

# 老年人中文语音数据库

**AISHELL-ASR0060**

产品说明书

北京希尔贝壳科技有限公司

**Beijing Shell Shell Technology Co.,Ltd**

Add: Room 3-621, 6F, Zhongguancun Lifangting No. 1, Shanyuan Road, Haidian District, Beijing 100080, P.R.China

Tel: +86 10 80225006 E-mail: [bd@aishelldata.com](mailto:bd@aishelldata.com)

## 目录

1 产品概述.....	3
2 产品目录结构.....	3
2.1 目录结构.....	3
2.2 命名规则.....	3
3 文本设计.....	4
3.1 语料制作.....	4
3.2 文本结构.....	4
4 数据录制环境.....	5
4.1 录制环境.....	5
4.2 录制设备.....	5
4.3 录制方法.....	5
5 标注转写规范.....	6
6 发音人信息.....	7
7 版权声明.....	7

# 1 产品概述

AISHELL-ASR0060 老年人中文语音数据库共 780 小时。录音语言，中文；录音地区，中国。录音文本包含智能控制词、POI 等 12 个分类。邀请 503 名 50 岁以上老年人参与录制。录制过程在安静室内环境中，同时使用 3 种不同设备：高保真麦克风（44.1kHz，16bit，260H）；Android 系统手机（16kHz，16bit，260H）；iOS 系统手机（16kHz，16bit，260H）。

此数据库经过专业语音校对人员转写标注，并通过严格质量检验，文本正确率在 97% 以上。

## 2 产品目录结构

### 2.1 目录结构

数据目录结构	
<b>数据目录结构</b>	
AISHELL-ASR0060.pdf	(数据库简介)
└─DOC	(文本说明文件)
├─wav_list.txt	(音频列表)
├─content.txt	(语音文本内容)
├─spk_info.xlsx	(录音人信息)
└─SPEECHDATA	(数据文件夹)
├─0001	(录音人文件夹)
0001_Android_0001.wav	(音频文件)
0001.txt	(语音文本内容)

图表 2-1-1 数据目录结构

### 2.2 命名规则

CORPUS/USAGE/SPEAKER\_NUM/SPEECH\_ID

e.g.AISHELL-ASR0060/SPEECHDATA/0001/0001\_Android\_0001.wav

目录名称	内容	备注
CORPUS	AISHELL-ASR0060	语音数据库编号

USAGE	SPEECHDATA	文件夹名称
SPEAKER_NUM	0001	录音人文件夹名称
SPEECH_ID	0001_Android_0001.wav	WAV 文件

图表 2-2-1 命名规则

## 3 文本设计

### 3.1 语料制作

语料在 12 个领域中选定。按照规则处理语料池。

序号	领域
1	家居家电名称和控制命令
2	POI (地理信息)
3	音乐、收音机类 (语控)
4	数字类 (语控)
5	电视、电影名称
6	财经
7	科技
8	体育
9	娱乐
10	时事新闻
11	英文拼读
12	广域文本

图表 3-1-1 语料池内容

- 脱敏处理。删除政治敏感、个人隐私、色情暴力等内容。
- 删除 <, >, [, ], ~, /, \, = 等符号。
- 删除含有中文和英文以外语言的内容。
- 删除单句含有 25 字以上的内容。
- 统一格式。

### 3.2 文本结构

考虑到语音覆盖及音素平衡，此数据库录音文本采用每份 500 句的分配方式设计，从语料池中抽取，结构如下。

序号	领域	句数
----	----	----

1	家居家电名称和控制命令	10
2	POI (地理信息)	63
3	音乐、收音机类 (语控)	56
4	数字类 (语控)	34
5	电视、电影名称	55
6	财经	52
7	科技	47
8	体育	45
9	娱乐	38
10	时事新闻	57
11	英文拼读	5
12	广域文本	38
合计	12 项	500

图表 3-2-1 文本结构

## 4 数据录制环境

### 4.1 录制环境

安静室内，不包括明显的其他人说话声音及其他噪音，无回音。发言人按照正常语速，朗读录音文本。

### 4.2 录制设备

录制设备包括高保真麦克风和录音机、手机。本数据库数据存储格式为高保真录制数据 44.1kHz、16bit 单声道和手机录制数据 16kHz、16bit 单声道格式。

### 4.3 录制方法

发音人距离高保真麦克风 20 厘米，以讲话正常音量，正常语速，朗读录音文本。Android 系统手机与 iOS 系统手机分别与麦克风间隔 20 厘米布置。



图表 4-3 录制示意图

## 5 标注转写规范

数据转写人员根据所听到的音频写出内容，力求使文本内容与音频发音内容保持一致。

准则如下：

- 1) 转写的内容必须和听到的语音完全一致，不能多字、少字、错字。
- 2) 数字要转写为汉字形式，如“一二三”，而不是“123”。注意区分“一”和“幺”，“二”和“两”。
- 3) 句中出现的英文按照发音写出单词，如“thank you”。按拼读朗读的字母，需转写成大写字母加空格的形式。如，“NBA”、“UFO”。
- 4) 句中包含的符号，按实际发言人发音转写。如“三 W 点 百度 点 com”。没有发音的符号，需要删掉。品牌名称，专有名称等按照实际惯用格式转写，如“QQ 空间”、“iPhone”、“喜马拉雅”。
- 5) 标注内容的完整性要与实际发音一致，不得删减。

## 6 发音人信息

发音人信息记录内容包括任务编号、性别、口音区域、年龄。

任务编号	性别	口音区域	年龄区间
0001	男	北方	58

图表 6-1-1 基本信息表例

**任务编号：**每个发言人领取 1 个任务编号，每个任务编号对应 1 份录音文本。每个发言人只能参加一次录制。

**性别：**男性 239 人，女性 264 人。

**口音区域：**按照发言人原生语言所属区域，分为北方、南方、其他。

**年龄区间：**发言人参加录音时的年龄，此数据库发言人年龄均为 55 岁以上。

## 7 版权声明

本文内容禁止转载，AISHELL(北京希尔贝壳科技有限公司)对本文拥有修改权、更新权及最终解释权。



**希尔贝壳**  
A I S H E L L

**Copyright**