



中华人民共和国档案行业标准

DA/T 33—2005

明清档案目录中心数据采集标准 明清档案机读目录数据交换格式

Data collection standard of China catalogue center for Ming-Qing
archives—The machine-readable catalogue exchange for
Ming-Qing archives

2005-04-30 发布

2005-09-01 实施

前 言

本标准针对国家档案局全国历史档案资料目录领导小组关于建立完善全国明清档案目录数据库的要求,提出了全国明清档案案卷级目录数据采集标准中数据交换的格式,为全国各级各类档案馆明清档案著录、数据采集、交换等工作的业务技术人员提供可操作的标准。

本标准由国家档案局提出并归口。

本标准起草单位:中国第一历史档案馆全国明清档案资料目录中心。

本标准主要起草人:邢永福、冯伯群、胡忠良、苏文英、郭慧。

明清档案目录中心数据采集标准

明清档案机读目录数据交换格式

1 范围

本标准规定了软磁盘、光盘等作为载体交换明清档案机读目录数据时所使用的格式。

本标准规定的明清档案机读目录数据文件的组织形式为交换格式的文本文件。

本标准适用于全国明清档案目录中心机读目录数据信息交换。

2 规范性引用文件

下列文件中的条款通过本标准的引用而成为本标准的条款。凡是注日期的引用文件，其随后所有的修改单(不包括勘误的内容)或修订版均不适用于本标准，然而，鼓励根据本标准达成协议的各方研究是否可使用这些文件的最新版本。凡是不注日期的引用文件，其最新版本适用于本标准。

GB/T 1988—1998 信息技术 信息交换用七位编码字符集

GB/T 2901—1992 书目信息交换用磁带格式

GB 2312—1980 信息交换用汉字编码字符集 基本集

GB 18030—2000 信息技术 信息交换用汉字编码字符集 基本集的扩充

DA/T 8—1994 明清档案著录细则

3 术语和定义

下列术语和定义适用于本标准。

3.1

档案目录 catalogue of archives

按照一定的秩序编排而成的条目组合，是档案管理、检索和报道的工具。

3.2

机读目录 machine-readable catalogue

将档案目录按照一定格式输入计算机内，可以由计算机读取、管理、检索的档案目录信息。

3.3

数据 data

对事物的概念和属性进行描述的规范化表示。

3.4

记录 record

字段的集合，完整描述一事物相关数据处理单元。

3.5

字段 field

记录中可标识的特定范围数据，可包含一个或多个子字段。

3.6

子字段 sub-field

字段的一部分，它含有一个确定的数据单元。

3.7

数据元素 data-element

被明确标识的最小数据信息单位。

3.8

字段标识符 tag

识别字段的一组三位数字符号。

3.9

子字段标识 sub-field identifier

识别子字段的符号,由(b~j)字母标识。

3.10

指示符 indicator

字段标识符后,提供有关字段内容的附加信息。用(0~9)数字标识。

3.11

分隔符 separator

用来限定记录、字段、数据单元的控制符。

记录分隔符 &, 字段分隔符 @, 子字段标识符 \$。

4 交换格式的记录构成

4.1 记录结构

记录由一组按一定格式写入的物理记录序列构成。

一条记录由记录头标、若干字段和记录尾标组成。

记录头标	字段 1	字段 N	记录尾标
------	------	-------	-------	------	------

记录头标 标识记录的开始信息,由 1 位字符组成。

其取值, M 代表明朝档案; Q 代表清朝档案。

记录尾标 标识记录的结束信息,由 1 位字符组成。

其取值, "&".

4.2 字段结构

字段标识符	指示符	子字段标识符(可选)	数据区	字段分隔符
3 位字符	1 位字符	2 位字符	N 位字符	1 个字符

字段标识符 标识该字段数据区所属的字段,其取值为三位数字。

指示符 标识对应字段的属性。其取值为数字,空时为“#”。

子字段标识符 由 2 位字符组成,标识字段中所属的子字段,

其第一位取值“\$”。

第二位取值为小写字母。

字段分隔符 指示该字段数据区结束。其取值为“@”。

4.3 文件结构

由若干记录组成一个文件,文件起始处为说明记录。

说明记录
记录
.....
记录

4.4 字段标识符说明

下列各字段定义除有专门说明外,均与 DA/T 8—1994 相应的著录项的定义保持一致。

- 101 档号
指示符为空
子字段表
- | 标识符 | 内容 | 说明 |
|-----|-----|--------|
| \$b | 全宗号 | 字段不可重复 |
| \$c | 目录号 | 字段不可重复 |
| \$d | 案卷号 | 字段不可重复 |
| \$e | 件号 | 字段不可重复 |
| \$f | 页号 | 字段不可重复 |
- 102 缩微号
指示符为空 字段不可重复
- 103 档案馆代码
指示符为空 字段不可重复
- 104 电子文件号
指示符为空 字段不可重复
- 105 画幅数
指示符为空 字段不可重复
- 106 密级
指示符为空 字段不可重复
- 201 题名
指示符为空 字段不可重复
- 202 责任者
- | | | |
|-----|---|---------|
| 指示符 | 1 | 表示第一责任者 |
| 指示符 | 2 | 表示第二责任者 |
| 指示符 | 3 | 表示第三责任者 |
- 子字段表
- | 标识符 | 内容 | 说明 |
|-----|------|--------|
| \$b | 官职爵位 | 字段不可重复 |
| \$c | 责任人 | 字段不可重复 |
- 203 文本
指示符为空 字段不可重复
- 204 时间
- | | | |
|-----|---|-----------------|
| 指示符 | 1 | 表示文件原纪年(明、清朝纪年) |
| 指示符 | 2 | 表示公元纪年 |
| 指示符 | 3 | 表示外国纪年 |
- 子字段表
- | 标识符 | 内容 | 说明 |
|-----|--------|--------|
| \$b | 文件起始时间 | 字段不可重复 |
| \$c | 文件终止时间 | 字段不可重复 |
- 301 附注
指示符为空

子字段表

标识符	内容	说明
\$b	责任者注	字段不可重复
\$c	载体形态	字段不可重复
\$d	档案数量	字段不可重复
\$e	数量单位	字段不可重复
\$f	附件	字段不可重复
\$g	规格	字段不可重复
\$h	语种	字段不可重复
\$i	档案实体	字段不可重复
\$j	时间注	字段不可重复

601 主题词

指示符为空

子字段表

标识符	内容	说明
\$b	主表主题词	字段可重复
\$c	人名主题词	字段可重复
\$d	国家机构主题词	字段可重复
\$e	官职爵位主题词	字段可重复
\$f	清代民族主题词	字段可重复
\$g	行政区划主题词	字段可重复

602 分类号

指示符 1 表示第一分类主号

指示符 2 表示第二分类主号

指示符 3 表示第三分类主号

子字段表

标识符	内容	说明
\$b	分类主号	字段不可重复
\$c	地区复分号	字段可重复
\$d	民族复分号	字段可重复
\$e	国家复分号	字段可重复

901 提要项

指示符为空 字段不可重复

902 朝年政权标识

指示符为空 字段不可重复

903 控制符

指示符为空 字段不可重复

5 XML Schema 模式定义

<schema>
<element name="明清档案目录">
<complexType>

续表

<sequence>
<element name="档案馆代码" type="string" />
<element name="全宗号" type="string"/>
<element name="目录号" type="string" />
<element name="案卷号" type="string" />
<element name="件号" type="string"/>
<element name="页号" type="string"/>
<element name="缩微号" type="string"/>
<element name="主分类号 A" type="string"/>
<element name="A 地区 1" type="string"/>
<element name="A 地区 2" type="string"/>
<element name="A 地区 3" type="string"/>
<element name="A 地区 4" type="string"/>
<element name="A 地区 5" type="string"/>
<element name="A 国家或民族 1" type="string"/>
<element name="A 国家或民族 2" type="string"/>
<element name="A 国家或民族 3" type="string"/>
<element name="A 国家或民族 4" type="string"/>
<element name="A 国家或民族 5" type="string"/>
<element name="主分类号 B" type="string"/>
<element name="B 地区 1" type="string"/>
<element name="B 地区 2" type="string"/>
<element name="B 地区 3" type="string"/>
<element name="B 地区 4" type="string"/>
<element name="B 地区 5" type="string"/>
<element name="B 国家或民族 1" type="string"/>
<element name="B 国家或民族 2" type="string"/>
<element name="B 国家或民族 3" type="string"/>
<element name="B 国家或民族 4" type="string"/>
<element name="B 国家或民族 5" type="string"/>
<element name="主分类号 C" type="string"/>
<element name="C 地区 1" type="string"/>
<element name="C 地区 2" type="string"/>
<element name="C 地区 3" type="string"/>
<element name="C 地区 4" type="string"/>
<element name="C 地区 5" type="string"/>

续表

<element name=" C 国家或民族 1" type="string"/>
<element name=" C 国家或民族 2" type="string"/>
<element name=" C 国家或民族 3" type="string"/>
<element name=" C 国家或民族 4" type="string"/>
<element name=" C 国家或民族 5" type="string"/>
<element name="责任者 A" type="string"/>
<element name="责任者 B" type="string"/>
<element name="责任者 C" type="string"/>
<element name="责任者注" type="string"/>
<element name="官职爵位 A" type="string"/>
<element name="官职爵位 B" type="string"/>
<element name="官职爵位 C" type="string"/>
<element name="题名" type="string"/>
<element name="文本" type="string"/>
<element name="语种" type="string"/>
<element name="原纪年起始时间" type="string"/>
<element name="原纪年终止时间" type="string"/>
<element name="公元纪年起始时间" type="string"/>
<element name="公元纪年终止时间" type="string"/>
<element name="外国纪年起始时间" type="string"/>
<element name="外国纪年终止时间" type="string"/>
<element name="主题词" type="string"/>
<element name="附注" type="string"/>
<element name="提要项" type="string"/>
<element name="档案数量" type="integer"/>
<element name="数量单位" type="string"/>
<element name="档案属性" type="string"/>
<element name="朝代政权标识" type="string"/>
<element name="控制符" type="string"/>
<element name="电子文件号" type="string"/>
<element name="画幅数" type="integer"/>
<element name="密级" type="string"/>
</sequence>
</complexType>
</schema>

6 文件数据字段映射表

序号	字段名称	字段类型	字段标识符
1	<档案馆代码>	字符	103
2	<全宗号>	字符	101 \$ b
3	<目录号>	字符	101 \$ c
4	<案卷号>	字符	101 \$ d
5	<件号>	字符	101 \$ e
6	<页号>	字符	101 \$ f
7	<缩微号>	字符	102
8	<主分类号 A>	字符	6021 \$ b
9	<A 地区 1>	字符	6021 \$ c
10	<A 地区 2>	字符	6021 \$ c
11	<A 地区 3>	字符	6021 \$ c
12	<A 地区 4>	字符	6021 \$ c
13	<A 地区 5>	字符	6021 \$ c
14	<A 民族 1>	字符	6021 \$ d
15	<A 民族 2>	字符	6021 \$ d
16	<A 民族 3>	字符	6021 \$ d
17	<A 民族 4>	字符	6021 \$ d
18	<A 民族 5>	字符	6021 \$ d
19	<A 国家 1>	字符	6021 \$ e
20	<A 国家 2>	字符	6021 \$ e
21	<A 国家 3>	字符	6021 \$ e
22	<A 国家 4>	字符	6021 \$ e
23	<A 国家 5>	字符	6021 \$ e
24	<主分类号 B>	字符	6022 \$ b
25	<B 地区 1>	字符	6022 \$ c
26	<B 地区 2>	字符	6022 \$ c
27	<B 地区 3>	字符	6022 \$ c
28	<B 地区 4>	字符	6022 \$ c
29	<B 地区 5>	字符	6022 \$ c
30	<B 民族 1>	字符	6022 \$ d
31	<B 民族 2>	字符	6022 \$ d
32	<B 民族 3>	字符	6022 \$ d
33	<B 民族 4>	字符	6022 \$ d
34	<B 民族 5>	字符	6022 \$ d

续表

序号	字段名称	字段类型	字段标识符
35	<B 国家 1>	字符	6022 \$ e
36	<B 国家 2>	字符	6022 \$ e
37	<B 国家 3>	字符	6022 \$ e
38	<B 国家 4>	字符	6022 \$ e
39	<B 国家 5>	字符	6022 \$ e
40	<主分类号 C>	字符	6023 \$ b
41	<C 地区 1>	字符	6023 \$ c
42	<C 地区 2>	字符	6023 \$ c
43	<C 地区 3>	字符	6023 \$ c
44	<C 地区 4>	字符	6023 \$ c
45	<C 地区 5>	字符	6023 \$ c
46	<C 民族 1>	字符	6023 \$ d
47	<C 民族 2>	字符	6023 \$ d
48	<C 民族 3>	字符	6023 \$ d
49	<C 民族 4>	字符	6023 \$ d
50	<C 民族 5>	字符	6023 \$ d
51	<C 国家 1>	字符	6023 \$ e
52	<C 国家 2>	字符	6023 \$ e
53	<C 国家 3>	字符	6023 \$ e
54	<C 国家 4>	字符	6023 \$ e
55	<C 国家 5>	字符	6023 \$ e
56	<责任者 A>	汉字	2021 \$ c
57	<责任者 B>	汉字	2022 \$ c
58	<责任者 C>	汉字	2023 \$ c
59	<责任者注>	字符	301 \$ b
60	<官职爵位 A>	汉字	2021 \$ b
61	<官职爵位 B>	汉字	2022 \$ b
62	<官职爵位 C>	汉字	2023 \$ b
63	<题名>	汉字	201
64	<文本>	汉字	203
65	<语种>	汉字	301 \$ h
66	<原纪年起始时间>	字符	2041 \$ b
67	<原纪年终止时间>	字符	2041 \$ c
68	<公元纪年起始时间>	字符	2042 \$ b
69	<公元纪年终止时间>	字符	2042 \$ c

续表

序号	字段名称	字段类型	字段标识符
70	<外国纪年起始时间>	字符	2043 \$ b
71	<外国纪年终止时间>	字符	2043 \$ c
72	<主题词>	汉字	601
73	<附注>	汉字	301
74	<提要项>	汉字	901
75	<档案数量>	数字	301 \$ d
76	<数量单位>	字符	301 \$ e
77	<档案属性>	字符	无标识符,内容为记录头标信息
78	<朝代政权标识>	字符	902
79	<控制符>	字符	903
80	<电子文件号>	字符	104
81	<画幅数>	数字	105
82	<密级>	字符	106

实例：清代军机处录副奏折著录条目

档号:0001-001

缩微号:001-0015

文本:奏折

官职爵位 A:都察院左副都御史

责任者 A:二格

文件起始时间:061101026

题名:奏为请设口外道员以专责成事

分类主号 A:C111

A 地区 1:61

A 地区 2:62

记录格式:

Q101# \$ d0001 \$ e001@102#001-0015@201#奏为请设口外道员以专责成事@2021 \$ b 都察院左副都御史 \$ c 二格@203#奏折@2041 \$ b061101026@6021 \$ bC111 \$ c61 \$ c62&