ICS 01.140.20
CCS A 14

DA

中华人民共和国档案行业标准

DA/T 110—2026

录音录像档案语音识别转写工作规范

Specification for speech recognition and transcription of audio and audio-visual records

2026-06-22 发布　　2026-09-01 实施

国家档案局　　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　总体原则和要求 …… 2
5　组织与管理 …… 3
5.1　工作组织 …… 3
5.2　工作人员 …… 3
5.3　工作场地 …… 3
5.4　管理制度 …… 3
5.5　工作流程 …… 3
5.6　工作文件 …… 3
5.7　元数据 …… 3
6　方案制定 …… 3
6.1　制定工作方案 …… 3
6.2　制定技术方案 …… 4
7　语音识别转写的实施 …… 4
7.1　档案出库 …… 4
7.2　评估与导入 …… 4
7.3　识别转写 …… 5
7.4　成果修改校正 …… 5
7.5　成果整理输出 …… 6
7.6　质量控制 …… 6
7.7　成果挂接 …… 7
7.8　成果验收 …… 7
7.9　档案入库 …… 7
8　语音识别转写成果管理与应用 …… 7
8.1　成果管理 …… 7
8.2　成果应用 …… 8
附录 A（资料性）　录音录像档案语音识别转写工作流程 …… 9
附录 B（资料性）　录音录像档案语音识别转写工作流程单 …… 10
参考文献 …… 11

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。

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。

本文件由国家档案局提出并归口。

本文件起草单位：安徽省档案馆、讯飞智元信息科技有限公司、中国人民大学信息资源管理学院。

本文件主要起草人：黄玉明、张海剑、朱霖露、邓晓文、徐拥军、牛力、陈姗姗、李衍、高峰。

录音录像档案语音识别转写工作规范

1 范围

本文件规定了录音录像档案语音识别转写的组织与管理、方案制定、实施、成果管理与应用的要求。

本文件适用于各级各类档案馆(室)以汉语、常见少数民族语言及国际通用语言录制的录音录像档案机器语音识别转写及相关系统建设应用。

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。

DA/T 18 档案著录规则

DA/T 62 录音录像档案数字化规范

DA/T 63 录音录像类电子档案元数据方案

DA/T 68.2 档案服务外包工作规范 第2部分:档案数字化服务

3 术语和定义

下列术语和定义适用于本文件。

3.1

录音录像档案 audio and audio-visual record

具有凭证、查考和保存价值并归档保存的录音录像文件。

[来源:DA/T 78—2019,3.3]

3.2

档案数字资源 digital archival resource

以数字形式存在的各类档案信息资源。

注:包括电子档案及其元数据、档案目录数据、传统载体档案数字化成果等。

[来源:DA/T 99—2024,3.1,有修改]

3.3

声纹 voiceprint

人的语音中所蕴含的、能用以表征和标识数据主体的生物学特性和行为特性的总称。

[来源:GB/T 41807—2022,3.3]

3.4

时间码 time code

对语音信号中出现的单词、短语或句子等单元进行时间标记的信息标识符。

3.5

语音识别 speech recognition

将人类的声音信号转化为文字,同时进行声纹(3.3)等特征提取的过程。

[来源:GB/T 21023—2007,3.1,有修改]

3.6

转写　transcription

将语音识别(3.5)结果进行书面文本规整、翻译及特征标记的过程。

[来源:GB/T 4894—2024,3.1.6.10,有修改]

3.7

语音识别转写成果　speech recognition and transcription result

语音信号经语音识别(3.5)转写(3.6)后形成的文件。

3.8

识别转写准确率　recognition and transcription accuracy

通过语音识别(3.5)转写(3.6)技术能够正确识别出的字符数占实际语音内容中应识别字符总数的百分比。

注:识别转写准确率=(识别正确字符数/应识别字符总数)×100%。

3.9

识别转写速度　recognition and transcription speed

单位时间长度录音录像档案通过语音识别(3.5)转写(3.6)获取语音识别转写成果(3.7)所需要的时间。

3.10

音素　speech

最小语音单位。

[来源:GB/T 3947—1996,10.3]

3.11

声学模型　acoustic model

将输入的语音信号转换为对应的音素(3.10)或音节序列的模型。

3.12

语言模型　language model

将声学模型(3.11)输出的音素(3.10)或音节序列转换为自然语言文本的模型。

3.13

热词集　hot words

由用户或应用程序定义的结构化关键词集合。

[来源:GB/T 34083—2017,3.6]

4　总体原则和要求

4.1　语音识别转写工作应纳入档案信息化建设范畴,统筹规划、有序推进。

4.2　语音识别转写工作宜根据录音录像档案的珍贵程度、开放程度、利用率、亟待抢救程度等因素科学实施,确定为语音识别转写对象的录音录像档案应全部识别转写。

4.3　语音识别转写成果应与录音录像档案建立一一对应关系,保持语义信息一致。

4.4　语音识别转写工作应采取可靠的技术和措施,确保工作过程规范、成果可靠、数据安全,并捕获相关元数据。语音识别转写工作实行服务外包的,应按 DA/T 68.2 的要求执行。

4.5　语音识别转写工作宜以有利于档案智能检索、数据挖掘、开发利用为原则。

4.6　语音识别转写工作宜重视涉密档案管理、知识产权保护、个人信息保护的问题。

5 组织与管理

5.1 工作组织

应建立语音识别转写工作组织，对组织实施、协调管理、安全保障、保密要求、技术支持、监督检查、成果验收等进行统筹规划，确保语音识别转写工作顺利开展。

5.2 工作人员

应配备熟悉档案业务且具备语音识别转写工作基础知识的工作人员，并明确岗位职责和保密要求。

5.3 工作场地

应设置专门语音识别转写工作场地，按 DA/T 62、DA/T 68.2 的要求，结合语音识别转写工作需要，合理布局档案存放、评估与导入、识别转写、成果修改校正、成果整理输出等工作区域。

5.4 管理制度

应制定并实施语音识别转写工作管理制度，包括岗位管理、人员管理、场地管理、设备管理、工作文件管理、存储载体管理、数据管理、质量管理、进度管理、安全管理、保密管理。语音识别转写工作管理制度宜科学、规范，有效保障档案信息安全和成果质量。

5.5 工作流程

语音识别转写工作流程包括方案制定、档案出库、评估与导入、识别转写、成果修改校正、成果整理输出、质量控制、成果挂接、成果验收、档案入库。录音录像档案语音识别转写工作流程见附录 A。

5.6 工作文件

语音识别转写工作文件主要包括工作方案、技术方案、审批书、流程单、验收报告、成果移交清单等，采取服务外包的还应至少包括项目招标文件、投标文件、中标通知书、项目合同、保密协议。录音录像档案语音识别转写工作流程单见附录 B。

5.7 元数据

语音识别转写过程中应完整捕获元数据，元数据项的设置应符合 DA/T 63 的规定，此外还应包括但不限于转写时间长度、转写环境、转写引擎、转写时间、音频信噪比、说话人人名、语种、成果文件名称、成果文件大小、成果文件格式。

6 方案制定

6.1 制定工作方案

6.1.1 依据语音识别转写工作的计划、目标等有关要求，在充分调研的基础上制定工作方案。

6.1.2 语音识别转写工作方案内容包括语音识别转写对象、工作目标、工作内容、经费核算、工作流程、技术方法、人员安排、验收依据、安全管理、保密管理等。

6.1.3 宜对工作方案进行专家论证，确保其科学、规范、合理。

6.1.4 工作方案应经审批后严格执行。

6.2 制定技术方案

6.2.1 应根据语音识别转写工作方案制定语音识别转写过程中各系统技术方案，包含语音识别转写引擎、语音识别转写软件、基础设施与网络等，所涉及的系统应满足如下条件：

——与管理录音录像档案的信息管理系统具备相同的安全保护等级或涉密保护等级；

——系统接口、数据格式及技术服务框架具有可兼容性、可扩展性；

——对其他系统、网络的功能和性能无明显影响；

——具备高安全性、可靠性，能在国产软硬件环境下稳定运行。

6.2.2 装备的语音识别转写引擎，能对录音录像档案进行高准确率和高速度语音识别转写，具备如下能力：

——支持语音识别转写文本、翻译文本、摘要文本及声纹的提取；

——具备对噪声的强抵抗能力，能有效屏蔽较大程度的噪声干扰；

——具备对异常语音方式（口吃、哼唱、口语化、复述等）的处理能力；

——能准确判别干扰词（停顿词、语气词、重复词等）。

6.2.3 装备的语音识别转写软件，能高速、准确输出语音识别转写成果，实现以下功能：

——对录音录像档案进行导入及预处理，录像档案自动提取音频；

——识别判断语种及支持手动选择语音识别转写引擎；

——自动语音识别转写；

——语音识别转写文本的处理及自动校验；

——输出语音识别转写成果，文本成果支持人工校正，支持热词集管理；

——数据阅览及导出，转写文本成果及翻译成果支持与档案同步对应阅览；

——说话人声纹识别、说话人标记、说话人声纹管理。

6.2.4 装备的基础设施与网络，包括引擎运行服务器、软件应用服务器、工作计算机、有线网络等，应支持系统正常运行。

7 语音识别转写的实施

7.1 档案出库

7.1.1 按照录音录像档案语音识别转写工作方案确定语音识别转写对象，对象包括原生录音录像电子档案和数字化形成的数字音视频文件。

7.1.2 根据语音识别转写对象开展档案数字资源调取、清点、登记等工作。档案载体出库的，按照本单位的档案库房管理规定执行；从电子档案管理信息系统调取的，按照本单位的信息系统管理相关规定执行。

7.2 评估与导入

7.2.1 语音识别转写实施前，应对语音识别转写对象的质量进行评估。评估内容一般包括语言信息、格式、编码格式、采样率、量化位数及声道等，宜满足以下各项要求：

——语言信息：以汉语、常见少数民族语言及国际通用语言录制；

——录音档案格式：WAV、MP3 等格式；

——录像档案格式：AVI、MPEG、MP4、FLV、MXF 等格式；

——编码格式：视频 H.264，音频 PCM；

——音频采样率：不低于 16 kHz，宜为 44.1 kHz；

——音频量化位数：不低于 16 bit；

——声道：单声道。

7.2.2 评估通过后，应将符合语音识别转写要求的录音录像档案导入语音识别转写软件。对评估不满足的，应采用技术手段处理，以满足 7.2.1 的要求。

7.3 识别转写

7.3.1 预处理

7.3.1.1 语音识别转写处理前，应根据音频特征，对输入的语音进行信号增强、噪声抑制、混响消除处理。

7.3.1.2 应对音频流分析检测，划分音频信号中语音信号和非语音信号（无声段或背景噪声），确定有效语音起止点，提取语音信号部分。

7.3.1.3 应对声学特征分析提取，自动检测筛选高抗噪性和强区分度的有效特征，特征参数经归一化处理后提取，以提升语音识别转写的准确性和稳定性。

7.3.2 识别

7.3.2.1 语音识别转写时，应支持自动检测或手动选择待识别音频的语种，确定语音识别转写引擎类型。

7.3.2.2 应通过声学模型和语言模型的联合建模完成音频识别。

7.3.2.3 声学模型应对语音信号自动抽取特征信息，并映射到语音的音素或词汇上，识别为初步文本。

7.3.2.4 语言模型应对语音信号自动语义分析，结合上下文信息推测出最可能的文本结果。应支持通过系统自动学习或人工自定义方式补充热词集，优化识别效果，以提升文本的生成准确率。

7.3.2.5 语言模型应对语音信号中不同语音片段进行声纹识别，根据识别后的声纹区分说话人和完成声纹注册，声纹识别注册方法见 SJ/T 11380 的规定。

7.3.3 转写

7.3.3.1 应按符合阅读习惯的格式处理转写文本，支持自动或手动语法修正，并对齐时间码。

7.3.3.2 分句与分段。应按语义自动将转写文本切分成若干个子句与段落，并添加标点。

7.3.3.3 文本顺滑。应自动过滤转写文本中无意义的字或词，如停顿词、语气词、非强调性重复词等。

7.3.3.4 非汉字字符转换。应自动按照常用书面表达方式转换转写文本中的非汉字字符，设置数字、字母、日期、货币、号码等不同场景的表达方式。

7.3.3.5 说话人标记。宜自动按照语音片段标记转写文本中对应的说话人。

7.3.3.6 文本翻译。宜自动对少数民族语言及国际通用语言转写文本进行简体中文翻译及规范化整理。

7.3.3.7 文本摘要。宜自动从转写文本中提取关键信息，归纳为简短的内容概述。

7.4 成果修改校正

7.4.1 应自动对转写文本进行语义识别和校正，通过热词集对转写文本中的特殊词汇及语句自动分析更正，如人名、地名、机构名、专有名词、多音字、同音不同字等。热词集应具备手动更新、自动更新和自动学习功能。

7.4.2 转写文本应与音频、视频同步阅览，为实现更高识别转写准确率，应以人工方式进行对照、修正。

7.4.3 宜对语音识别转写成果按说话人分离、聚类同步阅览，以实现快速校正。

7.4.4 宜对语音识别转写成果文本摘要内容进行手动调节或重新自动生成，以提高摘要的准确性。

7.5 成果整理输出

7.5.1 成果整理

7.5.1.1 语音识别转写成果主要包括纯文本文件、声纹音频文件、视频字幕文件等。

7.5.1.2 语音识别转写成果应记录不同说话人时间码，应与档案中说话人说话时间起止点一致，并支持同步阅览。录像档案语音识别转写成果，还应支持转写文本以字幕形式同步阅览，成果含翻译文本的应支持双语字幕。

7.5.1.3 语音识别转写成果应能被调用、编辑、导出、搜索。

7.5.2 成果输出

7.5.2.1 录音档案语音识别转写成果应保存为纯文本文件、声纹音频文件。录像档案语音识别转写成果应保存为纯文本文件、声纹音频文件、视频字幕文件。

7.5.2.2 应以录音录像档案的件为单位输出并保存语音识别转写成果，语音识别转写成果保存规则应符合表1的规定。

表1 语音识别转写成果保存规则

档案类型	录音录像档案语音识别转写成果	成果文件类型代码	成果文件格式
录音档案	纯文本文件(转写文本、翻译文本、摘要文本)	WB	TXT文件/XML文件/JSON文件
	声纹音频文件(按说话人)	SW	PCM文件
录像档案	纯文本文件(转写文本、翻译文本、摘要文本)	WB	TXT文件/XML文件/JSON文件
	声纹音频文件(按说话人)	SW	PCM文件
	视频字幕文件(转写文本、翻译文本)	ZM	SRT文件

7.5.2.3 应以档号为基础对语音识别转写成果命名，应确保语音识别转写成果命名的唯一性。一件档案形成多个语音识别转写成果时，应按档号结合成果文件类型代码和顺序流水号为语音识别转写成果命名。

示例1：档号为X001-LY·2025·D30-0001的单人录音档案，对应的语音识别转写成果文件名为X001LY2025D300001-WB01.txt和X001LY2025D300001-SW01.pcm。

示例2：档号为X001-LX·2025·Y-0001的两人录像档案，对应的语音识别转写成果文件名为X001LX2025Y0001-WB01.txt、X001LX2025Y0001-SW01.pcm、X001LX2025Y0001-SW02.pcm和X001LX2025Y0001-ZM01.srt。

7.5.2.4 应保持档案数字资源中时间码与语音识别转写成果中时间码的一致性，方便对档案的快速检索定位阅览。

7.5.2.5 应按DA/T 18和DA/T 63的要求，自动保存相关元数据。相关元数据应纳入电子档案管理信息系统中统一管理。

7.6 质量控制

7.6.1 质量控制要求

7.6.1.1 应结合工作方案中目标要求，对语音识别转写过程，包括评估与导入、识别转写、成果修改校正、成果整理输出等业务环节进行质量控制，并形成质量评价意见。

7.6.1.2 质量控制指标主要为识别转写准确率和识别转写速度，识别转写准确率指标与识别转写速度指标应同时适用。

7.6.2 识别转写准确率

7.6.2.1 语音识别转写对以汉语、常见少数民族语言及国际通用语言录制，声音质量好、说话人发音标准的识别转写准确率应不低于95%；以标准汉语普通话录制，识别转写准确率应不低于98%。

7.6.2.2 语音识别转写对声音质量差、杂音多、说话人发音不标准的识别转写准确率应不低于80%。

7.6.3 识别转写速度

应控制语音识别转写速度，1 h录音档案不超过10 min完成语音识别转写，1 h录像档案不超过30 min完成语音识别转写。

7.7 成果挂接

7.7.1 应对目录数据和与其对应的语音识别转写成果进行挂接，以实现目录数据与成果的关联。

7.7.2 应逐条检查挂接结果，包括目录数据与语音识别转写成果对应的准确性，已挂接语音识别转写成果与实际语音识别转写的档案数字资源数量的一致性、语音识别转写成果是否能正常打开，发现错误应及时纠正。

7.8 成果验收

7.8.1 应采用计算机自动检验与人工检验相结合的方式对语音识别转写成果进行验收检验。

7.8.2 验收检验内容包括语音识别转写文本、翻译文本、摘要文本、声纹音频文件、视频字幕文件、说话人标记情况、语音文字同步阅览情况、成果挂接情况、元数据、工作文件和存储载体等。

7.8.3 能采用更换语音识别转写引擎再次进行语音识别转写，两次成果自动比对检验的，应采用计算机自动检验方式进行100%检验；无法用计算机自动检验的，应采用人工抽检，被抽检录音录像档案件数及时间长度总和均不应低于录音录像档案的5%。

7.9 档案入库

7.9.1 按照本单位的录音录像档案入库相关要求，对完成语音识别转写的档案数字资源及语音识别转写成果进行处理和清点，并履行档案入库手续。档案载体入库的，按照本单位的档案库房管理规定执行；导入电子档案管理信息系统的，按照本单位的信息系统管理相关规定执行。

7.9.2 采用服务外包的，应对语音识别转写工作中使用的各种硬件设备和软件系统进行检查，以确保其无信息留存。凡存有信息的，应作清除信息的安全处理。服务外包验收后，外包方自带设备的硬盘予以拆除并与语音识别转写过程中使用过的其他移动存储介质一并予以接收。

8 语音识别转写成果管理与应用

8.1 成果管理

8.1.1 录音录像档案数字资源为电子档案的，语音识别转写成果应与录音录像电子档案分开存放，在存储载体或电子档案管理信息系统中以档号为基础逐级建立层次文件夹单独保存，两者之间建立关联关系。

8.1.2 录音录像档案数字资源为数字化成果的，语音识别转写成果宜与录音录像档案数字化成果统一保存，并保持两者之间的逻辑层次和关联关系。

8.1.3 应建立语音识别转写成果权限管理机制，与录音录像档案相同。

8.1.4 宜重视语音识别转写成果中声纹音频文件存储、管理、传输的合规问题。

8.1.5 语音识别转写成果应与录音录像档案数字资源同步开展数据备份、长期保存等工作。

8.2 成果应用

8.2.1 应建立语音识别转写成果索引库并与录音录像档案数字资源对应，通过电子档案管理信息系统实现全文检索，支持档案原文溯源查询及文本摘要信息的快速阅览。

8.2.2 可基于纯文本文件，结合数据挖掘技术，开展关键信息提取，辅助档案自动著录及编目等工作。

8.2.3 可基于声纹音频文件，开展说话人信息维护及检索等应用工作。

8.2.4 可利用语音识别转写成果，结合人工智能技术，开展知识管理、内容推荐、展览展陈、辅助开放审核、辅助档案编研等工作。

附　录　A
（资料性）
录音录像档案语音识别转写工作流程

图 A.1 给出了录音录像档案语音识别转写工作流程。

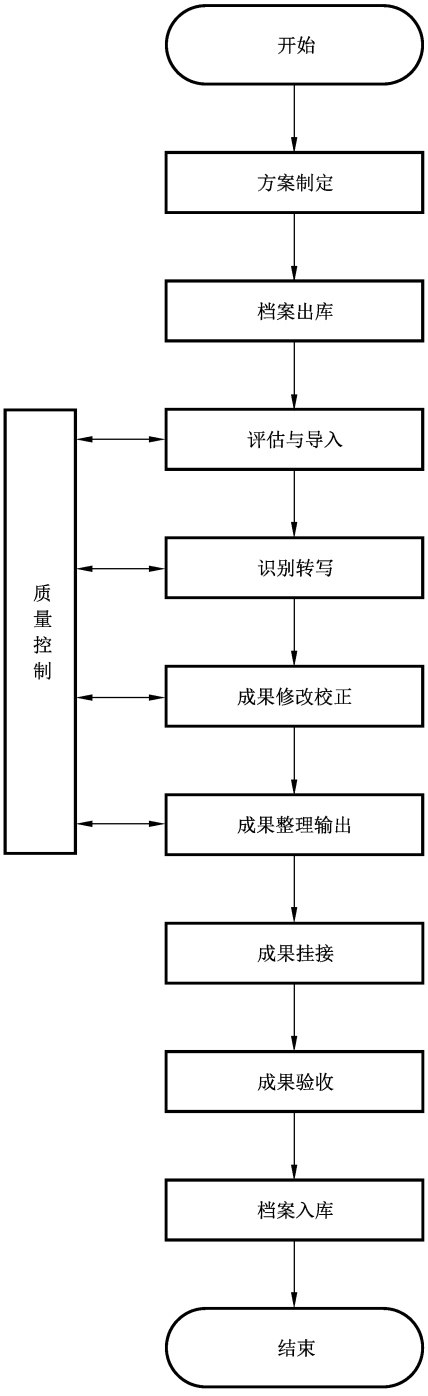

图 A.1　录音录像档案语音识别转写工作流程

附　录　B
（资料性）
录音录像档案语音识别转写工作流程单

表 B.1 给出了录音录像档案语音识别转写工作流程单的参考样式。

表 B.1　录音录像档案语音识别转写工作流程单

档号：　　档案类型：□录音档案　□录像档案

评估与导入	评估情况： 导入情况： 操作人： 年　　月　　日
识别转写	预处理情况：□完成　□未完成 识别情况：□完成　□未完成 转写情况： □分句与分段　□文本顺滑　□非汉字字符转换　□说话人标记　□文本翻译　□文本摘要 操作人： 年　　月　　日
成果修改校正	成果修改校正情况： 操作人： 年　　月　　日
成果整理输出	成果整理情况： 操作人： 年　　月　　日
	成果输出情况： 纯文本文件(数量)_____声纹音频文件(数量)_____视频字幕文件(数量)_____ 操作人： 年　　月　　日
质量控制	质量评价意见： 识别转写准确率： 识别转写速度： 操作人： 年　　月　　日
语音识别转写处理人：　　完成时间： 质检人：　　完成时间：	

注：预处理情况与识别情况根据软件反馈结果登记。

参 考 文 献

[1] GB/T 3947—1996 声学名词术语
[2] GB/T 4894—2024 信息与文献 基础和术语
[3] GB/T 18894—2016 电子文件归档与电子档案管理规范
[4] GB/T 21023—2007 中文语音识别系统通用技术规范
[5] GB/T 34083—2017 中文语音识别互联网服务接口规范
[6] GB/T 36464.1—2020 信息技术 智能语音交互系统 第1部分:通用规范
[7] GB/T 37036.5—2023 信息技术 移动设备生物特征识别 第5部分:声纹
[8] GB/T 41807—2022 信息安全技术 声纹识别数据安全要求
[9] GB/T 41813.1—2022 信息技术 智能语音交互测试方法 第1部分:语音识别
[10] DA/T 31—2017 纸质档案数字化规范
[11] DA/T 77—2019 纸质档案数字复制件光学字符识别(OCR)工作规范
[12] DA/T 78—2019 录音录像档案管理规范
[13] DA/T 99—2024 档案数字资源备份实施规范
[14] SJ/T 11380 自动声纹识别(说话人识别)技术规范
[15] 国家档案局办公室关于印发《档案数字化外包安全管理规范》的通知(档办发〔2014〕7号)