**雅砻江流域水电开发有限公司**

**电子文件封装技术规范**

紫光软件系统有限公司

2021年6月

1. 范围

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”）电子文件归档、长期保存过程中所需数据文件信息组织方式、存储结构及封装的规则和方法等。

本规范适用于电子文件归档、交换与长期保存，保证电子数据脱离系统的可读性。电子档案备份时数据组织与本规范长期保存信息包要求一致。

本规范与《电子文件元数据规范》（待补充）紧密联系，相辅相成，需一起使用。

1. 规范性引用文件

DA/T 48-2009 基于XML的电子文件封装规范

档发〔2012〕7号 电子档案移交与接收办法

档办发〔2015〕4号 企业电子文件归档和电子档案管理指南

YLJ-IM-17文书档案归档文件整理实施细则

YLJ-IM-20会计档案管理办法

YLJ-IM-16档案分类编号规则

雅砻江流域水电开发有限公司电子文件通用元数据规范

各业务系统元数据规范（名称待填）

1. 术语和定义

DA/T 48-2009、档办发〔2015〕4号中界定的下列术语和定义适用本规范。为方便适用，下列重复列出了某些术语和定义。

* 1. **封装**

将电子文件及其元数据按制定结构打包的过程。

* 1. **信息包**

包含电子档案元数据、内容数据、说明文件等，按照一定的结构组织，可用于不同环节之间信息传递的信息集合对象。

* 1. **封装格式**

电子文件归档、移交、长期保存过程中对数据文件信息的组织形式，本规范规定了以树型结构组织数据信息。

* 1. **归档信息包**

电子文件归档时立档单位内部业务部门向档案部门提交的信息包，简称ASIP。

* 1. **长期保存信息包**

按照长期保存要求形成的电子档案信息包，简称AIP。

1. 总则

根据公司电子文件管理阶段和管理对象的不同，为保证电子档案脱离系统的可读性，将电子文件与其元数据按照一定结构封装在一个数据包中，以维护电子文件与其元数据的完整性，并保障两者之间的可靠联系。

数据包采用压缩包形式，将压缩包内多个文档打成压缩包形式，其中元数据采用XML格式存储。信息存储结构按照归档环节、长期保存环节分别设定。

1. 封装格式说明
   1. 类型

公司电子文件封装工作在归档环节和长期保存环节开展，分别形成归档信息包和长期保存信息包。

* 1. 归档信息包格式说明

归档信息包按照归档方式，采用“一件一包”或“一卷一包”的信息包格式。具体封装结构如图1、图2所示，采用树型结构封装，根目录下可再设文件夹，分别管理不同件的电子文件。



合成文件.PDF



电子档案内容数据1.PDF



电子档案内容数据2.JPG



电子档案内容数据3.DOC



电子档案内容数据4.PDF



单件归档信息包



件元数据信息.XML

说明文件.txt

**图1 “一件一包”归档信息包封装格式**



合并文件.PDF



电子档案1.PDF



单卷归档信息包



案卷元数据信息.XML

说明文件TXT



卷内文件1

卷内文件2

卷内文件N

**……**



电子档案2..doc



电子档案3.xls



电子档案N.JPG

**……**



卷内文件目录.XML



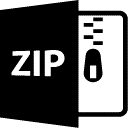
元数据.XML

**图2 “一卷一包”归档信息包封装格式**

* 1. 长期保存信息包格式说明

长期保存信息包采用树型结构封装，根据档案门类，按照档案分类建立树型结构。

（1）文书档案按照“全宗-年度-问题-保管期限”分别设立文件夹，保管不同年度、问题和保管期限的电子文件。一般一个年度为一个长期保存包，电子文件数量较少可多个年度为一个长期保存包，长期保存包封装格式见图3。



说明文件.TXT



目录文件.XML

其他文件夹



电子档案



年度1文件夹

年度2文件夹

年度N文件夹



问题1文件夹

问题2文件夹

问题N文件夹



保管期限1

保管期限2

保管期限3



归档文件文件夹1夹1

归档文件文件夹2

归档文件文件夹N



元数据信息.eep

合并版式文件.PDF



发文稿纸、办理单



正本



定稿



附件



草稿



目录文件.XLS

**图3文书档案长期保存信息包封装格式**

1. 科技档案按照“全宗-类目（6-9大类）-年度/项目-下一级类目”设立文件夹，一个年度或一个项目设立一个文件夹，按照类目划分逐级建立文件夹。
2. 声像档案按照“全宗-文件类型（照片、音频、视频）-年度-分类号”设立文件夹，按照照片、音频、视频分别设立文件夹，文件夹下按照年度、分类号逐级建立文件夹。
3. 会计档案按照“全宗-年度-会计档案类型”逐级设立文件夹。
4. 封装格式组成要素描述

6.1归档信息包组成要素描述

6.1.1归档信息包概述

归档信息包包括说明文件、元数据和数据内容三部分，长期保存信息包还包括目录文件。

6.1.2“一件一包”形式归档信息包组成要素描述

“一件一包”归档信息包应存放说明文件、元数据文件和电子档案三部分。

1. 说明文件存放与归档信息包有关的信息，包括归档信息包大小、生成归档包的软硬件环境描述及其他有助于说明归档信息包的内容。说明文件命名为“说明文件.TXT”。
2. 元数据文件存放该件的所有元数据信息，包括文件实体元数据、机构人员实体元数据、业务实体元数据和实体关系元数据四个域。元数据范围根据各业务系统建设情况和档案管理要求编制，具体见各类业务系统归档接口方案要求。元数据采用XML存储，文件名称命名为“元数据文件.XML”。
3. 电子档案文件夹存放电子档案内容数据，包括合并后的版式文件及各组件文件。以文书档案为例，电子档案文件夹包括版式文件、文件处理单、正本、定稿及附件等。电子档案命名和排序。合同结算类、会计凭证类归档信息包可按照《业务系统归档接口方案》要求执行。其他门类电子档案文件夹内电子文件内容、命名方式及排序应根据档案管理要求制定。

6.1.3“一卷一包”形式归档信息包组成要素描述

“一卷一包”归档信息包应存放说明文件、卷内目录文件、元数据文件和电子档案四部分。

（1）说明文件存放与归档信息包有关的信息，包括归档信息包大小、卷内文件数量、生成归档包的软硬件环境描述及其他有助于说明归档信息包的内容。说明文件命名为“说明文件.TXT”。

（2）案卷信息元数据文件存放该卷的案卷级所有元数据信息，包括文件实体元数据、机构人员实体元数据、业务实体元数据和实体关系元数据四个域。元数据范围根据各业务系统建设情况和档案管理要求编制，具体见各类业务系统归档接口方案要求。元数据采用XML存储，文件命名为“元数据文件.XML”。

（3）卷内目录存放该案卷内卷内文件信息，包括序号、文件编号、责任者、文件题名、备注等信息。卷内目录命名为“卷内目录.XML”。

（4）电子档案文件夹存放该卷电子档案内容数据，包括卷内每件电子文件。以会计凭证档案为例，电子档案文件夹包括该卷内每件会计凭证电子档案。会计凭证类电子档案命名和排序按照《业务系统归档接口方案》要求执行。其他门类电子档案文件夹内电子文件内容、命名方式及排序应根据档案管理要求制定。

6.2长期保存信息包组成要素描述

6.2.1长期保存信息包概述

长期保存信息包包括说明文件、目录文件、其他文件级和数据内容四部分。

6.2.2组成要素描述

（1）说明文件存放与长期信息包有关的信息，包括长期信息包大小、电子文件数量、存储载体参数、编号、读取本载体内档案所需要的软硬件环境及其他有助于说明信息包、载体的信息，说明文件命名为“说明文件.TXT”。

（2）目录文件存放有关电子档案的目录信息，目录文件与每卷、每件电子档案相对应。根据电子档案的整理方式进行文件级或案卷级描述。以“件”为单位整理的档案目录命名为“文件目录.XML”；以“案卷”为单位整理的档案目录命名为“案卷目录.XML”、“文件目录.XML”。根据实际利用需求，可增加XLS格式目录。案卷目录信息包括案卷档号、案卷题名、件数、保管期限、总页数、立卷单位等；文件目录包括序号、档号、文件编号、责任者、题名、页数/页号、备注等。

（3）其他文件夹存放各种其他附属的文件，包括所采用的元数据规范、数据封装规范、分类编号规则、文件命名规则、XML模式及交接信息等。这些文件采用TXT、XML或其他符合长期保存要求的格式。

（4）电子档案文件夹用于存放电子档案内容数据及其元数据。应根据电子档案的整理和分类方法以及实际情况，逐级建立文件夹，命名为“电子档案”，可对存储结构中的类别、案卷、文件等层级进行取舍。

1. 各组成要素XML描述

7.1元数据XML描述

7.2目录文件XML描述

以文书类电子档案长期保存信息包目录文件为例，XML描述如下。

<?xml version="1.0" encoding="UTF-8"?>

<文件目录>

<文件>

<年度></年度>

<顺序号></顺序号>

<档号></档号>

<责任者> </责任者>

<题名></题名>

<日期></日期>

<密级></密级>

<备注/>

<文件>

……

<文件>

……

</文件>

</文件目录>

7.2元数据XML描述见附件2。

元数据XML封装见《业务系统归档接口方案》。