

档案工作标准汇编

全国档案工作标准化技术委员会 编

档案工作标准汇编（9）

全国档案工作标准化技术委员会 编

中国档案出版社
2010

责任编辑/田小燕

装帧设计/汪俊宇

图书在版编目 (CIP) 数据

档案工作标准汇编. 9 / 全国档案工作标准化技术委员会编.
— 北京: 中国档案出版社, 2010. 4
ISBN 978-7-5105-0155-5

I. ①档… II. ①全… III. ①档案工作—标准—汇编—中国
IV. ①G27-65

中国版本图书馆 CIP 数据核字 (2010) 第 040993 号

出版/中国档案出版社 (北京市宣武区永安路 106 号 邮编 100050)

发行/中国档案出版社发行部

印刷/北京市鑫山源印刷有限公司

规格/880×1230 1/16 印张/10.5 字数/310 千字

版次/2010 年 6 月第 1 版 2010 年 6 月第 1 次印刷

印数/2500 册

定价/58.00 元

版权专有 侵权必究

出 版 说 明

为了加强对档案工作标准的宣传,方便各级档案部门和档案工作者在工作中贯彻实施标准,同时保持《档案工作标准汇编》的连续性,我们将近期发布的档案工作行业标准汇编成《档案工作标准汇编(9)》。

本汇编收录了国家档案局档发〔2009〕7号文件批准发布的3项行业标准,包括《文书类电子文件元数据方案》(DA/T 46—2009)、《版式电子文件长期保存格式需求》(DA/T 47—2009)、《基于XML的电子文件封装规范》(DA/T 48—2009)。

本书适合各级档案部门工作人员查考使用。

全国档案工作标准化技术委员会秘书处

2010年5月

目 录

DA/T 46—2009	文书类电子文件元数据方案	1
DA/T 47—2009	版式电子文件长期保存格式需求	67
DA/T 48—2009	基于 XML 的电子文件封装规范	75



中华人民共和国档案行业标准

DA/T 46—2009

文书类电子文件元数据方案

Metadata Standard for Administrative Electronic Records

2009-12-16 发布

2010-06-01 实施

国 家 档 案 局 发 布

前 言

本标准由安徽省档案局(馆)提出。

本标准由国家档案局归口。

本标准起草单位：安徽省档案局(馆)、北京北大方正技术研究院有限公司。

本标准主要起草人：李学香、黄玉明、周建武、胡华平、王学武、邓晓文、段丽琼、吴彬松、华婷、陈姗姗、朱霖露。

引 言

元数据是描述文件的背景、内容、结构及其管理过程的数据。元数据管理有助于保证电子文件的真实性、完整性、可靠性、可用性，保障电子文件的证据特性，便于对电子文件的理解、管理、交换和利用，是电子文件管理中必不可少的一部分。

随着电子文件的大量生成，有必要为创建、管理、应用元数据建立统一的方案，以系统有效地记录电子文件的内容特征、形式特征、背景和管理过程信息，为电子政务、办公自动化、文件及档案管理系统的设计提供依据，并全面实现元数据的多种功能和用途。本标准参照 ISO 15489《信息与文献 文件管理》、ISO 23081《信息与文献 文件管理过程 文件元数据》、档案行业标准 DA/T 18—1999《档案著录规则》、国际档案理事会标准《国际标准 档案著录规则(总则)》，并结合我国电子文件管理实际，规定了文书类电子文件元数据集以及元数据元素关系间的逻辑架构。

文书类电子文件元数据方案

1 范围

本标准规定了以文件和案卷形式的档案为对象的文书类电子文件形成、交换、归档、移交、保管、利用等全过程元数据设计、捕获、著录的一般要求。

本标准适用于各级各类档案馆、机关、团体、企业事业单位和其他社会组织对计算机系统中直接生成的文书类电子文件(含纸质文件/档案数字化图像、文本)的管理。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件,仅注日期的版本适用于本文件。凡是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 2312—1980 信息交换用汉字编码字符集 基本集

GB/T 4880.2—2000 语种代码名称 第2部分:3字母代码(ISO 639-2:1998,EQV)

GB/T 4881 中国语种代码

GB/T 7156—2003 文献保密等级代码与标识

GB/T 7408—2005 数据元和交换格式 信息交换 日期和时间表示法(ISO 8601:2000,IDT)

GB 11714—1997 全国组织机构代码编制规则

GB/T 15418 档案分类标引规则

GB 18030—2005 信息技术 中文编码字符集

GB/T 18894—2002 电子文件归档与管理规范

DA/T 1—2000 档案工作基本术语

DA/T 13—1994 档号编制规则

DA/T 14—1994 全宗指南编制规范

DA/T 18—1999 档案著录规则

DA/T 22—2000 归档文件整理规则

DA/T 31—2005 纸质档案数字化技术规范

ISO 11179-3:2003 信息技术 元数据注册系统(MDR) 第3部分:注册系统元模型和基本属性(Information technology — Metadata registries (MDR) — Part 3: Registry metamodel and basic attributes)

ISO 15489-1:2001 信息与文献 文件管理 第1部分:通则(Information and documentation — Records management — Part 1: General)

ISO 23081-1:2006 信息与文献 文件管理过程 文件元数据 第1部分:原则(Information and documentation — Records management processes — Metadata for records — Part 1: Principles)

ISO/TS 23081-2:2007 信息与文献 文件管理过程 文件元数据 第2部分:概念与实施(Information and documentation — Records management processes — Metadata for records — Part 2: Conceptual and implementation issues)

3 术语和定义

DA/T 1—2000 界定的以及下列术语和定义适用于本文件。

3.1

电子文件 electronic records

指在数字设备及环境中生成,以数码形式存储于磁带、磁盘、光盘等载体,依赖计算机等数字设备阅读、处理,并可在通信网络上传送的文件。

[GB/T 18894—2002,定义 3.1]

3.2

文书类电子文件 administrative electronic records

反映党务、政务、生产经营管理等各项管理活动的电子文件。

3.3

组合文件 combined records

由若干个联系紧密的文档组成、视为一个文件著录的文档集合体。

3.4

文档 document

归档文件最小的文件单元,可以是一个独立的自然件,也可以是组合文件中的某一自然件或某一自然件的附件。

3.5

元数据 metadata

描述文件的背景、内容、结构及其整个管理过程的数据。

[ISO 15489 - 1:2001,定义 3.12]

3.6

元素 element

通过标识、定义、约束性、值域等一组属性描述的数据单元。

3.7

简单型 simple type

不具有子元素的元素所对应的元素类型。

3.8

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。

3.9

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。

3.10

捕获 capture

适时获取电子文件及其元数据的方法和过程。

3.11

实体 entity

任何已经存在的、将要存在的或可能存在的具体的或抽象的事物,包括事物间的联系。

[ISO 23081 - 2:2007,定义 3.6]

4 元数据元素、结构及描述方法

4.1 元数据元素及其结构

本标准参考 ISO 23081,结合目前我国电子文件管理实际,将文书类电子文件元数据从概念层次上区分为文件实体元数据、机构人员实体元数据、业务实体元数据、实体关系元数据四个域,每个域包

含的元数据元素及其结构规定如下：

表 1 文件实体元数据

编号	元数据	编号	元数据
M1	聚合层次		
M2	来源	M3	档案馆名称
		M4	档案馆代码
		M5	全宗名称
		M6	立档单位名称
M7	电子文件号		
M8	档号	M9	全宗号
		M10	目录号
		M11	年度
		M12	保管期限
		M13	机构或问题
		M14	类别号
		M15	室编案卷号
		M16	馆编案卷号
		M17	室编件号
		M18	馆编件号
		M19	文档序号
		M20	页号
M21	内容描述	M22	题名
		M23	并列题名
		M24	副题名
		M25	说明题名文字
		M26	主题词
		M27	关键词
		M28	人名
		M29	摘要
		M30	分类号
		M31	文件编号
		M32	责任者
		M33	日期
		M34	文种
		M35	紧急程度
		M36	主送
		M37	抄送
		M38	密级
		M39	保密期限

表 1 文件实体元数据(续表)

编号	元数据	编号	元数据
M40	形式特征	M41	文件组合类型
		M42	件数
		M43	页数
		M44	语种
		M45	稿本
M46	电子属性	M47	格式信息
		M48	计算机文件名
		M49	计算机文件大小
		M50	文档创建程序
		M51	信息系统描述
M52	数字化属性	M53	数字化对象形态
		M54	扫描分辨率
		M55	扫描色彩模式
M57	电子签名	M56	图像压缩方案
		M58	签名规则
		M59	签名时间
		M60	签名人
		M61	签名结果
		M62	证书
		M63	证书引证
M65	存储位置	M64	签名算法标识
		M66	当前位置
		M67	脱机载体编号
		M68	脱机载体存址
M70	权限管理	M69	缩微号
		M71	知识产权说明
		M72	授权对象
		M73	授权行为
M75	附注	M74	控制标识

表 2 机构人员实体元数据

编号	元数据	编号	元数据
M76	机构人员类型		
M77	机构人员名称		
M78	组织机构代码		
M79	个人职位		

表 3 业务实体元数据

编号	元数据	编号	元数据
M80	业务状态		
M81	业务行为		
M82	行为时间		
M83	行为依据		
M84	行为描述		

表 4 实体关系元数据

编号	元数据	编号	元数据
M85	实体标识符		
M86	关系类型		
M87	关系		
M88	关系描述		

4.2 元数据的描述方法

本标准参考 ISO 11179.3:2003,采用表 5 所示方法对元数据元素进行描述。

本标准所描述的元数据元素有四个属性相同：

- 注册版本：1.0；
- 注册机构：中华人民共和国国家档案局；
- 字符集：GB 2312—1980、GB 18030—2005；
- 语言：中文。

表 5 元数据元素描述方法

编号	按一定规则排列的元数据的顺序号
中文名称	元数据元素的中文标识
英文名称	元数据元素的英文标识
定义	元数据元素含义的描述
目的	描述该元数据元素的必要性和作用
约束性	说明采用该元数据元素的强制性程度，分“必选”、“条件选”和“可选”。“必选”表示总是强制采用；“条件选”表示在特定环境和条件下必须采用；“可选”表示可采用也可不采用，由用户根据需要确定
可重复性	元数据元素是否可以重复出现

表 5 元数据元素描述方法(续表)

元素类型	元数据元素所属的类别。本标准将元素分为简单型、容器型和复合型
数据类型	为表达元数据元素值而规定的具有相同数学特性和相同操作集的数据类别。本标准数据类型包括字符型、数值型、日期时间型。容器型元素没有数据类型
编码修饰体系	对该元数据元素信息的描述应遵循的编码规则,包括编码修饰体系的标识和名称
值域	可以分配给元数据元素的值
缺省值	该元数据元素的默认值
子元素	该元数据元素具有的下属元素
信息来源	元数据元素值的来源
相关元素	与该元素有密切联系的元素
注释	对元素的进一步说明

5 文件实体元数据元素的描述

5.1 聚合层次

编号	M1
中文名称	聚合层次
英文名称	aggregation level
定义	电子文件在分类、整理、著录、保管和提供利用时,作为个体和特定群体的控制层次。如宗、目、卷、件等
目的	利于对电子文件的管理,为电子文件在某一层次的著录、检索提供条件
约束性	必选
可重复性	不可重复
元素类型	简单型
数据类型	字符型
编码修饰体系	——
值域	案卷 文件
缺省值	文件
子元素	——
信息来源	在形成、处理、管理电子文件的系统中通过预定义值域列表选择著录或手工著录
相关元素	——
注释	推荐使用预定义值域列表选择著录

5.2 来源

编号	M2
中文名称	来源
英文名称	provenance
定义	对形成、归档、移交、保管电子文件的单位(个人)和全宗的描述
目的	提供电子文件的背景信息, 利于对电子文件的理解、管理和利用
约束性	必选
可重复性	不可重复
元素类型	容器型
数据类型	——
编码修饰体系	——
值域	——
缺省值	——
子元素	档案馆名称(M3) 档案馆代码(M4) 全宗名称(M5) 立档单位名称(M6)
信息来源	——
相关元素	——
注释	——

5.2.1 档案馆名称

编号	M3
中文名称	档案馆名称
英文名称	archives name
定义	集中管理特定范围档案的专门机构的名称
目的	标识电子文件所属馆藏机构, 利于电子文件交换、公布和全国性、区域性目录中心建设
约束性	条件选
可重复性	不可重复
元素类型	简单型
数据类型	字符型
编码修饰体系	——
值域	——
缺省值	——
子元素	——
信息来源	在管理电子文件的系统中通过预定义值域列表选择著录或手工著录
相关元素	档案馆代码(M4)
注释	由档案馆著录, 在建立目录中心或公布、交换电子文件时, 本元素必选。推荐使用预定义值域列表选择著录

5.2.2 档案馆代码

编号	M4	
中文名称	档案馆代码	
英文名称	archives identifier	
定义	依据《编制全国档案馆名称代码实施细则》对档案馆所赋予的代码	
目的	唯一标识档案馆, 利于电子文件的管理和利用	
约束性	条件选	
可重复性	不可重复	
元素类型	简单型	
数据类型	字符型	
编码修饰体系	标识	名称
	国档发[1987]4号	编制全国档案馆名称代码实施细则
值域	——	
缺省值	——	
子元素	——	
信息来源	在管理电子文件的系统中通过预定义值域列表选择著录或手工著录	
相关元素	档案馆名称(M3)	
注释	由档案馆著录, 在建立目录中心或公布、交换电子文件时, 本元素必选。推荐使用预定义值域列表选择著录	

5.2.3 全宗名称

编号	M5	
中文名称	全宗名称	
英文名称	fonds name	
定义	档案馆(室)赋予全宗的标识	
目的	区分电子文件来源, 利于电子文件的管理, 提供检索点	
约束性	条件选	
可重复性	不可重复	
元素类型	简单型	
数据类型	字符型	
编码修饰体系	——	
值域	——	
缺省值	——	
子元素	——	
信息来源	在形成、处理、管理电子文件的系统中通过预定义值域列表选择著录或手工著录	
相关元素	立档单位名称(M6) 全宗号(M9)	
注释	全宗名称一般由接收、保管电子文件的档案馆(室)给定, 当给定全宗名称时, 本元素必选。推荐使用预定义值域列表选择著录	