



中华人民共和国国家标准

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

2014-09-30 发布

2015-01-01 实施

中华人民共和国国家质量监督检验检疫总局
中国国家标准化管理委员会 发布

目 次

前言	Ⅲ
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 加工级别及技术参数	2
5 加工准备	4
6 采集与处理	4
7 元数据加工	4
8 命名规则	7
9 质量管理	7
参考文献	8

前 言

GB/T 31219《图书馆馆藏资源数字化加工规范》分为五个部分：

- 第1部分：总则；
- 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。

本部分为 GB/T 31219 的第4部分。

本部分依据 GB/T 1.1—2009 给出的规则起草。

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部提出。

本部分由全国图书馆标准化技术委员会(SAT/TC 389)归口。

本部分起草单位：北京大学图书馆、国家图书馆、清华大学图书馆、天津音乐学院图书馆、中国音乐学院图书馆、国泰东方信息技术有限公司。

本部分主要起草人：聂华、朱本军、宋庆生、崔海媛、黄涛、龙伟、李若滨、郑小惠、童庆钧、王建欣、高媛、罗四洪、黄佳、孙超。

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

GB/T 31219 的本部分规定了音频资源数字化加工的技术参数和基本工作规范,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。

本部分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 4894—2009 信息与文献 术语

GB/T 7408—2005 数据元和交换格式 信息交换 日期和时间表示法

3 术语和定义

下列术语和定义适用于本文件。

3.1

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)。

3.2

量化精度 quantitative level

度量采样后离散信号幅度的分级数,通常用“位(bit)”作单位。

示例：16 位(bit)、24 位(bit)。

3.3

通道数 channels

传播音频信号的声道数。

示例：单声道(mono);多声道(multi channel)。

3.4

比特率 bit rate

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

示例：1.2 Mbps。

3.5

特殊记录 special recording

珍贵、稀有或有特别用途的音频资源。

3.6

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。

3.7

主文件 master file

以存档或长期保存为目的的音频数据。

3.8

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。

3.9

服务文件 service file

以获取或使用为目的的音频数据。

3.10

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。

3.11

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。

3.12

音频数据 audio data

数字加工后,符合加工标准、且以数字形态存在的声音文件。

3.13

载体 carrier

为表示信息而应用于数据媒介的材料。

[GB/T 4894—2009,定义 4.6.1.6]

示例:磁带、密纹唱片等。

4 加工级别及技术参数

4.1 加工级别

4.1.1 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。

4.1.2 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。

4.1.3 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。

4.2 主文件技术参数

4.2.1 标准主文件

编码:无压缩或无损压缩
通道数:单声道或多声道,取决于音频资源的声道数
采样率:44.1 kHz
量化位:16 bit
文件格式:WAVE 或 BWF

4.2.2 特殊记录主文件

编码:无压缩或无损压缩
通道数:单声道或多声道,取决于音频资源的声道数
采样率:96 kHz 或以上
量化位:24 bit
文件格式:WAVE 或 BWF

4.2.3 语音记录主文件

编码:无压缩或无损压缩
通道数:单声道或多声道,取决于音频资源的声道数
采样率:22.05 kHz 或以上
量化位:16 bit
文件格式:WAVE 或 BWF

4.3 服务文件技术参数

4.3.1 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(4.2.1)、特殊记录主文件(4.2.2)、语音记录主文件(4.2.3)。

4.3.2 标准音质服务文件

编码:MPEG Audio Layer3
通道数:单声道或多声道,取决于音频资源的声道数
比特率:128 kbps
文件格式:MP3

4.3.3 高保真音质服务文件

编码:MPEG Audio Layer3
通道数:单声道或多声道,取决于音频资源的声道数
比特率:320 kbps
文件格式:MP3

5 加工准备

- 5.1 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。
- 5.2 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。
- 5.3 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。

6 采集与处理

6.1 内容采集与转换

- 6.1.1 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。
- 6.1.2 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。

6.2 编辑与处理

- 6.2.1 加工为主文件时不进行任何处理。
- 6.2.2 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。
- 6.2.3 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。
- 6.2.4 作为服务文件的音频数据的内容单元前、后均应保留 2 s~3 s 空白。

7 元数据加工

7.1 元数据著录

- 7.1.1 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。
- 7.1.2 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 1 所列:

表 1 音频数据应标记的内容项

名称	标签	定义	注释
infoResourceIdentifier	信息资源标识符	唯一识别信息资源的标识	一般是特定应用系统内具有唯一识别性的标识符号
source	来源	对生成本信息资源的资源或其他实体的参照	对数字资源来说,指派生出本资源的非数字资源。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)。对元数据来说,指描述信息资源内容的元数据的原始编制机构。可用编制机构的名称或代码表示。对非数字资源或者派生自其他数字资源的数字资源来说,其来源信息一般在资源的描述元数据中反映,如 DC 的“关系(relation)”

表 1 (续)

名称	标签	定义	注释
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	也可嵌入或链接目前通行的技术元数据
eventIdentifier	事件标识符	唯一识别事件的标识	一般是特定应用系统内具有唯一识别性的标识符号。可由标识应用系统的前缀(即标识符的类型)与字符串(即标识符的值)组成。可由系统自动产生或由人工赋予
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	常见的类型有:采集、数字化、元数据加工等。建议建立受控词汇表,规范类型的取值
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或日期时间的范围	日期时间的表达采用 GB/T 7408—2005,或其他可转换成 GB/T 7408—2005 的表达方式。如:1994-11-05T08:15:30+08:00 [1994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	一般而言,格式可以包括资源的载体类型或资源的大小,格式元素可以用来决定展示或操作资源所需的软硬件或其他相应设备。建议采用来自于受控词表中的值(例如 Internet 媒体类型[MIME]定义的计算机媒体格式)
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:22.05 kHz、44.1 kHz、48 kHz 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(bit per second)作单位,如:1.5 Mbps、4 Mbps
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	量化精度	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量化精度,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	如:采集、格式转换等

表 1 (续)

名称	标签	定义	注释
equipment	设备	加工或使用信息资源的硬件环境	根据元素“事件”所规定的环境,区分加工资源与使用资源的设备。譬如:事件描述数字加工的环境,则设备用于信息资源的加工;事件描述信息服务的环境,则设备用于信息资源的使用。 数字信息资源的常用加工设备有:扫描仪、数码相机、数字摄像机、声卡、视频卡等
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。可由标识应用系统的前缀(即标识符的类型)与一字符串(即标识符的值)组成。可由系统自动产生或由人工赋予
agentName	代理名称	代理的名称	可以是个人、团体或自动装置的名称
agentType	代理类型	根据代理的定义对其划分的基本大类	建议建立受控词汇表,规范类型的取值。建议取值:个人、团体、软件

7.2 元数据关联

7.2.1 可在元数据著录时对不同对象进行关联。

7.2.2 关联的对象可包括,但不限于:

- 元数据之间;
- 元数据与音频数据来源资源之间;
- 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。

7.3 信息封装

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

7.3.2 所封装的信息应包括,但不限于表 2 所列:

表 2 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 示例:梁山伯与祝英台
日期	数字化日期,格式为 YYYY-MM-DD ^a 示例:1999-11-04
存档位置	机构的唯一标识符 示例:北京大学图书馆 ^b TK-0043 ^c
归属单位	加工后音频数据文件的归属单位 示例:北京大学图书馆
<p>^a 同时可见 GB/T 7408—2005。</p> <p>^b 为存档机构名。</p> <p>^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。</p>	

8 命名规则

- 8.1 应为加工后的音频数据文件或文件夹进行命名。
- 8.2 文件或文件夹的命名规则应符合,但不限于以下约束:
- 拥有唯一标识符、具备连续一致的结构;
 - 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;
 - 文件扩展名采用三位半角小写字母;
 - 服务文件的命名应与主文件的命名保持一定的继承性。

9 质量管理

- 9.1 应在加工过程中对音频数据文件的质量进行控制。
- 9.2 对已加工的音频数据文件的质量应符合,但不限于以下约束:
- 文件被完整采集;
 - 所有文件能正常打开和播放;
 - 噪音或失真不超过原有记录;
 - 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;
 - 服务文件音质符合重放标准;
 - 文件名、文件夹名和文件存储路径正确。

参 考 文 献

- [1] GB 3102(所有部分) 量和单位
- [2] 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
- [3] GY/T 168—2001 广播音频数据文件格式规范 广播波形格式(BWF)
- [4] GY/T 202.2—2007 广播电视音像资料编目规范 第2部分:广播资料
- [5] WH/T 52—2012 管理元数据规范
- [6] 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)
- [7] 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)
- [8] IASA-TC 03 The safeguarding of the Audio Heritage: Ethics, Principles and Preservation Strategy
- [9] IASA-TC 04 Guidelines on the Production and Preservation of Digital Audio Objects
- [10] PREMIS Data Dictionary for Preservation Metadata (Version 2.0) [EB/OL]. [2013-3-1]. <http://www.loc.gov/standards/premis/v2/premis-2-0.pdf>
- [11] IBM Corporation, Microsoft Corporation. Multimedia Programming Interface and Data Specifications (1.0) [EB/OL]. [2013-3-1]. <http://www-mmssp.ece.mcgill.ca/Documents/AudioFormats/WAVE/Docs/riffmci.pdf>
- [12] Library of Congress. AudioMD Data Dictionary [EB/OL]. [2013-3-1]. http://www.loc.gov/rr/mopic/avprot/DD_AMD.html
-