声音复刻

快速入门

腾讯云

版权所有：腾讯云计算（北京）有限责任公司

【版权声明】

©2013-2025 腾讯云版权所有

本文档（含所有文字、数据、图片等内容）完整的著作权归腾讯云计算（北京）有限责任公司单独所有，未经腾讯云事先明确书面许可，任何主体不得以任何形式复制、修改、使用、抄袭、传播本文档全部或部分内容。前述行为构成对腾讯云著作权的侵犯，腾讯云将依法采取措施追究法律责任。

【商标声明】

腾讯云

及其它腾讯云服务相关的商标均为腾讯云计算（北京）有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标，依法由权利人所有。未经腾讯云及有关权利人书面许可，任何主体不得以任何方式对前述商标进行使用、复制、修改、传播、抄录等行为，否则将构成对腾讯云及有关权利人商标权的侵犯，腾讯云将依法采取措施追究法律责任。

【服务声明】

本文档意在向您介绍腾讯云全部或部分产品、服务的当时的相关概况，部分产品、服务的内容可能不时有所调整。您所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定，除非双方另有约定，否则，腾讯云对本文档内容不做任何明示或默示的承诺或保证。

【联系我们】

我们致力于为您提供个性化的售前购买咨询服务，及相应的技术售后服务，任何问题请联系 4009100100或95716。

文档目录

快速入门
操作指引
训练数据准备指引

快速入门
操作指引

最近更新时间：2024-08-23 12:04:11

登录控制台

注册并通过实名认证后，登录腾讯云控制台。如果没有账号，请参考 账号注册教程 。

开通服务

单击进入 声音复刻控制台 ，即可开通相应服务。

购买

声音复刻提供预付费计费模式，您可以查看声音复刻的 计费概述 。

使用声音复刻服务

您可以根据以下两种方式的介绍选择合适的使用方式：

使用方式	适用对象	说明	相关文档
通过 声音复刻控制台 进行功能体验	非业务开发者，没有代码编写基础	此方式只能用于体验，不能用于开发。	功能体验
通过 API 3.0 Explorer 在线调用功能	开发初学者，有代码编写基础	此方式能够实现在线调用、签名验证、快速检索接口等能力	API 概览

说明：
若您需调用 API 实现功能开发，请提前对 HTTP 请求和 API 调用进行了解。

查看调用

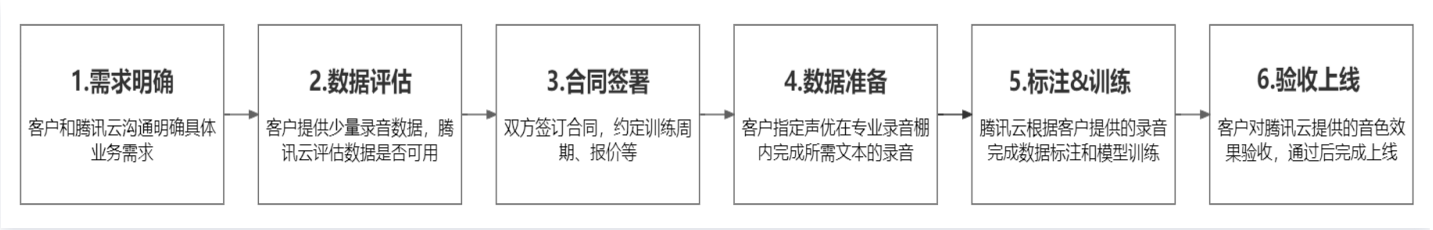
登录腾讯云 声音复刻控制台 ，查看声音复刻各服务的使用情况。

版权所有：腾讯云计算（北京）有限责任公司

训练数据准备指引

最近更新时间：2024-08-23 11:10:11

一、复刻流程



> **说明：**
> 轻量版声音复刻和基础版声音复刻流程参考 声音复刻控制台 。

二、训练数据

1. 数据准备前

【录音环境】建议使用和真实应用场景一致的环境进行录音采集，录音环境中避免出现回音、背景音乐、吵闹人声的情况；
【录音演员】定制方需拥有录音演员使用授权；
【录音设备】录音过程中，使用同一个录音设备，避免电流声等杂音。

2. 数据准备中

【录音质量】录制音频保证干净清晰，所有音频的情绪、语速、音量、音高一致稳定，避免语速忽快忽慢、音量忽大忽小；避免情绪表达过于夸张；朗读不出现卡顿、读错问题；避免方言口音朗读，避免音频包含口水音或较大的换气声；
【录音风格】按照正常业务场景对应的音色风格自然表达；
【音频格式】单/双声道16bit 固定采样率16k以上，44100、48000最优，音频非压缩保存，音频格式为无损音频wav格式；
【录音语种】中文、英文；场景包含英文播报的，需要单独提供中文音频、英文音频、中英文音频；
【数据确认】音频准备过程中，建议提供部分样例，由产品经理确认是否符合业务要求。

3. 数据准备后

【数据检查】数据文件提交给产品经理，进行数据检查，确保发音风格统一、数据质量可用。数据检查、筛选无误后，由产品经理给出声音定制排期。

版权所有：腾讯云计算（北京）有限责任公司

> **说明：**
> 轻量版声音复刻无需线下准备数据，基础版声音复刻数据准备后可直接在声音复刻控制台 创建训练任务。

三、标注&训练

标注及训练由腾讯完成，一般周期3-4周。

> **说明：**
> 轻量版声音复刻可30分钟内完成训练，基础版声音复刻可1-2天完成训练。

四、交付验收

【验收文本】客户提供20句验证文本模板，用于验证效果；
【模型优化】按报价包含的次数支持，超出需要额外报价付费，参考小规模定制报价。

可优化范围，以下为客观规则话播报，与定制模型无关，语音合成服务本身会持续优化：

a. 多音字播报，如「一行白鹭上青天」中的行（hang2）；
b. 特殊姓名，如「华」hua4；
c. 专有名词，如「腾讯」、「深圳大学」；
d. 专业术语，如「k8s」、「SLA」；
e. 时间，如「2022-3-21」、「19:01:02」；
f. 电话号码，13811227834、010-5114987；
g. 数学，3/4、0.18、23.78%；
h. 符号单位：「4:3」。

不可优化范围，以下为主观效果或个性化需求：

a. 情绪不可优化，需根据场景提前确认录音情绪，如金融催收、电商客服、武侠小说等；一种情绪为一个音色，如需提供多种情绪需要定制多个音色；
b. 语速不可优化，需要音频语速一致，tts接口通过SSM参数支持微调；
c. 音量不可优化，需要音频的音量一致，tts接口通过SSM参数支持微调；
d. 若音频采用和原始录音拼接，无法保证拼接效果，建议采用整句合成或全量合成；
e. 特定文本按特殊播报的，如 `sanzhang@tencent.com` 会按照拼音播报却要单个字母播报。

版权所有：腾讯云计算（北京）有限责任公司