**雅砻江流域水电开发有限公司**

**数字档案馆系统数据鉴定与迁移工作规范**

**2021年8月**

**目 录**

[1 概述 1](#_Toc3235)

[2 规范性应用文件 1](#_Toc23261)

[3 术语和定义 1](#_Toc16522)

[3.1 鉴定 1](#_Toc30434)

[3.2 迁移 1](#_Toc21616)

[4 总体实施方案 2](#_Toc22005)

[5 鉴定工作 3](#_Toc28428)

[6 迁移工作 3](#_Toc17983)

[6.1 迁移工作要求 3](#_Toc9968)

[6.2 数据迁移范围 4](#_Toc10482)

[6.3 迁移历史数据类型 6](#_Toc29149)

[6.4 总体迁移方案 7](#_Toc14521)

[6.4.1 结构化数据迁移 8](#_Toc5116)

[6.4.2非结构化数据迁移 8](#_Toc29852)

2. 概述

本规范规定了雅砻江流域水电开发有限公司（以下简称“公司”）档案数据鉴定与迁移工作的实施步骤。

本规范适用于公司档案数据由旧档案管理系统向数字档案馆系统迁移工作及迁移前的鉴定工作。

1. 规范性应用文件

GB/T 18894-2016 电子文件归档与电子档案管理规范

档办发〔2015〕4号 企业电子文件归档和电子档案管理管理指南

ISO 13008:2012 信息与文献 数字档案转换和迁移过程

1. 术语和定义

## 鉴定

在档案数据迁移前，对电子档案进行销毁鉴定、真实性、完整性、安全性和可用性审查等工作。

## 迁移

在维护真实性、完整性和可用性的前提下，将电子档案从一个系统转移到另一个系统的过程。

1. 总体实施方案

公司电子档案数据鉴定与迁移工作应在确保电子档案的真实、可靠、完整和可用的基础上，实施电子档案及其元数据的转换或迁移。

按照确认鉴定和迁移需求、评估转换或迁移风险、制定转换或迁移方案、审批转换或迁移方案、转换或迁移测试、实施转换或迁移、评估转换或迁移结果、报告转换或迁移结果等步骤实施电子档案及元数据的鉴定和迁移工作。实施步骤如图1所示。



**图1 档案数据鉴定和迁移流程**

1. 鉴定工作

（1）数据迁移前，档案室应根据本单位档案保管期限表进行电子档案销毁鉴定，提出被鉴定对象的续存或销毁意见，对确无保存价值的档案登记造册，经公司法定代表人或主管领导批准后进行监销。

（2）对于需续存的档案数据，应根据《数字档案馆四性检测方案》对电子档案及其元数据的真实性、完整性、可用性和安全性进行检测，检测通过后进行数据迁移。**也可进行数据迁移后，在数字档案馆系统中进行四性检测，根据检测结果完善档案数据。**

1. 迁移工作

## 迁移工作要求

为保证历史数据迁移的准确性、完整性、可用性和真实性，迁移工作应满足以下要求：

1. 迁移工作开展前应确定数据迁移范围、迁移数量、迁移时间；
2. 迁移工作应采用较少的停机时间，较少的实施时间顺利完成迁移任务；
3. 历史数据应从Oracle数据库迁移到MySql数据库，在数字档案馆系统部署完成后开展数据迁移工作；
4. 数据迁移前，做好数据迁移准备工作，包括数据鉴定、清洗等；
5. 保证数据迁移后，在数字档案馆系统可以对历史档案统一管理和利用。

## 数据迁移范围

本次需完成原公司档案管理信息系统全部历史数据迁移，包括38个全宗，约20T，300万个电子文件纸质档案数字化副本和两河口水电站及杨房沟水电站业务系统产生的原生电子文件。数据迁移后，档案数据需按照全宗分类，档案报表模板由原来的一个模板对应多个全宗的形式改为多个模板对应多个全宗的形式，例如原档案管理系统中文书档案是采用1套模板共对应38个全宗，迁移后数字档案馆系统采用文书档案38套模板分别对应38个全宗。需迁移全宗范围如表1所示：

**表1 迁移全宗一览表**

|  |  |  |
| --- | --- | --- |
| **全宗号** | **全宗名称** | **备注** |
| 00 | 雅砻江流域公司（总部）档案 |  |
| 01 | 二滩水电工程和二滩水力发电总厂档案 |  |
| 02 | 桐子林水电工程和桐子林建设管理局档案 |  |
| 03 | 官地水电工程、官地建设管理局和官地水力发电厂档案 |  |
| 04 | 锦屏一级水电工程、锦屏建设管理局和锦屏水力发电总厂档案 |  |
| 05 | 锦屏二级水电工程档案 |  |
| 06 | 两河口水电工程、两河口建设管理局和两河口水力发电厂档案 |  |
| 07 | 杨房沟水电工程档案、杨房沟建设管理局和杨房沟水力发电厂 |  |
| 08 | 卡拉水电工程档案 |  |
| 09 | 牙根（二级）水电工程档案（原牙根电站） |  |
| 10 | 楞古水电工程 |  |
| 11 | 孟底沟水电工程档案 |  |
| 12 | 牙根（一级）水电工程档案 |  |
| 13 | 甲西水电工程档案 |  |
| 14 | 共科水电工程档案 |  |
| 15 | 新龙水电站工程档案、上游建设管理局和新龙水力发电总厂档案 |  |
| 16 | 乐安水电工程档案 |  |
| 17 | 林达水电工程档案 |  |
| 18 | 仁达水电工程档案 |  |
| 19 | 木罗水电工程档案 |  |
| 20 | 格尼水电工程档案 |  |
| 21 | 木能达水电工程档案 |  |
| 22 | 温波水电工程档案 |  |
| 23 | 鄂曲水电工程档案 |  |
| 24 | 卡拉、杨房沟水电站交通专用公路工程档案 |  |
| 25 | 两河口库区代建工程两河口段档案 |  |
| 26 | 河口库区代建工程新龙段档案 |  |
| 27 | 会理光伏项目工程档案 |  |
| 28 | 冕宁光伏项目工程档案 |  |
| 29 | 凉山有限公司档案 |  |
| 30 | 雅砻江四川能源有限公司档案 |  |
| 31 | 新能源管理局档案 |  |
| 32 | 盐源光伏项目档案 |  |
| 33 | 西昌风电项目档案 |  |
| 34 | 上游光伏项目档案 |  |
| 35 | 上游风电项目档案 |  |
| 36 | 德昌风电项目档案 |  |
| 37 | 锦屏地下实验室一、二期项目档案 |  |
| 38 | 锦屏大设施项目档案 |  |

## 迁移历史数据类型

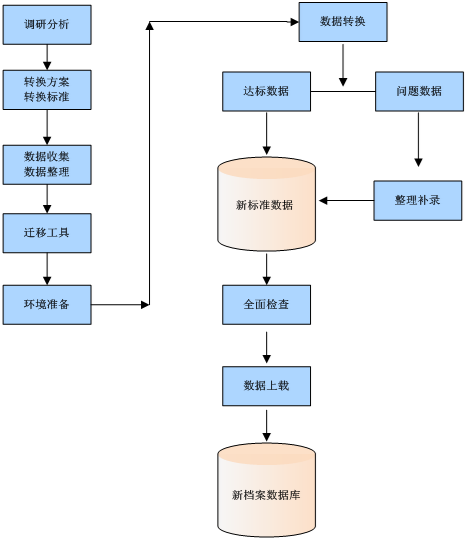
按照数据类型，需迁移的历史数据分为结构化数据和非结构化数据两类。

（1）结构化数据是指各全宗目录数据库信息、档案门类信息、档号规则、元数据、各类报表信息等；

（2）非结构化数据是指各全宗存储的电子档案，格式包括doc、pdf、xls、jpg等，目前电子档案共分为8个磁盘存储。

## 总体迁移方案

数字档案馆系统上线试运行前需要完成原档案管理系统中历史数据的迁移，导出条目数据和附件数据，在迁移实施期间，暂停原电子档案系统的手工采集、整理、归档业务，待所有数据迁移完成后，使用数字档案馆系统完成业务。数据迁移后，原档案管理系统并行一段时间，只可以查询档案不再接收新增和调整数据。具体实施步骤如图2所示。



**图2 数据迁移流程图**

## 6.4.1 结构化数据迁移

（1）准备数据验证工具

为保证数据的安全、准确导入，通过对比RazorSQL、Navicat、 DamengDTS 三种迁移工具，分别对三种迁移工具进行测试，根据测试数据分析迁移后数据的准确性、完整性及转换效率，得出DamengDTS迁移工具的性能最优，因此，采用该工具进行数据迁移。

1. 结构化数据迁移步骤

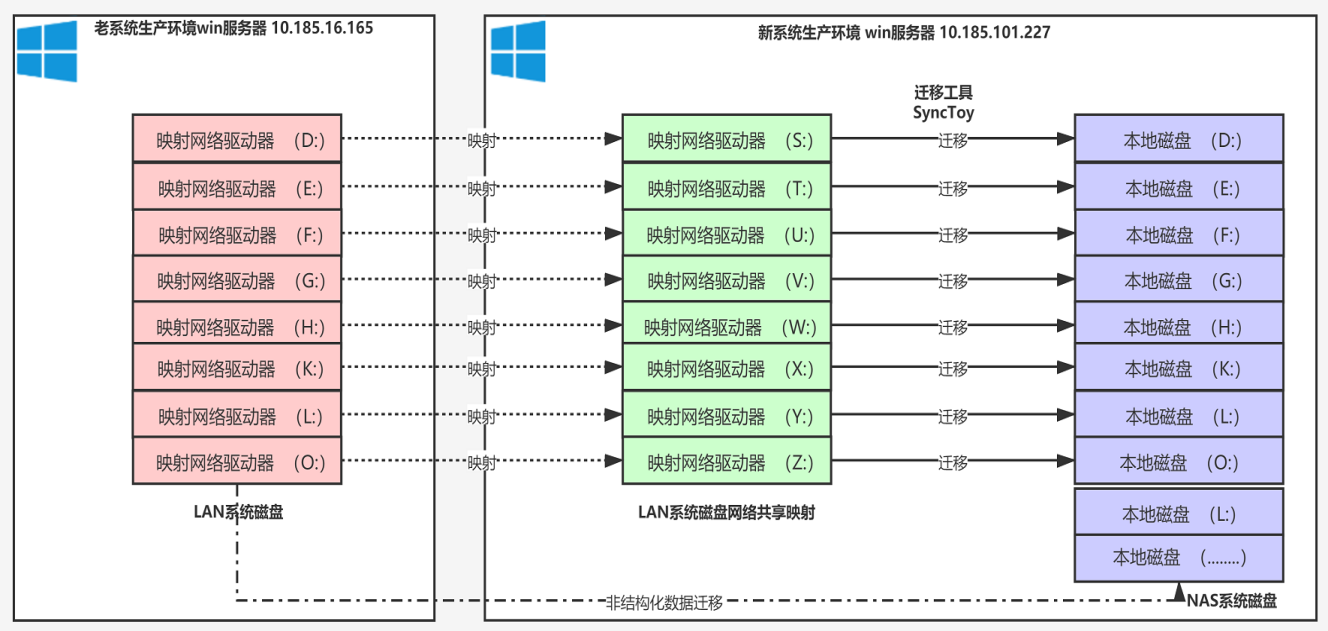
首先，将原档案管理系统的Oracle数据库结构转换为Dameng数据库结构；其次，将Dameng数据库结构转换为Mysql数据库结构；最后，迁移完成后进行数据结构验证确认，并且对数据库性能进行压测。

1. 迁移后数据检测

数据迁移到数字档案馆系统后，需检测各项配置是否正确、目录树、条目数据与原档案管理系统是否一致等。根据检测反馈的情况对其进行调整、数据清洗或再迁移。由甲方用户完成迁移结果的确认。

## 6.4.2非结构化数据迁移

1. 准备迁移环境

非结构数据具有如下特点：存储量大（20T）、文件数量多（300万个）、仍然在持续增长。针对这些特点，首先，要准备好新服务器的网络存储设备，总容量需要50T。然后，将原档案管理系统的8个数据存储磁盘以“网络磁盘映射”的方式，映射在新服务器上，如下图8个红色磁盘映射成8个绿色磁盘。

**图3 非结构化数据迁移流程图**

1. 非结构化数据迁移步骤

磁盘映射关系后访问绿色磁盘会自动链接到红色磁盘。使用数据同步工具“SyncToy”将数据同步到新磁盘（蓝色部分）。工具有增量同步和错误日志记录功能，可以借助此功能进行监控和人工copy错误文件。按照每秒30M的速度，公式：20T/（24天\*60分\*60秒\*30M）=8天，完成所有电子文件的同步。

1. 迁移后数据检测

数据迁移到新系统后，针对非结构化数据我们可以查看磁盘总量进行汇总，借助文件查询工具“Everything”进行抽查，确保新老系统电子文件总量一致、数量一致。