

腾讯云智能语音服务

产品介绍

产品文档



腾讯云

【版权声明】

©2013-2017 腾讯云版权所有

本文档著作权归腾讯云单独所有，未经腾讯云事先书面许可，任何主体不得以任何形式复制、修改、抄袭、传播全部或部分本文档内容。

【商标声明】



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。

【服务声明】

本文档意在向客户介绍腾讯云全部或部分产品、服务的当时的整体概况，部分产品、服务的内容可能有所调整。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模式的承诺或保证。

文档目录

文档声明.....	2
产品介绍.....	4
产品简介	4
基本概念	6
语音识别	8
语音合成	9
声纹识别.....	11
客服机器人.....	12

产品介绍

产品简介

智能语音服务 (Artificial Audio Intelligence) 满足语音识别、语音合成、声纹识别等语音处理需求。智能语音服务拥有强大的垂直领域定制化服务，打造专业高效的语音大脑，为企业提供全方位的智能语音解决方案。

产品功能

语音识别

智能语音服务具有语音转文字，根据关键词列表搜索语音偏移量，情绪识别，静音监测等功能；提供声学 and 语言模型定制化服务。

语音合成

智能语音服务可以将文本转成自然清晰的语音，支持多种音色选择、语速选择，并支持中文、英文和中英文混读等。

声纹识别

提供语音一对一识别功能，可以用来鉴定说话人身份，在金融行业可用于用户身份鉴定，保障用户资金安全。

云端 + 嵌入式

云端服务可随需求弹性扩缩容，并提供持久化存储能力；嵌入式引擎满足无网络场景下智能语音需求。云端和嵌入式的结合，可以满足智能硬件和智能家居的实时语音交互需求，提升用户体验。

应用场景

语音质检

语音转文字，情绪识别，静音监测，关键词搜索等多种能力结合，提高通话合规性检查效率；语料信息隐藏的市场信息亦可提高企业产能。

语音搜索

针对垂直领域定制化的语音识别引擎，在联网或无网络环境下，提供便捷人机交互方式，提升产品体验。

智能客服

通过机器学习挖掘用户问题，免去多级语音菜单导航烦恼，通过机器人回答代替 50% 以上人力，降低运营成本。

身份鉴定

声纹识别技术高达 99% 的身份鉴定准确率，为众多领域提供更可靠的安全保证。

基本概念

智能语音服务涉及到对象存储以及语音识别的相关基本概念，为了帮助用户理解后续内容，将智能语音服务涉及的基本概念作如下说明。

1. APPID

腾讯云应用 ID。

获取途径：新用户注册并登录 [腾讯云管理中心控制台](#)，立即生成 APPID。

2. ProjectId

腾讯云项目 ID，默认项目 ID 为 0。

获取途径：用户登录 [控制台](#)，选择【用户中心】栏下的【项目管理】，可查看默认项目，默认项目 ProjectId = 0。单击【新建】即可创建新项目，并获取新的 ProjectId。

3. SecretId 和 SecretKey

腾讯云安全凭证，其中，SecretId 是用于标识 API 调用者身份的，而 SecretKey

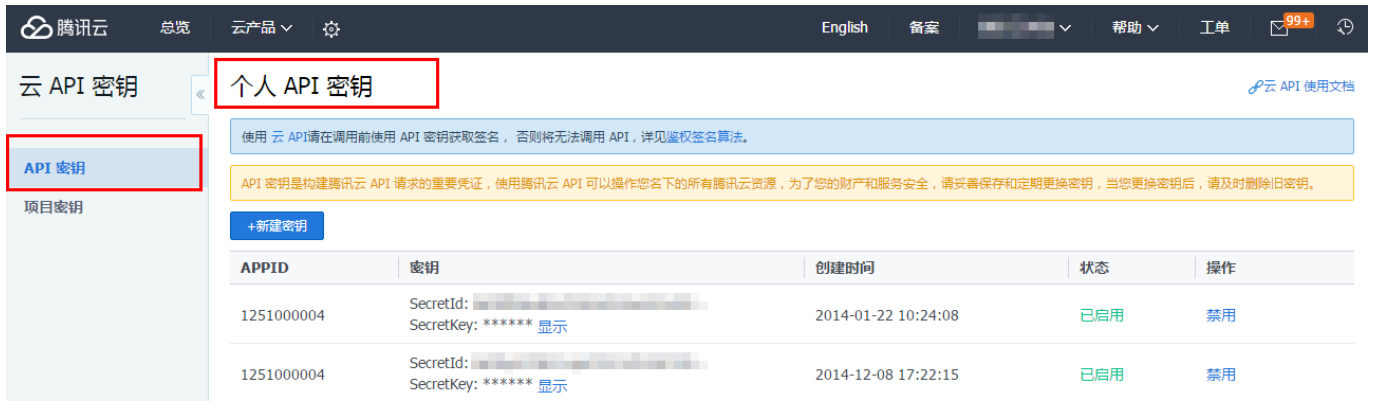
是用于加密签名字符串和服务器端验证签名字符串的密钥。用户应严格保管其 SecretKey，避免泄露。

获取途径：

1) 用户登录 [控制台](#)，单击【云产品】，选择【监控与管理】栏下的【云 API 密钥】，进入云 API 密钥管理页面。

2) 在云 API 访问密钥管理页面，单击【新建】即可以创建一对 SecretId/SecretKey，每个帐号最多可以拥有两对 SecretId/SecretKey。





4. 服务类型

智能语音提供的四种服务类型，包括语音识别（Automatic Speech Recognition，ASR）、语音合成（Text to Speech，TTS）、声纹识别（Voiceprint Recognize，VPR）、客服机器人（Customer Service Robot，CSR）。

5. 子服务类型

在智能语音服务 ASR、TTS、VRP、CSR 服务层级下，还细分出了服务子类型。例如：语音识别子服务类型当前包括离线语音识别和实时流式语音识别。

6. 模板

模板由一组指定语音处理的参数定义而成，模板的命名需要在 Project 内唯一，必备参数包括 APPID，ProjectId，服务类型、子服务类型，除此之外还包括对应子服务类型需要的其他处理参数。用户可以在控制台进行模板的创建、修改和删除。提交到智能语音服务的请求，可以通过指定模板的形式，来指定如何处理语音。

7. Bucket

用户存储在对象储存下的第一级目录。每个 Bucket 拥有自己的访问域名（外网、内网、加速访问域名）、访问控制权限（公有、私有）、回源设置、防盗链设置等属性配置项目。

语音识别

语音识别 (Automatic Speech Recognition) 服务，应用业界最先进的深度学习算法，具备出色的语音转文字、关键词检索、静音检测、语速检测、情绪识别能力。全面满足电话录音质检、实时语音输入、直播字幕及审核等多种场景下的语音处理需求。

优势

强大的识别能力：

立足腾讯社交数据大平台，积累了数十万小时标注语音数据，拥有海量的语言层数据，具备有竞争力的识别准确率。

方法最全：具备业内领先语音识别建模方法，包括 LSTM，BLSTM，Deep CNN，LFMMI，CTC 重估等多种声学模型建模技术。

应用广泛：经过内外部各大业务线的充分验证，珍爱网、知乎、北京知识产权法院、微信、QQ 等，覆盖行业广泛，日服务亿级用户，丰富的业务应用场景。

语音转文字

语音转文字功能可以实现成语音到文字的转换，语音输入设备可以是电话、手机、鹅颈麦克风、麦克风阵列等等，同时您可以根据业务指定通用或者合适的领域引擎。广泛应用于电话录音转写、医院电子病历输入、会议现场同步字幕、直播配字幕等场景。

关键词检索

关键词检索能够帮助您从语音中检索出业务关注的词汇。您可以同时提供多达 200 个关键词，关键词检索功能可以帮助您定位语音中出现了哪个或哪几个关键词以及它们分别相对语音开头的时间偏移位置，同时返回结果置信度。广泛应用于电话录音、直播、APP 音频质检审核场景。

语速检测

语速检测可以对说话人的语速进行检测，给出分钟内说出的字数。较人工检测方式，准确率更高，效率更高，成本更低。广泛应用于电话座席的语速监控，同声传译场景的发音人语速监控，电视广播播音员的语速评测等场景。

语音合成

语音合成 (Text To Speech) 满足已知文本生成语音的需求, 打通人机交互闭环。多种音色选择, 支持自定义音量、语速, 为企业客户提供定制自有领域词库和个性化发音人服务, 让发音更自然、更专业、更符合场景需求。语音合成广泛应用于语音导航、有声读物、标准发音领读、自动新闻播报等场景。

优势

a. 算法领先

腾讯自研的基于隐马尔可夫模型 (HMM) 的语音合成系统, 采用业内先进的深度学习合成算法, 具备合成快、合成语音自然流畅等优势, 让设备和应用轻松发声, 人机语音交互更加顺畅高效。

b. 高稳定性

经过海量用户的充分验证, 服务稳定高效, 具备更高的业务场景匹配度, 服务产品包括微信、腾讯地图、QQ 浏览器有声小说、2345 天气王等日服务亿级用户。

产品功能

支持中文、英文、中英文等多种语种的合成; 支持根据业务需求选择合适的音量、语速等属性; 更有多种发音人音色供选择。应用业界先进机器学习算法的语音合成引擎, 丰富的情感语料, 让合成的声音更加自然, 逼近普通人的朗读水平。

常见的应用场景

智能客服机器人

智能客服机器人通过自动识别客户语音及语义, 自动回复客户问题, 通过语音合成将回复以语音形式呈现, 让人机交互闭环。

电子有声读物

有声读物的出现让人们可以在开车、走路等不方便阅读的时刻也能享受学习的快乐。人工合成有声读物, 耗资费力, 且准确率难以保证。语音合成让有声读物的生成变得更简单, 情感合成技术让声音更自然动听, 自定义发音人支持个性化音色, 满足业务多方位合成需求。

智慧教育

配合智慧教育系统, 语音合成可以实现中英文音素、单词、词组、课文的标准朗读及带读。除公共基础教育课

堂应用外，还可以在课外教育培训机构及教辅软件中广泛应用。

声纹识别

声纹识别 (Voice Print Recognition) 是根据说话人的声波特性进行身份辨识的服务，身份辨识与口音无关，与语言无关，可以用于说话人辨认和说话人确认。

优势

a. 高稳定性

经过海量用户的充分验证，服务稳定可靠，能够随业务规模无限扩展，已经在微信声纹锁、微众银行远程开户等场景应用，服务亿级用户。

b. 高安全性

声纹特征和说话内容同时验证，自然结合，双重因子安全验证，进一步增强安全可靠。随机数字或文本的声纹校验，可有效地拒绝仿冒录音回放的欺诈。

基于随机数字密码识别方法

身份验证时，用户无需再牢记“数字密码”，用声音代替按键输入，验证时系统自动分配数字密码，用户根据分配读出指定数字，完成身份验证。该功能能够解决客户因遗忘或混淆密码而带来的困扰，实时识别，速度快，安全性高。

基于指定文本密码识别方法

身份注册时，用户需要按照系统要求读出指定文本，并需要牢记文本密码。身份认证时，系统会同时校验指定文本内容及声纹特性。文本密码需要用户妥善保存，并保密。同时支持系统自动分配及用户自定义文本密码，灵活方便，您可以根据业务需求进行选择。

说话人辨别方法

说话人辨认前，需要先给定一个目标说话人集合，包含所有用户的语音特征序列，将待测语音从说话人集合的用户中辨认出来，从而鉴别出说话人，是一个“多选一”的选择问题。

客服机器人

客服机器人 (Customer Service Robot) 依托腾讯领先的自然语言综合处理能力，深度学习企业自有知识库，构建能够在多渠道随时响应客户问题的智能客服机器人，同时支持与内部业务系统交互，帮助企业以更低的人力成本更智能地全方位触达客户。

机器人自动回复

客服不在线，支持全自动回复；基于上下文的多轮会话能力，不仅支持业务资讯类问题的自动回复，同时满足业务操作类问题的引导交互；提供 API 及 SDK

接入方式，与内部业务系统无缝对接，有效整合企业数据，让服务更智能。

人工在线客服

机器未解决问题或关键业务问题，无缝对接人工客服，平衡服务效率与满意度，可以根据业务体系、工作量、既往服务史等规则，自动分配人工客服。

数据运营分析

针对机器人及客服的会话量、服务质量、关键词进行实时统计分析；全方位分析客户满意度、情感、热点问题，实时掌握客户服务情况。

应用场景

金融行业

应用客服机器人可以帮助企业有效节省 50%

以上电话客服人力；多渠道接入能力，可以快速构建全媒体客服。降低企业成本的同时完善企业服务体系。

电商

统一人工及机器人客服后台运营系统，让电商客户轻松接入微信、APP、网页等多种服务渠道，降低人力消耗成本；历史消息留存机制，有效追踪用户问题，提升服务体验。

旅游

精准覆盖用户高频率咨询，同时支持与内部业务系统打通，为用户提供自助交互操作类服务，例如：退票、改签、查询等等；服务流程大大简化，用户操作快捷方便，服务标准统一，有效提升企业效率。

政务

随时关注广大人民群众关心的热点问题，机器人自动回复已知问题，智能地检索出未解决的问题，供政企工作人员进行调研讨论，并通过系统做出标准回复，快速精准触达人民群众，提升政企单位的服务体验。