

电话通信业务场景语音数据库（测试集）



北京希尔贝壳科技有限公司

Beijing Shell Shell Technology Co., Ltd.

Add: Room 813, Building No. 4, Shangdi 10th Street, Haidian District, Beijing 100080, P.R.China

Tel: 010-80225006 E-mail: bd@aishelldata.com

目录

1 产品概述.....	2
2 录音语料.....	2
3 发音人信息.....	3
3.1 口音区域	3
3.2 性别比例	3
3.2 年龄比例	3
4 数据录制环境.....	4
4.1 录制环境.....	4
4.2 录制设备.....	4
4.3 录制方法.....	4
5 语音数据转写.....	4
6 数据文件目录.....	5
7 版权声明.....	5



1 产品概述

AISHELL-2021D-EVAL 是 AISHELL-ASR0006 的子库，共 1.95 小时。AISHELL-ASR0006 口音对话语音数据库共 1400 小时，录音语言为中文，共邀请 1440 名来自中国不同口音区域发音人参与录制。录制过程为安静室内环境，录音人通过手机，以自然对话形式进行录制。数据存储格式为 8000Hz，16bit。

AISHELL-2021D-EVAL 在电商零售领域，针对模拟电信业务场景，随机抽取 5 段自然对话，共 6 名发音人。

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，文本正确率达到 100%。

2 录音语料

AISHELL-ASR0006 数据库为电话自然对话方式，录音人对话内容领域覆盖电商、教育等常见领域，具体结构如下。AISHELL-2021D-EVAL 抽取的是电商领域内容。

领域	内容	占比
电商、零售	退换货、推销	10
教育	课程咨询	10
生活服务	餐饮、旅游、娱乐、交通	10
汽车销售	买卖、维修、车险	10
房地产	房屋买卖、租赁	5
政府	线上政务咨询	5
物流	快递查询	5
其他	自由发挥	40
合计		100

图表 2-1 文本结构

3 发音人信息

3.1 口音区域

数据库录音人覆盖七大方言区，包括北方方言、吴方言、赣方言、湘方言、闽方言、粤方言以及客家方言。具体结构比例见下图。

AISHELL-2021D-EVAL 数据集抽取北方方言。

口音区	人数	占比 (%)
北方方言	218	15.1
吴方言	186	12.9
赣方言	263	18.3
湘方言	249	17.3
闽方言	174	12.1
粤方言	183	12.7
客家方言	167	11.6
合计	1440	100

图表 3-1-1 文本结构

3.2 性别比例

数据库总人数为 1440 人，男 604 人，女 836 人。AISHELL-2021D-EVAL 数据集为男 1 人，女 5 人。

性别	男性	女性	合计
比例	41.9	58.1	100%

图表 3-2-1

3.2 年龄比例

录音人年龄覆盖 18-60 岁，具体分布如下所示：A（18-25 岁）；B（26-40 岁）；C（41 岁以上）。AISHELL-2021D-EVAL 数据集的录音人年龄在 21-31 岁之间分布。

	年龄段	人数	比例 (%)
A	18-25	840	58.33
B	26-40	400	27.78
C	>41	200	13.89
合计		1440	100

图表 3-3-1

4 数据录制环境

4.1 录制环境

安静室内，无回音，无其他人说话。

4.2 录制设备

市场上常见的 Android 系统手机，如华为、oppo、vivo、小米等。数据存储格式为 16kHz & 8kHz, 16bit wav 文件。

4.3 录制方法

录音人在安静室内通过手机自然通话录制。

5 语音数据转写

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。一般准则如下：

- 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) 音频中有英文发音的应写成相应的汉字或英文。具体分为以下几种情况：

网址中包含的所有的字母或单词，均为大写。例如：发音内容为“www.abc.com”，应转写为“三 W 点 A B C 点 com”。

发音中包含的英文单词，转写时全部为小写。

发音中包含的英文字母，转写时全部为大写。

对于一些大写专有名词，或者一些英文缩写全部大写加空格，例如：CEO、CCTV 等。

4) 标注内容的完整性要与实际发音一致，不得删减。

5) AISHELL-ASR0006 数据库中，句中包含方言字词，不改变句意，用同音字代替，字前加<dialect>标记。

6) AISHELL-ASR0006 数据库中，如果两个人说话有重叠，只做主说话人内容转写，并同时加入<overlap>标记。无论是连续性重叠还是多字间断性重叠，都在每个重叠字处标记<overlap>。如下所示：

例 1：甲说：“今天的天气好热呀！”话还没完，乙说：“嗯。”，“嗯”字正好跟“热”字重叠了，且两个人声音大小差不多，则转写为“今天的天气好热<overlap>呀！”。

例 2：甲说：“我觉得是这样的，嗯。”，乙说：“我也觉得是。”，乙在甲没有说完“嗯”的时候说的话，覆盖了甲的“嗯”，则转写为“我觉得是这样的。”和“<overlap>我也觉得是。”。

6 数据文件目录

数据目录树	
数据目录结构	
AISHELL-2021D-EVAL.pdf	(数据库简介)
└─DOC	(文本说明文件夹)
├─wav_list.txt	(音频列表)
├─content.txt	(标注文本)
└─SPEECHDATA	(数据文件夹)
├─00004	(通话编号)
├─0004_00001.wav	(音频文件)

图表 6-1-1 数据目录结构

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



希尔贝壳
A I S H E L L

Copyright