门	字符型	20
21、	扫描部门	字符型	16

4.5.3 档号唯一性选择

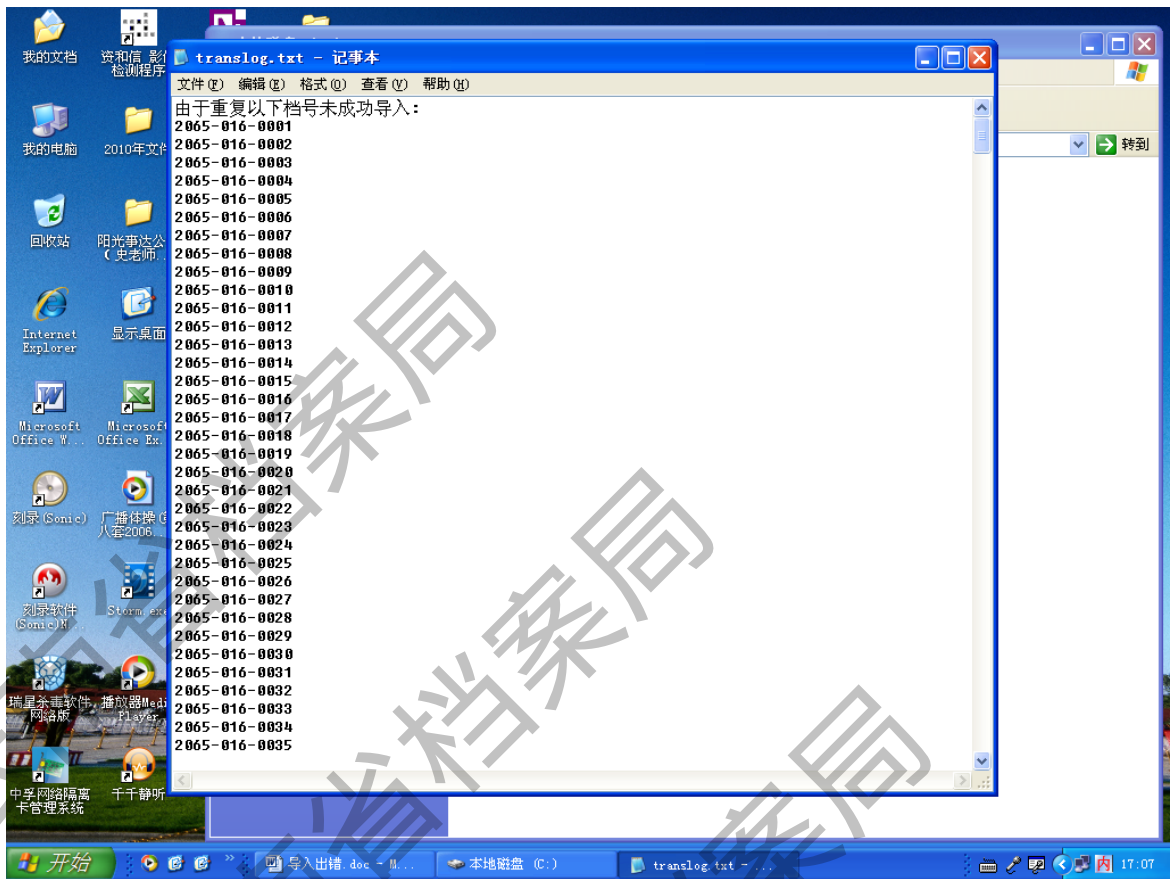
档案目录数据库中的每一份文件都有一个与之相对应的唯一档号，必须以该档号作为这份文件扫描后的图像文件命名。

【目录数据导入的字段对应关系】

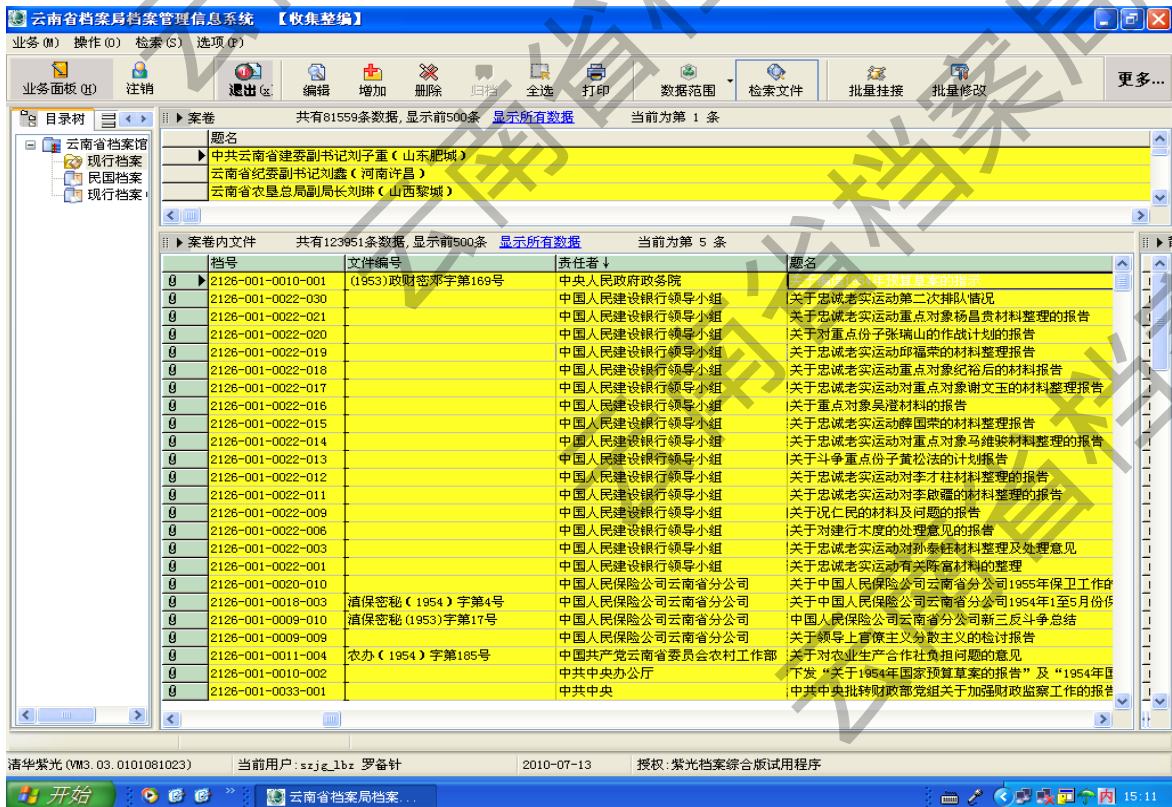


如果档号重复，系统提示退出不予以导入。如下图的数据就是档号未成功导入目录的提示：

【未导入数据详细记录提示】



【检查核对目录内容】

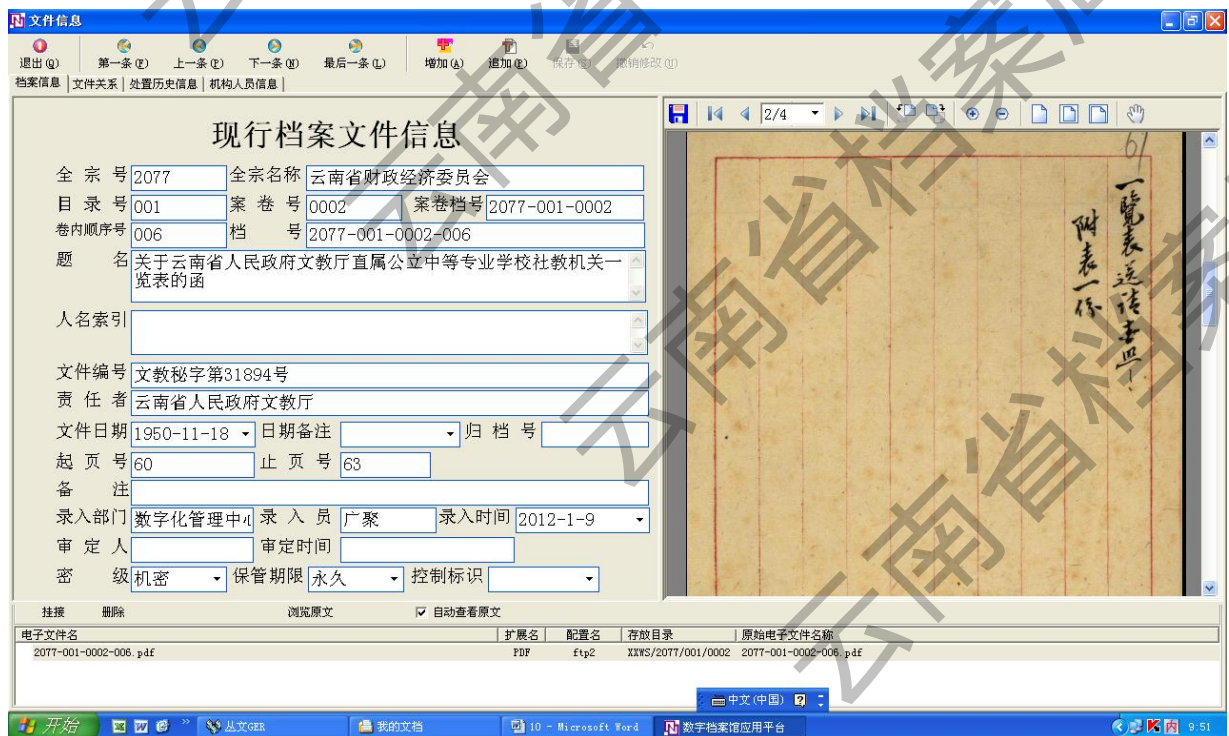


4.6 数据挂接验收

挂接是档案数字化转换过程中形成的目录数据库与图像数据库，通过质检环节确认为“合格”后，各数字化加工公司将其加载到云南省档案局（馆）的数据服务器上。通过紫光档案管理系统软件，以纸质档案目录数据库为依据，将每一份纸质档案文件扫描所得的一个或多个图像存储为一份图像文件。通过每一份图像文件的文件名与档案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，便可批量、快速地进行挂接。

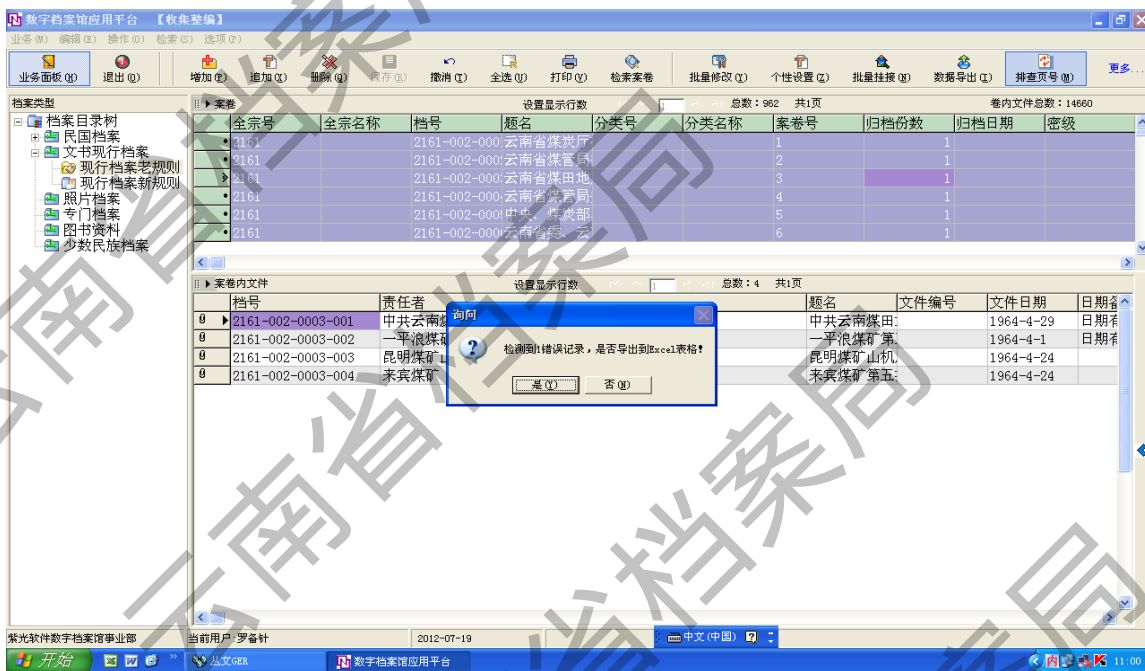
数字化加工公司将每一份纸质档案文件扫描所得的一个或多个图像采用多元 TIFF 格式存储成单个图像文件夹，同时与云南省档案馆档案目录数据库对接，将图像数据挂入数据库中。

[图像挂接后的结果]

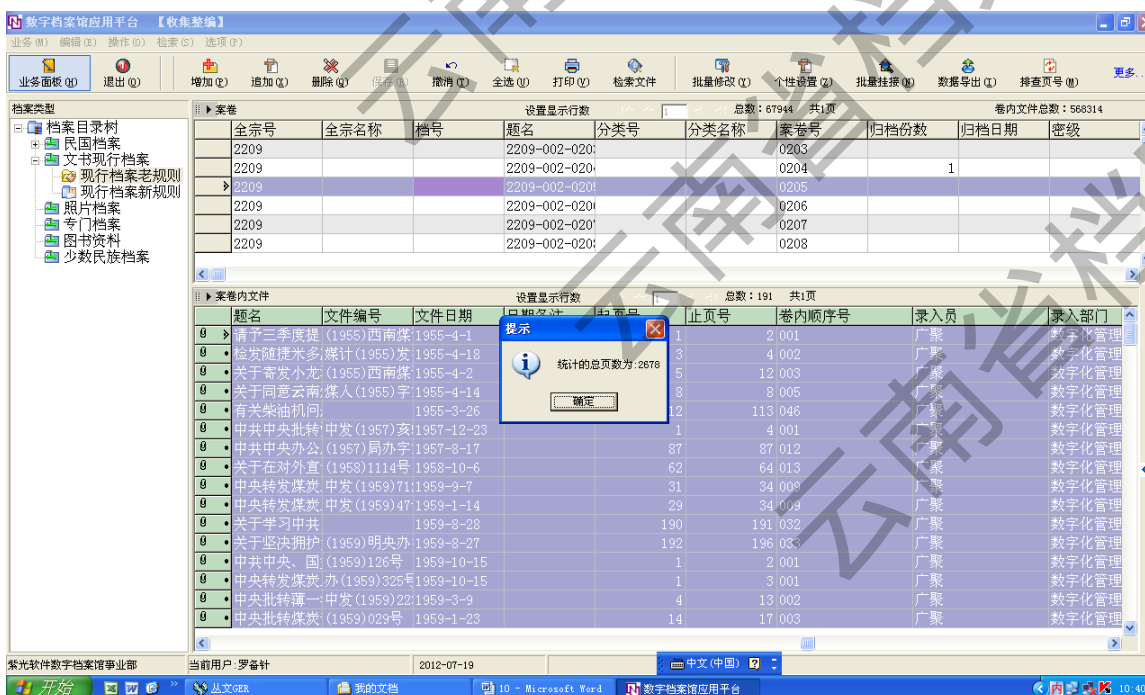


数据挂接验收，主要是检查图像数据与档案实体的页码总数是否一致，图像文件的名称与档案目录数据库中该份文件的档号、标题是否相同，图像文件总数与目录总数是否一致，挂接的数据是否有重复、遗漏等。

[目录与图像文件检查方法]

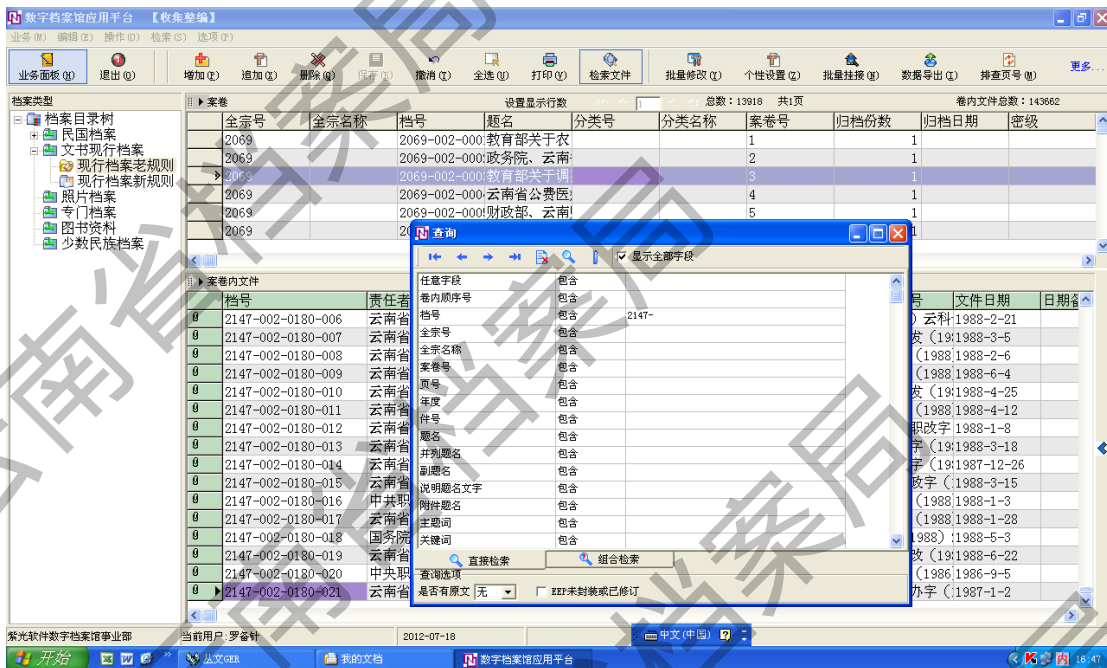


[图像文件总页统计方法]



在挂接后，我们使用档案管理软件即可测试出其是否多挂、少挂或漏挂。如下图是查找漏挂图像的检测方法：

[图像漏挂检测方法]



检查挂接情况，数据挂接后，以档案文件级目录数据库为依据，对挂接的图像文件进行检查，核对图像文件的命名格式是否符合要求，发现错误及时做出修正。

4.7 档案文件级目录及题名验收

在本类目的数字化档案验收中，我们根据国家档案局《DA/T18-1999 国家档案著录标准》，并结合《云南省档案馆现行档案文件级目录著录与扫描技术要求》和《云南省档案馆历史档案文件级目录著录及扫描要求》的具体规定对全宗号、目录号、案卷号、档号、责任者、题名、文号、文件日期、起止页号等项目著录质量分项进行。采用的方式是按照

抽查比例，从每个验收档案全宗中随机调取部分档案原件，对照数字化档案条目和档案原件逐条逐页进行对比验收。具体的验收标准主要有以下几个方面：

4.7.1 全宗号：是由档案馆给立档单位编制的代号。如：2074、2126等。全宗号的著录在系统软件设计中为自动继承，正常情况下不会有错漏。

4.7.2 目录号：对每一本案卷目录的编号。用阿拉伯数从1依次标注（1、2、3……）。目录号的著录在系统软件设计中为自动继承，正常情况下不会有错漏。

4.7.3 案卷号：对每一个案卷的编号。用阿拉伯数从1依次标注（1、2、3……）。案卷号的著录在系统软件设计中不得为空，也不能重复，因此可以杜绝空号和重复现象。

4.7.4 卷内顺序号：案卷内文件统一用阿拉伯数字排序，无空号、重号。卷内顺序号的著录在系统软件设计中不得为空，也不能重复，因此也可以杜绝空号和重复现象。

4.7.5 档号（系统自动生成）：全宗号-目录号-案卷号-卷内顺序号，例如：2065-001-0001-001，全宗号（4位）；目录号（3位）；案卷号（4位）；卷内顺序号（3位）。

4.7.6 案卷档号（系统自动生成）：全宗号-目录号-案卷号，例如：2065-001-0001）。

4.7.7 文件责任者：是指文件的形成者，它包括形成文件的机关或个人，可以是一个机关或机关的一个组织机构，也可以是两个以上的机

关或个人。

(1) 责任者按红头文件上的机关名称全称或落款印章上的机关名称全称以及个人名称照实填写，一般不得简写或省略。个人责任者要著录姓名并尽量考证著录其所在单位、部门及职务。对个别责任者全称过长的，可使用通常使用的规范化简称。

(2) 对红头文件机关名称上只有一个责任者的，而文件落款上有两个以上落款单位的，“责任者”应以落款单位为准。

(3) 多个责任者按主次顺序尽量写全，若责任者太多，超过规定字段，著录前两个责任者，责任者与责任者之间用顿号“、”隔开。

(4) 同一个全宗相同责任者名称应统一。

(5) 文件无责任者的或责任者无法考证的可以为空。

(6) 责任者不能出现“本省”、“本市”、“本县”、“本厅”、“本委”、“本局”、“本处”、“本院”、和“我省”、“我市”、“我县”、“我厅”、“我委”、“我局”、“我院”等。

(7) 对各种申报表、统计报表等的责任者应使用填报单位。

(8) 凡转发文件的以转发文单位作为责任者。

4.7.8 题名：文件题名是揭示文件内容与成份的重要著录项目，是表达档案中心内容形式特征的名称，能够准确、全面揭示文件内容，结构规范，文字通顺、简洁、精炼，采用中性文字，不带政治立场、观点，文字尽量与原文相吻合，还可使用模糊检索或关键字检索查找。本类目为数字化档案验收的重点部分，在验收过程中有以下具体的标准要求：

(1) 题名应忠实于原有纸质目录，原则上不进行改动。如原纸质

目录上有明显错、漏字的，可在目录上进行修改，但不得在原档案上作任何涂改。

(2) 文件题名主要以一事一件为主，具体要求如下：

①文件的正本与定稿（底稿）、附件应为一件，著录为一条目录。

②转发文与被转发文为一件，应著录为一条目录。

③来文与复文（即请示、报告与批复）为一件，应以批复来著录文件题名。

(3) 题名可使用顿号（、）、书名号（《》）、引号（‘、’、“”）、破折号（—）；但不能使用逗号（,）、分号（;）、句号（。）等。繁体字一律改为简写字，对于无法考证的字用“□”表明。

(4) 对部份文件题名上只写“通知”、“电报”、“请示”、“通报”、“简报”、“记录”、“纪要”、“电话记录”、“文件材料”、“情况反映”、“调查研究材料”、“密码电报”、“明传电报”等未能揭示文件内容的，要根据文件内容拟定文件题名。

(5) 对干部任免的文件，要将所有任免干部姓名一一录入。

(6) 对离退休、授予称号、干部档案、行政介绍信、组织关系介绍信、调整职工工资花名册以及某会议参会人员名单等材料均不录入姓名。

4.7.9 人名索引：用来录入人名。对题名项人名录不下的，均录在人名索引字段上。

4.7.10 文种：符合公文文种要求并与档案文件文种相一致，能够准确反映文件的时代特征、行文关系、重要程度和发文范围。

4.7.11 文号：文号是文件的编号。是文件制发过程中由制发机关、团体或个人赋予文件的顺序号，具体是指文件的发文号。文号一般由机关代字、年度、顺序号组成，是责任者制发文件时编写的文件顺序号，便于文件的查找利用。著录文号时要求照原文件发文号著录，年份要用4位数表示。要确保文号完整准确，不能随意省略，以免给查找利用档案造成困难。如有多个文号时录入主要责任者的文号，作者与文号要对应录入。

4.7.12 文件日期：就是文件形成时间，文件日期应填写文件落款日期。日期统一使用公元纪年日期，用阿拉伯数字标识，以通用国标8位数字为标准，对月、日是个位数的，要在数字前加“0”，如：1995-01-08

（一般是指在字符型字段或使用 Excel 办公软件录入的方式）。如果系统是日期型的，就直接录入“1995-1-8”即可；文件无时间但能够从档案内容中考证出来的，要求填注考证时间，年代无法考证的应在“日期备注”项内填写相邻时间，并注明“年代有待考证”。没有月份的就录入1月，注明“月份有待考证”；没有日期的就填上1日，注明“日期有待考证”。

4.7.13 日期备注：用于档案内容无时间时的注明。年代有待考证（参照上条目录年代）、月份有待考证（均填写1月）、日期有待考证（均填写1日）。

4.7.14 起止页号：目录中所填写的起止页码必须与档案实体页码相吻合，无差错，无空缺。

4.7.15 保管期限：照卷壳上录入，如卷壳上没有标明保管期限的，

该录入项为空。

4.7.16 密级：照录。如：绝密、机密、秘密等。

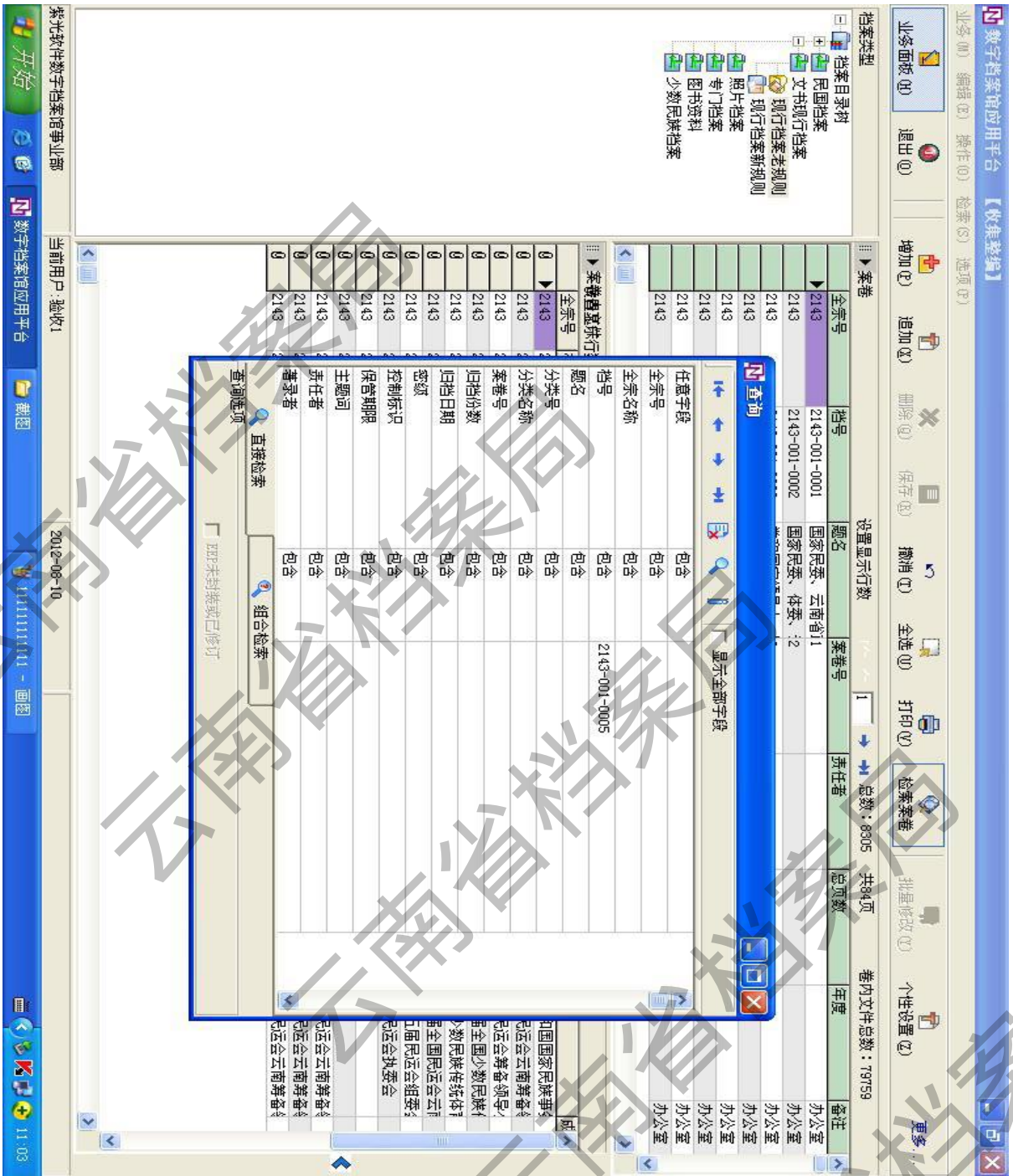
4.7.17 录入时间：当天的录入时间。由系统软件自动生成。

4.7.18 录入员：填写录入人员姓名或公司名称。

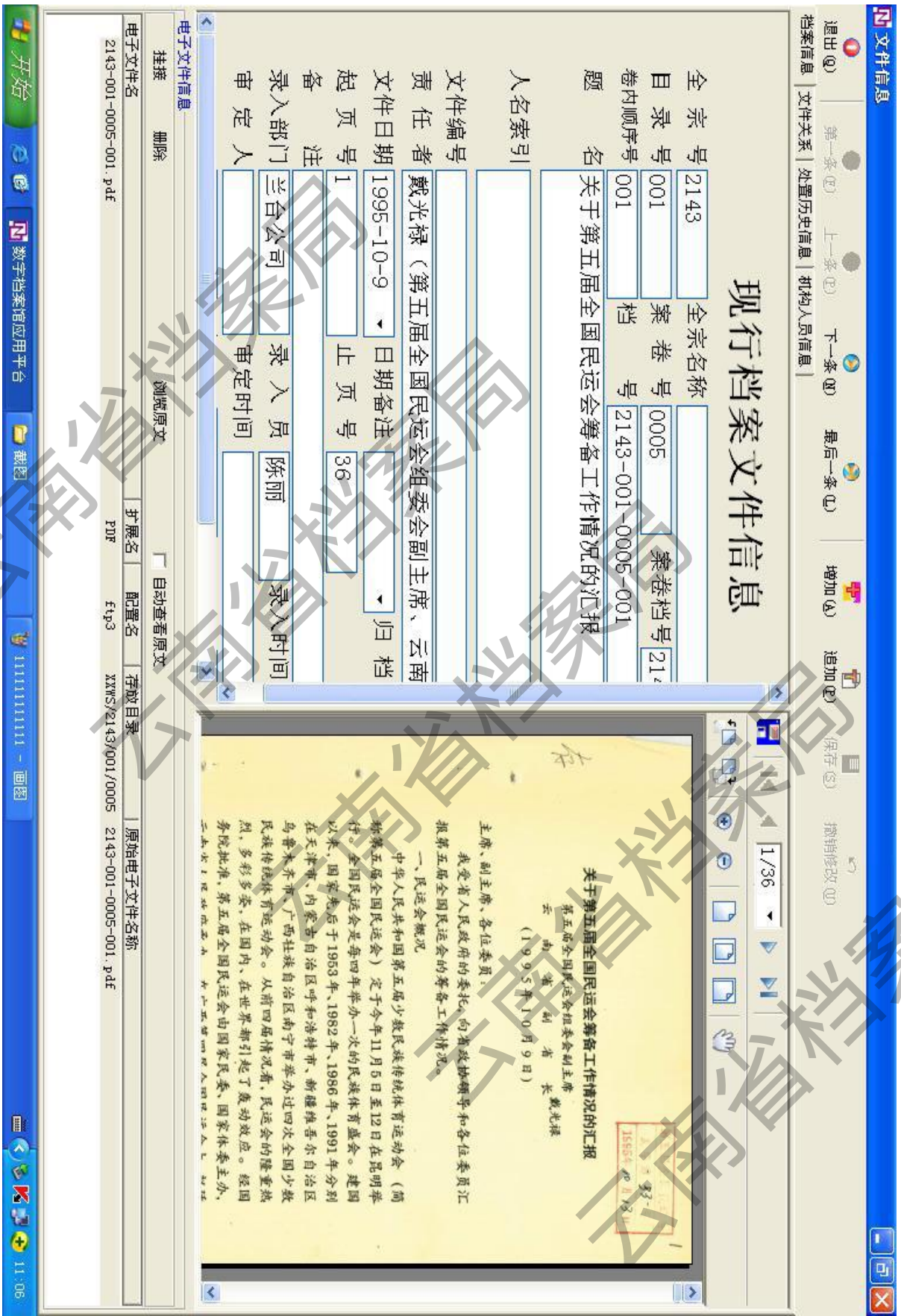
4.7.19 录入部门：填写加工公司名称。

4.7.20 备注：备注填写各份文件需要说明的问题或文件的载体形式。如：虫蛀、破损、字迹不清、需修裱和‘此页为特大图纸用相机拍摄未扫描’等，以方便进行下一步的管理及修复等工作。

以上著录项是根据国家档案局《DA/T18-1999 国家档案著录标准》，并结合《云南省档案馆现行档案文件级目录著录与扫描技术要求》和《云南省档案馆历史档案文件级目录著录及扫描要求》的具体规定对档号、责任者、题名、文号、文件日期、起止页号等项目著录质量分项进行验收的。通过这个类目的验收，可以完整揭示档案的基本内容，也为每一份档案的查找利用提供了完善的信息条件，大大提高档案的查全率和查准率。



云南省档案馆数字集成管理系统数字化验收模块操作界面（截图1）



云南省档案馆数字集成管理系统数字化验收模块操作界面（截图2）

4.8 档案整理验收

扫描工作完成后，加工公司对拆除过装订物的档案必须按档案保管要求采用三孔一线的方法重新装订，装订必须保护档案的实体不被损坏，保持案卷内文件的排列顺序不变，做到安全、准确、无遗漏，装订成品结实美观。

档案整理验收工作主要检查是否拆除装订物，并保护档案不受损害；是否按要求把同一案卷中的扫描件和非扫描件区分开；是否制作并填写数字化档案加工流程登记单并详细记录档案整理后每份文件的起始页（件）号和页数；扫描完成后，拆除过装订物的档案是否按档案保管的要求重新装订。

4.9 数据备份验收

以上各项数据验收合格后，还有最后一道验收工序，即对加工公司备份的数据验收。根据我们的要求，备份介质采用一次性写入光盘或移动硬盘。至少备份三套，分别为两套 TIFF 格式和一套 PDF 格式。在验收当中需要注意的是：两套 TIFF 格式的存储介质，必须是不同的品牌，以防止同批次的存储介质出现质量问题，从而影响所有数据都不能使用。

数据备份验收工作主要检查图片数据备份和目录数据备份和备份数据是否能被计算机正确打开和读取、数据信息是否完整、文件数量是否准确等，备份介质是否有揭示存储介质内容的标签等。

硬盘标签式样：

全宗号	××××全宗				
格式	硬盘型号	案卷数	条目数	影像数	数据容量
TIFF 盘 1	Seagate	3280	56189	400185	1.78TB
		新规则	11009	177971	
加工公司	×××公司				
加工时间	2010.02-2010.10		移交时间	2010.12	

实物图：



云南省档案馆纸质档案数字化成果移交表（光盘）

全宗号	全宗名称	目录数	案卷级目录数量 (条)	文件级目录数量 (条)	影像数量(幅)
光盘内容及数量	内容		数量		
			套(TIFF)	共 片	
			套(TIFF)	共 片	
			套(PDF)	共 片	
检查情况	检查项目		检查情况		
	载体外观检查				
	病毒检查				
	可读性检查				
	目录数据检查				
	图像数据检查				
	其他检查				
交接说明					
移交部门			接收部门		
公司		档案数字化管理中心	电子备份信息处		
经办人:		经办人:	经办人:		
负责人:		负责人:	负责人:		
移交日期: 年 月 日		日期: 年 月 日	接收日期: 年 月 日		
局领导审核意见					

附件：数字化数据清单

云南省档案馆纸质档案数字化数据移交表（移动硬盘）

全宗号	全宗名称	目录数	案卷级目录数量 (条)	文件级目录数量 (条)	影像数量(幅)
硬盘型号 及数量	硬盘型号		储存格式及数量		
			1套(TIFF)	共 盘	
			2套(TIFF)	共 盘	
			1套(PDF)	共 盘	
检查情况	检查项目		检查情况		
	载体外观检查				
	病毒检查				
	可读性检查				
	目录数据检查				
	图像数据检查				
	其他检查				
交接说明	以上档案原文数字化成果经云南省档案局档案数字化验收委员会验收，同意移交。				
移交部门			接收部门		
**** 公司		数字化管理中心		电子备份信息处	
经办人:		经办人:		经办人:	
负责人:		负责人:		负责人:	
移交日期: 年 月 日		日期: 年 月 日		接收日期: 年 月 日	
局领导审核意见					

附件：数字化数据清单

五、项目研究成果推广应用

档案数字化验收项目符合现代档案管理理念，注重档案数字化验收工作与档案数据库建设、档案信息网络开发的有机衔接，充分应用计算机软硬件功能，最大限度地发挥人力资源和数字化加工设备能力，较好地保证了数字化档案的真实可靠。项目组对档案数字化流程重组与再造、操作方法与步骤所进行的全面研究，实现了档案数字化数据的再次优化和管理系统的整体优化。数字化验收工作针对档案原文数字化转换后开展的全面的专业技术研究，较好地保证了数字化后档案实体整理工作的科学规范、数据著录的齐全完整、扫描对象的具体明确和数字化信息的经济实用。通过十八批次的数字化成果的测试与验收研究，我们逐步形成了一套行之有效的成果验收的标准和做法。从《纸质档案数字化验收标准体系研究》项目的学术价值研究表明，本项目研究成果系统完整，进一步补充与完善了国内档案数据质量检验细化标准及流程管理，进一步延伸并补充了《纸质档案数字化技术规范》内容，对目前全国档案系统数字化成果质量检验的技术指标和具体工作流程规范具有一定指导意义，对全国各档案馆的数字化成果测试评价及验收工作具有较大的参考价值。

六、项目研究报告及相关验收规定、办法、教材及表单

6.1 档案数字化验收规定、办法及教材

6.1.1 《云南省档案馆档案数字化管理中心工作规范》

- 6.1.2 《云南省档案馆数字化加工管理暂行办法》
- 6.1.3 《云南省档案馆数字化范围划定及数字化方法运用暂行规定（试行）》
- 6.1.4 《云南省档案馆数字化档案验收暂行办法》
- 6.1.5 《云南省档案馆数字化档案验收工作实施细则（暂行）》
- 6.1.6 《云南省档案馆馆藏档案卷内文件状况登记备查办法（暂行）》
- 6.1.7 《关于数字化档案数据光盘标识的规定（暂行）》
- 6.1.8 《云南省档案馆现行档案文件级目录著录与扫描技术要求》
- 6.1.9 《云南省档案馆历史档案文件级目录著录及扫描要求（教材）》
- 6.1.10 《关于档案数字化加工及管理过程中保密及安全管理方面的规定（教材）》
- 6.1.11 云南省档案馆数字化管理中心安全管理要求（挂图）（教材）

6.2 档案数字化验收申请、报告及表单

- 6.2.1 关于开展第 n 期数字化档案验收工作的请示
- 6.2.2 云南省档案馆第 n 期数字化档案验收小组名单
- 6.2.3 数字化档案验收会议议程与验收方式、验收总结
- 6.2.4 云南省档案馆馆藏档案数字化对象审批单
- 6.2.5 《云南省数字化档案质量验收申请表》
- 6.2.6 《云南省档案馆数字化档案验收流程（挂图）》
- 6.2.7 《数字化档案质量验收记录表》
- 6.2.8 《数字化档案质量验收记录汇总表》

6.2.9 《云南省档案馆数字化档案质量验收报告》

6.2.10 《云南省档案馆纸质档案数字化数据移交表（移动硬盘）》

6.2.11 《云南省档案馆纸质档案数字化成果移交表（光盘）》

6.2.12 《云南省档案馆数字化档案流程移交单（纸质）》

6.2.13 《云南省档案馆档案原文数字化加工项目竣工报告表》

6.2.14 《数字化加工备查表》

6.2.15 《档案数字化流程工作单》

6.2.16 《云南省档案馆馆藏档案卷内文件材料登记备查表》

6.2.17 《云南省档案馆档案数字化加工周工作量统计表》

6.2.18 《各公司移动硬盘及U盘交接登记表》