ICS 35.240
CCS L 70

# 中华人民共和国国家标准

GB/T XXXXX—XXXX
代替 GB/T 31219.4-2014

# 图书馆馆藏资源数字化加工规范 第4部分：音频资源

Specification of library collections digitization Part 4 : Audio resources

（征求意见稿）

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上。

XXXX - XX - XX 发布　　XXXX - XX - XX 实施

国家市场监督管理总局
国家标准化管理委员会　发布

# 目　　次

前言 …… III
引言 …… IV
1　范围 …… 1
2　规范性引用文件 …… 1
3　术语和定义 …… 1
4　缩略语 …… 2
5　加工级别及技术参数 …… 3
5.1　加工级别 …… 3
5.2　主文件技术参数 …… 3
5.3　服务文件技术参数 …… 3
6　加工准备 …… 4
6.1　应对拟加工音频资源进行盘点、核查和筛选，必要时对音频资源的载体进行适度的清洗和修复。 …… 4
6.2　应为需加工的音频资源选择合适的播放设备，对播放设备进行清洁、调整，设定符合正确播放的参数。 …… 4
6.3　应在加工之前对加工级别及技术参数、采集与处理、元数据加工、命名、质量管理等进行一致性约定。 …… 4
7　采集与处理 …… 4
7.1　内容采集与转换 …… 4
7.2　编辑与处理 …… 4
8　元数据加工 …… 4
8.1　元数据著录 …… 4
8.2　元数据关联 …… 7
8.3　信息封装 …… 7
9　命名规则 …… 8
9.1　应为加工后的音频数据文件或文件夹进行命名。 …… 8
9.2　文件或文件夹的命名规则应至少符合以下规则： …… 8
10　质量管理 …… 8
10.1　应在加工过程中对音频数据文件的质量进行控制。 …… 8
10.2　对已加工的音频数据文件的质量应符合以下要求： …… 8
参考文献 …… 9

表 1　音频资源加工技术参数 …… 3
表 2　音频数据应标记的内容项 …… 4
表 3　音频数据文件信息块应封装的信息 …… 7

# 前　言

本部分依据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》给出的规则起草。

本文件代替 GB/T 31219—2014《图书馆馆藏资源数字化加工规范 第4部分：音频资源》，与GB/T 31219—2014相比，除编辑性改动外，主要技术变化如下：

1)按GB/T 1.1—2020修改文件格式和内容，修改“前言”、增加“引言”。

2)修改和增加术语：

a） 修改：术语中的sample precision原标准翻译为量化精度，现在普遍翻译为采样精度（见2014版的3.3）

b） 修改：术语中的Channel number（通道数）改为Num channels（见2014版的3.5）

c） 修改：术语中的“位（bit）”改为比特（见2014版的3.4）

d） 修改：主文件格式增加FLAC格式（见2014版的4.2）

e） 修改：服务文件格式增加AAC格式（见2014版的3.5）

f） 修改：MPEG Audio Layer3改为有损压缩（见2014版的4.3）

3)替换和增加规范性引用文件：

a） 替换：GB/T 4894-2009 《信息与文献 术语》替换为GB/T 4894-2024 《信息与文献 基础和术语》（见2014版的2）

b） 替换：IASA-TC 03The safeguarding of the Audio Heritage: Ethics, Principles and Preservation Strategy（2009，2th edition）替换为IASA-TC 03The safeguarding of the Audio Heritage: Ethics, Principles and Preservation Strategy（2017, 4th edition）

c） 替换：PREMIS Data Dictionary for Preservation Metadata(Version 2.0)[EB/OL].[2013-3-1]替换为PREMIS Data Dictionary for Preservation Metadata(Version 2.0)[EB/OL].[2015-6-1]

d） 增加：GB/T 25100.1-2025 信息与文献 都柏林核心元数据元素集 第1部分：核心元素（和GB/T 25100.2-2025《信息与文献 都柏林核心元数据元素集 第2部分：DCMI属性和类》（见2014版的2）

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这些专利的责任。

本部分由中华人民共和国文化和旅游部提出。

本部分由全国图书馆标准化技术委员会(SAT/TC389)归口。

本标准起草单位：北京大学、国家图书馆、清华大学、中国音乐学院。

本标准的主要起草人：童云海、宋庆生、曲天书、周建超、龙伟、童庆钧、罗四洪。

本文件首次发布于2014年，本次为第一次修订。

# 引　言

本文件是GB/T 31219《图书馆馆藏资源数字化加工规范》系列标准的第4部分。GB/T 31219系列标准已发布了以下部分：

— 第2部分：文本资源；

— 第3部分：图像资源；

— 第4部分：音频资源；

— 第5部分：视频资源。

图书馆馆藏资源数字化加工依托数字化设备与相关技术开展，二者发展水平直接影响资源数字化成果质量。随着数字化硬件持续迭代、技术指标不断提升，馆藏资源数字化加工的标准体系与实操规范也需要动态调整、与时俱进。本分册为馆藏资源数字化加工标准的第 4 部分，规定音频类馆藏资源数字化加工的推荐性要求与工作规范。

本文件为推荐性国家标准。各图书馆在实施音频资源数字化加工项目时，可结合自身设备技术参数、资源建设用途等现实条件酌情调整；元数据著录、文件命名等相关工作，可参照并引用现行有效的国际、国家及行业相关标准。

# 图书馆馆藏资源数字化加工规范 第4部分：音频资源

## 1 范围

本文件规定了音频资源转化为数字资源的加工要求和工作规范，包括音频资源数字化加工的加工级别及技术参数、加工准备、采集与处理、元数据加工、命名规则、质量管理。

本文件适用于图书馆音频资源数字化加工，其他文献信息机构音频资源的数字化加工也可参照使用。

## 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。

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/T 7408—2005 数据元和交换格式 信息交换 日期和时间表示法（见本文件的8）

GB/T 25100（所有部分） 信息与文献 都柏林核心元数据元素集（见本文件的8）

## 3 术语和定义

下列术语和定义适用于本文件。

3.1

**音频资源 audio resource**

数字加工前，不符合加工标准，以模拟或数字形态存在的，可被人耳感知的声音文件。

3.2

**音频数据 audio data**

数字加工后，符合加工标准，以数字形态存在的声音文件。

3.3

**采样率 sampling frequency**

每秒从连续信号中提取并组成离散信号的采样个数。

**示例：**44.1千赫兹（kHz）。

3.4

**采样精度 sample precision**

度量采样后离散信号幅度的分级数。

**注：**通常用“比特（bit）”作单位。

16比特（bit）、24比特（bit）。

3.5

**通道数 Num channels**

描述音频信号的声道数。

示例：单声道（mono）；多声道（multichannel）。

3.6

**比特率 bit rate**

每秒传送或处理的数据位数，也称为“码率”，用 bps(bits per second)作单位。

示例：1.2Mbps。

3.7

**特殊记录 special recording**

珍贵、稀有或有特别用途的音频资源。

3.8

**语音记录 voice recording**

只有人的声音，不包含其他任何乐音、动物声音的音频资源。

注：采访或访谈录音中除了受访者的声音，应无其他任何虫鸣、鸟叫、乐器以及背景音的音频资源。

3.9

**主文件 master file**

以存档或长期保存为目的的音频数据。

3.10

**标准主文件 standard master file**

除了特殊记录、语音记录之外的以存档或长期保存为目的的音频数据。

注：通常情况下，以存档或长期保存为目的的数字音频数据都加工为标准主文件。

3.11

**服务文件 service file**

以获取或使用为目的的音频数据。

3.12

**封装 wrap**

将元数据信息与数字音频数据相嵌入，使之形成一个有机的整体。

3.13

**载体 carrier**

为表示信息而应用于数据媒介的材料。

注：来源于 GB/T 4894-2024。

示例：磁带、密纹唱片等。

## 4 缩略语

下列缩略语适用于本文件。

WAVE 波形音频文件格式（waveform audio file format）

BWF 广播波形格式（broadcast wave format）

FLAC 自由无损音频编码（Free Lossless Audio Code）

MP3 有损音频编码（MPEG-1 Audio Layer 3）

AAC 高级音频编码（Advanced Audio Coding）

## 5 加工级别及技术参数

### 5.1 加工级别

#### 5.1.1 音频数据的加工级别

应根据音频资源的加工目的和使用要求设定加工级别为主文件、服务文件。

#### 5.1.2 主文件的类别

主文件可分为标准主文件、特殊记录主文件和语音记录主文件。

#### 5.1.3 服务文件的类别

服务文件可分为无损服务文件、标准音质服务文件和高保真音质服务文件。

### 5.2 主文件技术参数

**表 1 音频资源加工技术参数**

| 应用级别 | 通道数 | 采样率 | 量化位 | 文件格式 | 压缩算法 |
|---|---|---|---|---|---|
| 标准主文件 | 单声道或多声道，取决于音频资源的声道数 | 44.1kHz | 16bit | WAVE<br>BWF<br>FLAC | 无压缩或无损压缩 |
| 特殊记录主文件 | 单声道或多声道，取决于音频资源的声道数 | 96kHz 或以上 | 96kHz 或以上 | WAVE<br>BWF<br>FLAC | 无压缩或无损压缩 |
| 语音记录主文件 | 单声道或多声道，取决于音频资源的声道数 | 22.5kHz 或以上 | 16bit | WAVE<br>BWF<br>FLAC | 无压缩或无损压缩 |

### 5.3 服务文件技术参数

#### 5.3.1 无损服务文件

可将主文件的副本作为服务文件交付使用。无损服务文件的技术参数参见标准主文件（5.2）、特殊记录主文件（5.2）、语音记录主文件（5.2）。

#### 5.3.2 标准音质服务文件

编码：有损压缩

通道数：单声道或多声道，取决于音频资源的声道数

比特率：128kbps

文件格式：MP3、AAC

#### 5.3.3 高保真音质服务文件

编码：有损压缩

通道数：单声道或多声道，取决于音频资源的声道数
比特率：320kbps
文件格式：MP3、AAC

## 6 加工准备

6.1 应对拟加工音频资源进行盘点、核查和筛选，必要时对音频资源的载体进行适度的清洗和修复。
6.2 应为需加工的音频资源选择合适的播放设备，对播放设备进行清洁、调整，设定符合正确播放的参数。
6.3 应在加工之前对加工级别及技术参数、采集与处理、元数据加工、命名、质量管理等进行一致性约定。

## 7 采集与处理

### 7.1 内容采集与转换

7.1.1 可先将拟加工的音频资源加工成主文件，并由主文件转换或派生服务文件。

7.1.2 在采集为主文件时应将音频资源的内容进行整体采集。

示例 1：磁带 A 面作为一个整体进行采集，包含 A 面全部曲目。
示例 2：如果采集设备有自动翻面功能，可将磁带 A 面、B 面合为一个整体进行采集，包含 A 面、B 面全部曲目。

### 7.2 编辑与处理

7.2.1 加工为主文件时不进行任何处理。
7.2.2 加工为服务文件时可进行必要的处理，包括降噪等技术处理。

7.2.3 加工为服务文件时应按照音频数据的内容单元进行剪切或拼接。

示例：主文件包含 10 首曲目，转换或派生为服务文件时应分别剪切为独立的 10 首曲目。
7.2.4 作为服务文件的音频数据的内容单元前、后均应保留 2-3 秒空白。

## 8 元数据加工

### 8.1 元数据著录

8.1.1 应在加工过程中为主文件和服务文件著录描述元数据、管理元数据或结构元数据。

8.1.2 应在加工过程中对主文件和服务文件的加工信息进行标记。标记的信息应包括但不限于表 2 所列：

表 2 音频数据应标记的内容项

| 元素/修饰词名称 | 元素标签 | 修饰词标签 | 定义 | 注释 |
| --- | --- | --- | --- | --- |
| identifier | 标识符 | | 唯一识别一个管理元数据记录的标识 | 指特定应用系统内具有唯一识别性的标识 |

表 2 音频数据应标记的内容项 （续）

| 元素/修饰词名称 | 元素标签 | 修饰词标签 | 定义 | 注释 |
|---|---|---|---|---|
| | | | | 符号。由标识应用系统的前缀（即标识符的类型）和一个字符串（即标识符的值）组成。由系统自动产生或由人工赋予。 |
| infoResource | 信息资源 | | 任何可标识的信息对象 | 须通过信息系统加以管理的基本对象。 |
| infoResource Identifier | | 信息资源标识符 | 唯一识别信息资源的标识 | 指特定应用系统内具有唯一识别性的标识符号。由标识应用系统的前缀（即标识符的类型）和一个字符串（即标识符的值）组成。由系统自动产生或由人工赋予。 |
| source | | 来源 | 对生成当前信息资源的资源或其他实体的参照 | 可用正式标识体系的字符串表示，如 URI 以及其他标识非数字资源的编码体系（ISBN，ISSN，ISRC 等）。对于派生自其他数字资源的数字对象来说，其来源信息在描述元数据中反映，如 DC 的“关系（relation）”。 |
| technicalInfo | 技术信息 | | 与信息资源的创建、加工、使用相关的物理参数、技术手段与标准以及硬件环境。 | 可嵌入或链接通行的技术元数据。 |
| eventIdentifier | | 事件标识符 | 唯一识别事件的标识。 | 主要用于管理元数据记录内部关联元素之间的链接，如代理与其相关的事件，建议由系统自动生成。 |
| eventType | | 事件类型 | 根据信息资源生命周期的基本阶段对事件划分的大类。 | 采集、数字化、元数据加工等。 |
| action | | 操作 | 一个事件中有特定意义的细分的行动。 | 常见的操作有：资源选择、资源评估、资源供应，原始编目、套录、扫描、查询、创建、复制、修改、删除、合并等。建议建立受控词汇表或代码表，规范操作的取值。 根据实际应用的需求对操作的细化程度作出规定。建议 |

表 2 音频数据应标记的内容项 （续）

| 元素/修饰词名称 | 元素标签 | 修饰词标签 | 定义 | 注释 |
|---|---|---|---|---|
| | | | | 用层次表明细化的程度，如：原始编目.修改。 |
| actionDateTime | | 操作日期时间 | 操作发生的日期时间，如图像采集时间、图像处理日期等。 | 日期时间的表达采用GB/T7408-2005，或其它可转换成GB/T7408-2005的表达方式。如：2005-11-05T08:15:30+08:00 [2005年11月5日北京时间16时15分30秒]。 |
| format | | 格式 | 信息资源的物理或数字表现形式。 | 信息资源的内容形式，包括资源内容与其元数据的类型。 |
| samplingFrequency | | 采样频率 | 每秒从连续信号中提取并组成离散信号的采样个数。 | 音频信号采样频率，如22.05KHz、44.1KHz、48KHz等。 |
| bitRate | | 比特率 | 单位时间内传送或处理的数据位数。 | 也称为数据速率。用bps（bits per second）作单位，如：1.5Mbps、4Mbps。 |
| numberOfChannels | | 通道数 | 传播音频信号的声道数。 | 一般有单声道、立体声、四声道环绕、5.1声道等。 |
| quantizationPrecision | | 采样精度 | 度量采样后离散信号幅度时的分级数。 | 把采样得到的信号幅度的样本值从模拟量转换成数字量，数字量的二进制位数即采样精度，如：8位、16位、或8-bit、16-bit。 |
| processingMode | | 加工方式 | 数字信息资源的加工方式。 | 数字信息资源的加工方式。如：采集，格式转换等。 |
| rights | 权限 | | 与某一信息资源或代理有关的一种或多种权利或许可的说明 | 包括知识产权或其他各种权利的说明。 |
| agent | 代理 | | 有行为或有行为能力的一类资源。 | 是在信息资源的生命周期中发挥某种作用的一个人、一群人或一个自动装置。 |
| agentIdentifier | | 代理标识符 | 唯一识别代理的标识。 | 一般是特定应用系统内具有唯一识别性的标识符号,建议由系统自动生成。 |

表 2 音频数据应标记的内容项 （续）

| 元素/修饰词名称 | 元素标签 | 修饰词标签 | 定义 | 注释 |
|---|---|---|---|---|
| agentName | | 代理名称 | 代理的名称，包括职称、所在单位等。 | 可以是个人、团体或者自动装置。代理可以没有名称或可以不同步其名称。 |
| agentType | | 代理类型 | 根据代理的定义对其划分的基本大类。 | 建立受控词汇表，规范类型的取值，建议为：个人，团体，软件。 |
| event | 事件 | | 一个行动或一件发生的事情。 | 在信息资源的管理过程中，特指与信息资源相关联的行动或事情。 |

## 8.2 元数据关联

### 8.2.1 可在元数据著录时对不同对象进行关联。

### 8.2.2 关联的对象可包括，但不限于：

a)元数据之间；

b)元数据与音频数据来源资源之间；

c)对象数据与来源资源之间。

示例：可通过在元数据中设置一个“关联（Relation）”字段来专门描述不同对象之间的关联关系。

## 8.3 信息封装

### 8.3.1 可为符合标准技术参数的 WAVE 或 BWF、FLAC、MP3、AAC 等音频数据封装基本信息。

### 8.3.2 所封装的信息应包括但不限于表 3 所列：

表 3 音频数据文件信息块应封装的信息

| 信息块 | 说明和示例 |
|---|---|
| 题名 | 音频数据文件对应的名称<br>示例：梁山伯与祝英台 |
| 日期 | 数字化日期，格式为 YYYY-MM-DD<br>1999-11-04 |
| 存档位置 | 机构的唯一标识符<br>北京大学图书馆 TK-0043 |
| 归属单位 | 加工后音频数据文件的归属单位<br>北京大学图书馆 |
| a） 同时可参见 GB/T 7408-2005<br>b） 为存档机构名<br>c） 为数字资源原始文件的唯一标识符，一般为存档机构内的典藏号 | |

## 9 命名规则

9.1 应为加工后的音频数据文件或文件夹进行命名。

9.2 文件或文件夹的命名规则应至少符合以下规则：

a） 拥有唯一标识符、具备连续一致的结构；
b） 文件或文件夹命名应严格遵守计算机系统对文件命名的限制；
c） 文件扩展名采用三位半角小写字母；
d） 服务文件的命名应与主文件的命名保持一定的继承性。

## 10 质量管理

10.1 应在加工过程中对音频数据文件的质量进行控制。

10.2 对已加工的音频数据文件的质量应符合以下要求：

a） 文件被完整采集；
b） 所有文件能正常打开和播放；
c） 噪音或失真不超过原有记录；
d） 所有文件达到本标准所规定的采样率、量化位等技术参数规格；
e） 服务文件音质符合重放标准；
f） 文件名、文件夹名和文件存储路径正确。

# 参 考 文 献

[1] GB/T 3102（所有部分） 量和单位

[2] GB/T 4894—2024 信息与文献 基础和术语

[3] GB/T 17975.3-2002 信息技术 运动图像及其伴音信号的通用编码 第3部分：音频

[4] GB/T 17975.7-2002 信息技术 运动图像及其伴音信息的通用编码 第7部分:先进音频编码(AAC)

[5] GB/T 33475.3-2018 信息技术 高效多媒体编码 第3部分：音频

[6] GY/T 168-2001 广播音频数据文件格式规范—广播波形格式（BWF）

[7] GY/T 202.2-2016 广播电视音像资料编目规范 第2部分：音频资料

[8] WH/T 52-2012 管理元数据规范

[9] ISO/IEC 11172-3:1993 信息技术 用于数字存储媒体的运动图像及其伴音编码（最高约1.5 Mbit/s）——第3部分：音频（Information technology — Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s — Part 3: Audio）

[10] ISO/IEC 15938-2:2002 信息技术 多媒体内容描述界面 第2部分：描述定义语言（Information technology - Multimedia content description interface - Part 2：Description definition language）

[11] ISO/IEC TR 15938-11:2005 信息技术 多媒体内容描述界面 第11部分：MPEG-7的配置模式（Information technology — Multimedia content description interface — Part 11 MPEG-7 profile schemas）

[12] RFC 9639 Free Lossless Audio Codec (FLAC)

[13] IASA-TC 03 The safeguarding of the Audio Heritage: Ethics, Principles and Preservation Strategy(2017, 4th edition)

[14] IASA-TC 04 Guidelines on the Production and Preservation of Digital Audio Objects(2009, 2th edition) Guidelines on the Production and Preservation of Digital Audio Objects (web edition) | International Association of Sound and Audiovisual Archives (iasa-web.org)

[15] PREMIS Data Dictionary for Preservation Metadata(Version 2.0)[EB/OL].[2015-6-1].http://www.loc.gov/standards/premis/v2/premis-2-0.pdf

[16] PREMIS Data Dictionary for Preservation（ version 3.0 ）[2015-11].http://www.loc.gov/standards/premis/v3/index.html

[17] Library of Congress.AudioMD Data Dictionary[EB/OL].[2013-3-1].http://www.loc.gov/rr/mopic/avprot/DD_AMD.html

[18] The Safeguarding of the Audiovisual Heritage: Ethics, Principles and Preservation Strategy (web edition) | International Association of Sound and Audiovisual Archives (iasa-web.org)

[19] 朱强，张春红，龙伟.国家图书馆音频数据加工标准和操作指南.北京：国家图书馆出版社，2011