

## 标贝科技数据产品说明书

### 【中文女声 DB-30 音库】

# 数据产品概述

## 数据产品介绍

语音合成是通过机械的、电子的方法产生人造语音的技术。TTS 技术（又称文语转换技术）隶属于语音合成，它是将计算机自己产生的、或外部输入的文字信息转变为可以听得懂的、流利的口语输出的技术。

TTS 语音合成技术是实现人机语音通信关键技术之一。使电脑具有类似于人一样的说话能力，是当今时代信息产业的重要竞争市场。和语音识别 ASR 相比，语音合成的技术相对来说要成熟一些，是应用范围较广的技术。

随着人工智能产业的飞速发展，语音合成系统也得到了更加广泛的应用。除了语音合成初期的清晰度、可懂度以外，人们对语音合成的自然度、节奏感以及音质的要求也越来越高。而语音库的质量也是决定语音合成效果的关键因素。

【中文女声 DB-30】采集对象为中文女声，拟人口语对话风格。录制环境为专业录音室和录音软件，录音环境和设备自始至终保持不变，录音环境的信噪比不低于 35dB;单声道录音，用 48KHz 16 比特采样频率、PCM WAV 格式。录音采用真实对话及自然闲聊方式录制，录音风格自然。根据合成语音标注标准，依据发音人实际发音对音库进行文本音字标注、韵律层级标注、口语化行为标签标注。

## 数据应用场景

本语音数据库应用广泛，可用于以下方面：

1. 科研，可用于语音合成系统模型训练；
2. 生活出行；
3. 智能科技；
4. 教育、娱乐等。

# 数据产品详情

## 技术参数

数据规格	
数据内容	中文女声 DB-30
录音语料	综合多领域语料样本；音节音子的数量、类型、音调、音连以及韵律等进行覆盖。
有效时长	<p>总 <b>30.24 小时</b></p> <p>【基础对话 13.3H+自然闲聊 4.49H+中英混 2H+8 种情感 3.18H（厌、喜、安慰鼓励、崇拜、怒、悲、惊、撒娇卖萌）+个性化播报 4.83H+自然情感 2.44H（喜、强喜、怒、强怒、哀、强哀、害怕）】</p>
语言类型	中文普通话
发音人	中文女声，20-30 岁
录音环境	<p>声音采集环境为专业录音棚环境</p> <ol style="list-style-type: none"><li>1) 录音棚符合专业音库录制标准；</li><li>2) 录音环境和设备自始至终保持不变；</li><li>3) 录音环境的信噪比不低于 35dB。</li></ol>
录制工具	专业录音设备及录音软件
采样格式	无压缩 PCM WAV 格式，采样率为 48KHz 16bit。
标注内容	音字标注、韵律标注、口语化行为标签标注（不含中英混）
标注格式	音字韵律文本及口语化特征标签标注为.txt 格式文档；语气词情感标注为.interval 格式文档

质量标准	1. 语音文件为 48KHz 格式，音色、音量、语速一致，无截幅； 2. 标注文件字准确率不低于 98%； 3. 韵律无明显不合理的标注错误。
存储方式	FTP 存储
文件格式	音频文件：WAV 文本及口语化特征标签标注文件：TXT 语气词情感标注文件：interval
数据脱敏	
数据敏感项	无
产品应用	
适用领域	科研、智能科技、生活、教育、娱乐等多领域应用
版权归属	
版权所有者	标贝（青岛）科技有限公司

## 数据层次结构

### 数据目录树

#### 数据目录结构

- | 中文女声 DB-30
  - | | 基础对话
    - | | | \*.wav （音频文件）
    - | | | \*.txt （标注文本文件）
    - | | | \*.txt （口语化特征标签标注文件）
    - | | | \*.interval （语气词情感标注文件）
  - | | 闲聊
    - | | | \*.wav （音频文件）

| | | \*.txt (标注文本文件)  
 | | | \*.txt (口语化特征标签标注文件)  
 | | | \*.interval (语气词情感标注文件)  
 | | 中英混  
 | | | \*.wav (音频文件)  
 | | | \*.txt (标注文本文件)  
 | | 多情感  
 | | | \*.wav (音频文件)  
 | | | \*.txt (标注文本文件)  
 | | | \*.txt (口语化特征标签标注文件)  
 | | | \*.interval (语气词标注文件)  
 | | 个性化播报  
 | | | \*.wav (音频文件)  
 | | | \*.txt (标注文本文件)  
 | | | \*.txt (口语化特征标签标注文件)  
 | | | \*.interval (语气词情感标注文件)  
 | | 自然情感  
 | | | \*.wav (音频文件)  
 | | | \*.txt (标注文本文件)  
 | | | \*.txt (口语化特征标签标注文件)  
 | | | \*.interval (语气词标注文件)

## 文件命名规则

内容	详细描述	格式
语音文件	录音单句文件	WAV 48k 16bit
文本文件	音字、韵律标注后的文本文件；口语化特征标签标注文件	TXT 文档
标签文件	语气词情感标注文件	Interval

例如：000001.wav

000001-010000.txt

000001.interval

## 标注文件内容格式

### 一、音字韵律标注：

## 1-1 中文标注格式

1.文本格式为 \*.txt，一行文字，一行拼音。文字行句首为句子号，句子号由六位半角阿拉伯数字组成，以 Tab 键隔开，后接文本内容，句尾以回车换行符结束；拼音行句首为 Tab 键，后接文本拼音，拼音之间以空格分开，句尾以回车换行符结束。

2.声调：声调以 1-6 标记，1-4 对应阴平、阳平、上声、去声，5 表示轻声，6 表示上声变调。

3.韵律标注：

中文韵律结构标注包含韵律词(#1)、韵律短语(#2)、语调短语(#3)、句末(#4)四个层级的标注。

例：

100001 该公司#1 当时#1 表示#3，将于#1 本周一#2 公布#1 正式#1 消息#4。

gai1 gong1 si1 dang1 shi2 biao3 shi4 jiang1 yu2 ben3 zhou1 yi1 gong1 bu4 zheng4 shi4 xiao1 xi5

## 1-2 中英混合标注格式

1.文本格式为\*.txt，一行文字，一行拼音。文字行句首为句子号，句子号由六位半角阿拉伯数字组成，以 Tab 键隔开，后接文本内容，句尾以回车换行符结束；拼音行句首为 Tab 键，后接文本拼音，拼音之间以空格分开，句尾以回车换行符结束；英文单词之间以 “/” 分开，音素之间以空格分开,音节以"."分隔。

2.中文声调：声调以 1-6 标记，1-4 对应阴平、阳平、上声、去声，5 表示轻声，6 表示上声变调

3.英文重音：0 代表非重音，1 代表重音，2 代表次重音。

4.韵律标注：

中英混合文本，按中文韵律结构标注，包含韵律词(#1)、韵律短语(#2)、语调短语(#3)、句末(#4)四个层级的标注。

（注：一个英文单词划分为一个韵律词，以“空格”隔开）

例：

100001 阿凡达的#1 主题曲#2I See You#4。

a1 fan2 da2 de5 zhu3 ti2 qu3 / AY1 / S IY1 / Y UW1

## 二、口语化特征标签标注：

特殊口语化特征标签标注符号说明

- 1.主重音：@，标在文本行，主重音词后；
- 2.次重音：\$，标在文本行，次重音词后；
- 3.拖音：<r>，标在拼音行 字后；
- 4.非常规韵律停顿：/，体现在文本行；
- 5.语速加快：#，精确到每一个加快的字上，标在字后，加快的每个字都相对饱满；
- 6.连读：<>，把连读的词括起来，有吞音现象标注

## 三、语气词情感标注：

标注格式为\*.interval 格式；

第一行标注为语气词，以阿拉伯数字标注，阿拉伯数字代表不同种类语气词对应的编号（例如：1-停顿；2-疑问；3-应答；4-不同意或出乎意料；5-感叹.....）；

第二行为情感标注。



