

SZSD

数 字 山 东 工 程 标 准

SZSD 0046.2—2024

电子档案 第2部分：数据规范

Electronic records—Part 2: Data specification

2024-04-15 发布

2024-06-01 实施

山东省大数据局 发 布

目 次

前言..... II

引言..... III

1 范围..... 1

2 规范性引用文件..... 1

3 术语和定义..... 1

4 缩略语..... 1

5 电子档案元数据..... 2

 5.1 元数据基本集..... 2

 5.2 元数据扩展..... 4

6 电子公文归档数据组织..... 4

 6.1 归档信息包..... 4

 6.2 说明文件..... 5

 6.3 电子档案文件夹..... 5

7 电子公文归档格式要求..... 5

附录 A（资料性） 说明文件示例..... 7

附录 B（资料性） 电子档案元数据示例..... 8

参考文献..... 10

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

本文件是SZSD 0046《电子档案》的第2部分。SZSD 0046已经发布了以下部分：

- 第1部分：接口规范；
- 第2部分：数据规范。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本标准由山东省大数据局提出并归口。

本标准由山东省大数据局、山东省档案馆组织实施。

本标准起草单位：山东省大数据局、山东省档案馆。

本标准主要起草人：李世华、王宪东、赵琳、刘虎、储牧原、马学刚、武伟、李文姣、张学干、郑华鑫

引 言

为贯彻落实国家机关电子档案和电子公文管理的政策文件要求，提升电子档案的法律效力，发挥电子档案在政府流程再造的关键作用，加快建设“数字山东”和“互联网+政务服务”，对电子档案相关工作提出要求。

针对当前电子档案中接口和数据管理等工作存在的普遍问题，制定SZSD 0046《电子档案》系列标准。SZSD 0046拟由以下两个部分构成。

——第1部分：接口规范。目的在于规范山东省省直单位的电子公文归档系统及档案管理系统的接口的相关要求。

——第2部分：数据规范。目的在于规范山东省省直单位电子档案的元数据、电子公文归档的数据组织和归档格式等要求。

电子档案 第2部分：数据规范

1 范围

本文件规定了山东省省直单位电子档案的元数据、电子公文归档的数据组织和归档格式等要求。
本文件适用于山东省省直单位电子公文归档系统及档案管理系统产生的电子公文的归档和处理。
注：市级单位电子档案的数据规范可参照本标准。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

- GB/T 18894—2016 电子文件归档与电子档案管理规范
- GB/T 33480—2016 党政电子公文元数据规范
- GB/T 39362—2020 党政机关电子公文归档规范
- DA/T 22—2015 归档文件整理规则
- DA/T 92—2022 电子档案单套管理一般要求

国家档案局. 关于全面推行机关档案分类方案、文件材料归档范围和档案保管期限表三合一制度的通知：档函〔2022〕58号.（2022-5-10）[2023-2-24].

3 术语和定义

GB/T 39362—2020和DA/T 58—2014界定的以及下列术语和定义适用于本文件。

3.1

简单型 simple type

不具有子元素的元素所对应的元素类型。

[来源：DA/T 46—2009, 3.7]

3.2

容器型 container type

具有子元素且本身不能被赋值的元素所对应的元素类型。

[来源：DA/T 46—2009, 3.7]

3.3

复合型 complex type

本身可以被赋值且在一定条件下可以具有子元素的元素所对应的元素类型。

[来源：DA/T 46—2009, 3.7]

4 缩略语

下列缩略语适用于本文件。

OCR：光学字符识别（Optical Character Recognition）

OFD：开放式版式文档（Open Fixed layout Document）
XML：可扩展标记语言（Extensible Markup Language）

5 电子档案元数据

5.1 元数据基本集

电子档案元数据基本集中的元素及其结构描述见表1。

表 1 电子档案元数据基本集

编号	元数据元素名称	英文名称	约束性	可重复性	元素类型	数据类型	备注
M1	公文标识	document identifier	必选	不可重复	简单型	字符型	GB/T 33480—2016中“公文标识”
M2	文种	document type	可选	不可重复	简单型	字符型	GB/T 33480—2016中“文种”
M3	份号	serial number of copies	可选	不可重复	简单型	字符型	GB/T 33480—2016中“份号”
M4	密级和保密期限	security classification and duration	必选	不可重复	简单型	字符型	DB37/T 536—2019中“密级”和“保密期限”； GB/T 33480—2016中“密级和保密期限”
M5	紧急程度	emergency degree	可选	不可重复	简单型	字符型	GB/T 33480—2016中“紧急程度”
M6	发文机关标志	identification of document issuer	可选	不可重复	简单型	字符型	GB/T 33480—2016中“发文机关标志”
M7	发文字号	issued number of document	可选	不可重复	简单型	字符型	DB37/T 536—2019中“文件编号”； GB/T 33480—2016中“发文字号”
M8	签发人	signer	可选	不可重复	简单型	字符型	GB/T 33480—2016中“签发人”
M9	标题	document title	必选	不可重复	简单型	字符型	DB37/T 536—2019中“题名”； GB/T 33480—2016中“标题”
M10	主送机关	department	可选	不可重复	简单型	字符型	GB/T 33480—2016中“主送机关”
M11	附件说明	attachment	可选	不可重复	简单型	字符型	DB37/T 536—2019中“附件”； GB/T 33480—2016中“附件说明”

表1 电子档案元数据基本集（续）

编号	元数据元素名称	英文名称	约束性	可重复性	元素类型	数据类型	备注
M12	发文机关或签发 人署名	signature of document issuing agency	可选	不可重复	简单型	字符型	DB37/T 536—2019中 “责任者”； GB/T 33480—2016中 “发文机关或签发人署 名”
M13	成文日期	issued date of document	必选	不可重复	简单型	日期型	DB37/T 536—2019中 “时间”； GB/T 33480—2016中 “成文日期”
M14	附注	annotation	可选	可重复	简单型	字符型	DB37/T 536—2019中 “附注”； GB/T 33480—2016中 “附注”
M15	抄送机关	copy of department	可选	不可重复	简单型	字符型	GB/T 33480—2016中 “抄送机关”
M16	印发机关	printing and sending department	可选	不可重复	简单型	字符型	GB/T 33480—2016中 “印发机关”
M17	印发日期	printing date	可选	不可重复	简单型	日期型	DB37/T 536—2019中 “时间”； GB/T 33480—2016中 “印发日期”
M18	发布层次	release level	可选	不可重复	简单型	字符型	GB/T 33480—2016中 “发布层次”
M19	主题词	descriptor	可选	可重复	简单型	字符型	DA/T 46—2006中 “主题词”； DB37/T 536—2019中 “主题词”
M20	关键词	keyword	可选	不可重复	简单型	字符型	DA/T 46—2006中 “关键词”； DB37/T 536—2019中 “关键词”
M21	档号	archival code	必选	不可重复	复合型	字符型	DB37/T 536—2019中 “档号”
M22	全宗号	fonds identifier	必选	不可重复	简单型	字符型	DB37/T 536—2019中 “全宗号”
M23	年度	year	必选	不可重复	简单型	数值型	DB37/T 536—2019中 “年度”
M24	保管期限	period	必选	不可重复	简单型	字符型	DB37/T 536—2019中 “保管期限”

表1 电子档案元数据基本集（续）

编号	元数据元素名称	英文名称	约束性	可重复性	元素类型	数据类型	备注
M25	机构或问题	organizational structure or function	必选	不可重复	简单型	字符型	DB37/T 536—2019中 “机构或问题”
M26	件号	item number	必选	不可重复	简单型	字符型	DB37/T 536—2019中 “文件号”
M27	控制标识	control identifier	必选	不可重复	简单型	字符型	DB37/T 536—2019中 “控制标识”
M28	页数	total number of pages	必选	不可重复	简单型	数值型	DB37/T 536—2019中 “页数”
M29	页号	page number	可选	不可重复	简单型	数值型	DA/T 46—2006中 “页号”； DB37/T 536—2019中 “页（张）号”
M30	数字对象标识	electronic file identifier	必选	不可重复	简单型	字符型	—
M31	环境信息	environment information	可选	不可重复	容器型	—	—
M32	软件环境	software environment	可选	不可重复	简单型	字符型	—
M33	硬件环境	hardware environment	可选	不可重复	简单型	字符型	—
M34	处理类型	disposal type	可选	不可重复	简单型	字符型	—
M35	处理者	disposal actor	可选	不可重复	简单型	字符型	—
M36	处理部门	disposal department	可选	不可重复	简单型	字符型	—
M37	处理时间	disposal time	可选	不可重复	简单型	日期型	—
M38	处理结果	disposal result	可选	不可重复	简单型	字符型	—

5.2 元数据扩展

实际工作中可根据具体情况扩展元数据集合，元数据扩展应符合GB/T 33480—2016中的相关要求。

6 电子公文归档数据组织

6.1 归档信息包

电子公文及其元数据应按照GB/T 39362—2020中的要求形成归档信息包。归档信息包中一般包含同一批次归档的多份归档电子公文，电子公文应以DA/T 22—2015规定的分类方式为依据进行组织。以采用年度—保管期限—机构（问题）三级分类方式为例，归档信息包结构示例见图1。

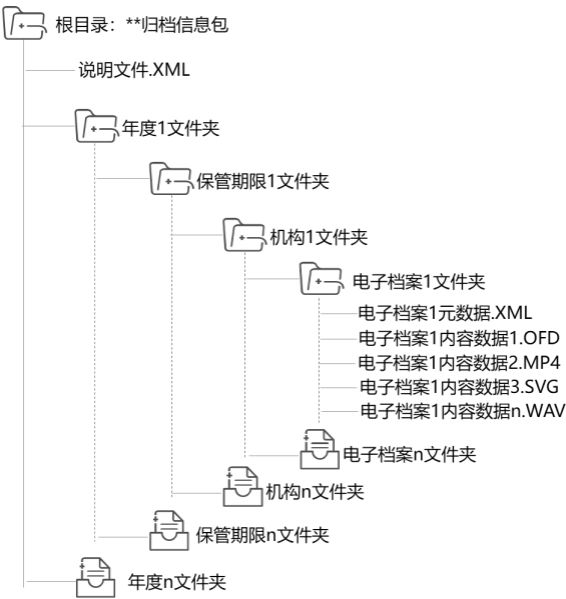


图 1 归档信息包结构

6.2 说明文件

说明文件以XML文件方式存放信息包的元数据信息，包括信息包标识、信息包类型、创建者、创建时间、创建环境、存储位置、信息包说明等，示例见附录A。

6.3 电子档案文件夹

电子档案文件夹一般以档号命名，档号应符合《关于全面推行机关档案分类方案、文件材料归档范围和档案保管期限表三合一制度的通知》（档函〔2022〕58号）的编号规范，例如“A007-WS•2011-Y-BGS-0001”。文件夹中存放电子档案内容数据及其元数据，示例见附录B，具体要求如下：

- a) 电子档案内容数据按照“档号”+“-”+“xx”（xx指件内顺序号）的方式命名，例如“A007-WS•2011-Y-BGS-0001-01.OFD”，包括OFD格式的内容数据以及其他附属文件：
 - 1) OFD文件：内容数据的主体部分，包括文件处理单和公文主体（包含正本，定稿及历次修改稿），多个OFD文件按正本、文件处理单、定稿、历次修改稿的顺序排列；
 - 2) 其他附属文件：部分无法转换成OFD格式的文件（比如图形文件、音视频文件、其他特殊格式文件等）可根据第7章要求按照一定顺序置于电子档案文件夹中。
- b) 电子档案元数据信息应以XML文件方式存放，存在多个OFD文件时，元数据文件独立存在并按照一定顺序置于电子档案文件夹中。

7 电子公文归档格式要求

根据构成文件内容的数字对象类型的不同可将电子公文分成文件处理单、公文主体，其他附属文件三个部分，双层OFD文件由扫描形成的电子公文内容图像，通过OCR技术提取全文内容数据和扫描图像合并生成。详细格式要求见表2。其他电子公文格式要求应符合GB/T 18894—2016中8.3和DA/T 92—2022中8.1的规定。

表 2 电子公文归档格式要求

文件构成	文件形态	归档格式
文件处理单	纸质(扫描)	双层 OFD
	电子(网页, 文本)	OFD
公文主体(含正本、定稿、历次修改稿等多个版本)	纸质(扫描)	双层 OFD
	电子(文本)	OFD
其他附属文件	电子	文本类:OFD
		图像类:JPG、TIF
		图形类: SVG、STEP
其他附属文件	电子	音频类:WAV、MP3
		视频类:AVI、MP4、MPG
		社交媒体类: HTML、MHT

附 录 A
(资料性)
说明文件示例

说明文示例为：

```
<xs:element name="信息包元数据">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="信息包标识" type="xs:string"/>
      <xs:element name="信息包类型" type="xs:string"/>
      <xs:element name="创建者" type="xs:string"/>
      <xs:element name="创建时间" type="xs:date"/>
      <xs:element name="创建环境" type="xs:date"/>
      <xs:element name="存储位置" type="xs:string"/>
      <xs:element name="信息包说明" type="xs:string"/>
    </xs:sequence>
  </xs:complexType>
```

附 录 B
(资料性)
电子档案元数据示例

电子档案文件夹中的元数据示例为：

```
<xs:element name="电子档案元数据">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="公文标识" type="xs:string"/>
      <xs:element name="文种" type="xs:string"/>
      <xs:element name="份号" type="xs:string"/>
      <xs:element name="密级和保密期限" type="xs:string"/>
      <xs:element name="紧急程度" type="xs:string"/>
      <xs:element name="发文机关标志" type="xs:string"/>
      <xs:element name="发文字号" type="xs:string"/>
      <xs:element name="签发人" type="xs:string"/>
      <xs:element name="标题" type="xs:string"/>
      <xs:element name="主送机关" type="xs:string"/>
      <xs:element name="附件说明" type="xs:string"/>
      <xs:element name="发文机关或签发人署名" type="xs:string"/>
      <xs:element name="成文日期" type="xs:date"/>
      <xs:element name="附注" type="xs:string"/>
      <xs:element name="抄送机关" type="xs:string"/>
      <xs:element name="印发机关" type="xs:string"/>
      <xs:element name="印发日期" type="xs:date"/>
      <xs:element name="发布层次" type="xs:string"/>
      <xs:element name="主题词" type="xs:string"/>
      <xs:element name="关键词" type="xs:string"/>
      <xs:element name="档号" type="xs:string"/>
      <xs:complexType>
        <xs:sequence>
          <xs:element name="全宗号" type="xs:integer"/>
          <xs:element name="年度" type="xs:integer"/>
          <xs:element name="保管期限" type="xs:string"/>
          <xs:element name="机构或问题" type="xs:string"/>
          <xs:element name="件号" type="xs:string"/>
        </xs:sequence>
      </xs:complexType>
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="控制标识" type="xs:string"/>
<xs:element name="页数" type="xs:integer"/>
<xs:element name="页号" type="xs:integer"/>
```

```

<xs:element name="数字对象标识" type="xs:string"/>
<xs:element name="环境信息">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="软件环境" type="xs:string"/>
      <xs:element name="硬件环境" type="xs:string"/>
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="处理类型" type="xs:string" maxOccurs="unbounded"/>
<xs:element name="处理者" type="xs:string" maxOccurs="unbounded"/>
<xs:element name="处理部门" type="xs:string" maxOccurs="unbounded"/>
<xs:element name="处理时间" type="xs:date" maxOccurs="unbounded"/>
<xs:element name="处理结果" type="xs:string" maxOccurs="unbounded"/>
</xs:sequence>
</xs:complexType>
</xs:element>

```

参 考 文 献

- [1] GB/T 33476.1—2016 党政机关电子公文格式规范 第1部分：公文结构
 - [2] DA/T 58—2014 电子档案管理基本术语
 - [3] DA/T 70—2018 文书类电子档案检测一般要求
 - [4] DB37/T 536—2019 文书档案目录数据采集规范
-