语音交互服务

常见问题

文档版本	01
发布日期	2024-06-20

华为技术有限公司

版权所有 © 华为技术有限公司 2024。 保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标声明

和其他华为商标均为华为技术有限公司的商标。
本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为公司对本文档内容不做任何明示或暗示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

华为技术有限公司

地址： 深圳市龙岗区坂田华为总部办公楼 邮编：518129

网址： https://www.huawei.com

客户服务邮箱： support@huawei.com

客户服务电话： 4008302118

安全声明

漏洞处理流程

华为公司对产品漏洞管理的规定以“漏洞处理流程”为准，该流程的详细内容请参见如下网址：
https://www.huawei.com/cn/psirt/vul-response-process
如企业客户须获取漏洞信息，请参见如下网址：
https://securitybulletin.huawei.com/enterprise/cn/security-advisory

 版权所有 © 华为技术有限公司

目 录

1 产品咨询类 1
1.1 什么是语音交互服务 1
1.2 支持哪些语言 1
1.3 是否支持 aac 格式的语音文件转文字 1
2 计费类 2
2.1 服务当前处于商用还是免费阶段 2
3 API 使用类 3
3.1 为什么会出现识别结果非常差的情况 3
3.2 有哪些途径可以使用语音交互的 API 3
3.3 返回状态正常，但识别结果为空 3
3.4 Token 消息体中 user name，domain name 和 project name 分别指的是什么 4
3.5 如何调整并发 4
3.6 调用语音交互的 API 显示 auth failed 4
3.7 调用语音交互的 API 超时怎么处理（connect timeout） 5
3.8 如何解决“The silence time is too long, and the audio will not be recognized”报错 5
3.9 Token 怎么获取 6
4 SDK 使用类 8
4.1 Java SDK 是否支持使用 Maven 导入 8
4.2 Java SDK 是否可在 Android 使用 8
4.3 SDK 支持哪些语言 8
5 错误码类 9
5.1 SIS.0304 报错 9
5.2 SIS.0032 报错 9
5.3 APIG.0101 报错 10
5.4 APIG.0201 报错 10
5.5 APIG.0301 报错 10
5.6 Token 与 Project ID 不匹配报错 11
5.7 calc ak sk signature fail:signature expired 报错 11
6 部署类 13
6.1 语音交互服务是否支持私有化部署或离线部署 13
7 区域与可用区 14

 版权所有 © 华为技术有限公司

7.1 什么是区域与可用区 14
8 权限类 16
8.1 语音交互服务需要申请什么权限 16
8.2 如何通过主账号创建一个子账号 16

 版权所有 © 华为技术有限公司

1 产品咨询类

1.1 什么是语音交互服务

语音交互服务（Speech Interaction Service，简称SIS）是一种人机交互方式，用户通过实时访问和调用API（Application Programming Interface，应用程序编程接口）获取语音交互结果。例如用户通过语音识别功能，将口述音频、普通话或者带有一定方言的语音文件识别成可编辑的文本，同时也支持通过语音合成功能将文本转换成逼真的语音等提升用户体验。适用场景如语音客服质检、会议记录、语音短消息、有声读物、电话回访等。

1.2 支持哪些语言

- 一句话识别支持中文普通话，带方言口音的普通话和方言（四川话、粤语和上海话）。

1.3 是否支持 aac 格式的语音文件转文字

一句话识别可实现语音转文字，一句话识别支持aac格式。

版权所有 © 华为技术有限公司

2 计费类

2.1 服务当前处于商用还是免费阶段

当前语音交互服务在**新加坡region暂未商用，**开通服务和获取详细收费明细请**联系华为云客服**。

 版权所有 © 华为技术有限公司

3 API 使用类

3.1 为什么会出现识别结果非常差的情况

问题现象

调用语音识别接口，识别结果同真实结果差别很大，或者服务端报音频格式错误。

解决方案

1. 检查音频采样率是否符合。

   对于裸音频，可采用toolsoft Audio player等工具进行试听，通过设置不同的采样率，播放正常的即为音频正常采样率。

2. 如果检查参数“property”是否与采样率一致，如“chinese_8k_common”， *8k* 即采样率。

3.2 有哪些途径可以使用语音交互的 API

共有三种方式可以基于已构建好的请求消息发起请求。

- cURL

  cURL是一个命令行工具，用来执行各种URL操作和信息传输。cURL充当的是HTTP客户端，可以发送HTTP请求给服务端，并接收响应消息。cURL适用于接口调试。关于cURL详细信息请参见**https://curl.haxx.se/**。

- 编码

  通过编码调用接口，组装请求消息，并发送处理请求消息。

- REST客户端

  Mozilla、Google都为REST提供了图形化的浏览器插件，发送处理请求消息。下载postman请参见**https://www.getpostman.com/**。

3.3 返回状态正常，但识别结果为空

一般由于格式不匹配造成的。请按照以下情况进行排查。

 版权所有 © 华为技术有限公司

1. 请确保音频格式和请求格式参数保持一致，音频采样率和选择“property”参数中采样率保持一致。

2. 请确保音频位宽为16bit，目前仅支持16bit位宽的音频，如果低于该位宽的音频，则无法正常识别。

3.4 Token 消息体中 user name，domain name 和 project name 分别指的是什么

user name指用户名称，domain name指用户所属的账号名称。如果是账号获取token，账号的user name和domain name相同。如果是IAM用户获取token（账号可以创建多个IAM用户），user name和domain name不相同，user name为实际的用户名称。

project name指项目名称，如“ap-southeast-3”。。

图 3-1 我的凭证

3.5 如何调整并发

API接口支持并发调用。具体可咨询商务合作。

3.6 调用语音交互的 API 显示 auth failed

问题现象

校验失败服务端通常是使用不正确的Token进行访问，可能产生的原因有多种。

解决方案

1. 检查用户账号是否正常。

   当用户账号存在未充值或者属于冻结状态。您可登录账号，在用户检查是否有余额（以上信息在用户Console均可以查到）。

2. 检查获取Token是否正常。
   - 确保获取Token为24小时之内的Token。
   - 确保获取Token的参数嵌套是否正确。

     正确的嵌套应该如下列代码示例所示，错误的嵌套也能获取Token，但无法访问服务。

3. 使用AK/SK方式访问，检查AK和SK是否填写正确。

版权所有 © 华为技术有限公司

4. 检查“project_id”是否填写正确。

   说明

   如果无法确定原因和解决问题，请联系技术服务人员。

```
auth_data = {
  "auth": {
    "identity": {
      "password": {
        "user": {
          "name": user_name,
          "password": password,
          "domain": {
            "name": domain_name
          }
        }
      },
      "methods": ["password"]
    },
    "scope": {
      "project": {
        "name": region
      }
    }
  }
}
```

3.7 调用语音交互的 API 超时怎么处理（connect timeout）

问题现象

超时通常是由于网络不通造成的，需要对网络环境和代码进行具体分析。

解决方案

1. 检查网络是否正常。

   确保网络正常连接互联网，正常访问华为云官网，成功获取Token等。如果使用代理，需要保证代理可用性。

2. 检查url是否存在错误。

   如书写出错，http方法出错（get和post混淆）等。

3.8 如何解决“The silence time is too long, and the audio will not be recognized”报错

问题描述

识别文件错误，提示“The silence time is too long, and the audio will not be recognized”？

可能原因

实时语音识别的初始静默时间过长无法识别。

 版权所有 © 华为技术有限公司

解决方案

方法一：将参数设置中静默时间的参数vad_head 的值设置大一些，详细请参见API参考的实时语音识别的请求参数。

方法二：处理音频文件，将识别文件初始的静默音频剪切掉。

3.9 Token 怎么获取

Token接口请求构造如下所示：

username、***domainname***、***project name***可登录控制台“**我的凭证**”页面获取。***password***为用户密码。

图 3-2 我的凭证

如果您的华为云账号已升级为华为账号，将不支持获取Token。建议为您自己创建一个IAM用户，获取IAM用户的Token。

```
POST https://iam.ap-southeast-3.myhuaweicloud.com/v3/auth/tokens
Content-Type: application/json

{
    "auth": {
        "identity": {
            "methods": [
                "password"
            ],
            "password": {
                "user": {
                    "name": "username",
                    "password": "********",
                    "domain": {
                        "name": "domainname"
                    }
                }
            }
        },
        "scope": {
            "project": {
                "name": "region"
            }
        }
    }
}
```

如下图所示，返回的响应消息头中“x-subject-token”就是需要获取的用户Token。获取Token之后，您就可以使用Token认证调用SIS服务API。

版权所有 © 华为技术有限公司

获取Token后，再调用SIS服务API接口时，您需要在请求消息头中添加“X-Auth-Token”，其值即为Token。例如Token值为“ABCDEFJ....”，则调用接口时将“X-Auth-Token: ABCDEFJ....”加到请求消息头即可，如下所示。

```
Content-Type: application/json
X-Auth-Token: ABCDEFJ....
```


版权所有 © 华为技术有限公司

4 SDK 使用类

4.1 Java SDK 是否支持使用 Maven 导入

支持Maven引入，引入方式详见《 SDK参考 》中章节。

4.2 Java SDK 是否可在 Android 使用

目前由于兼容性问题，Java SDK不支持在android使用。

4.3 SDK 支持哪些语言

目前union sdk（ 仅支持https接口 ）已支持java、python，后续根据需求考虑支持其他语言。

版权所有 © 华为技术有限公司

5 错误码类

5.1 SIS.0304 报错

问题现象

客户端长时间没有发送语音，服务端超过20s没有收到语音，就会报这个错误。

解决方案

1. 检查代码是否存在发送音频后睡眠过多时间。

2. 检查是否发送end请求，是否关闭客户端。超过20s没有发送数据同时也没有断开连接，就会报该错误。在使用完毕后一定要记得断开与服务端连接。

5.2 SIS.0032 报错

问题描述

为何调用一句话识别的api接口，返回错误码"SIS.0032"，错误信息为:"'audio_format' is invalid"?

解决方案

确保设置的audio_format格式的值是合法，待识别的音频格式和audio_format参数格式要保证一致。

一句话识别支持的音频格式详见表5-1。

表 5-1 audio_format 取值范围

audio_format取值	说明
pcm16k16bit	16k16bit单通道录音数据。
pcm8k16bit	8k16bit单通道录音数据。

 版权所有 © 华为技术有限公司

audio_format取值	说明
wav	带wav封装头的格式，从封装头中自动确定格式，目前仅支持8k/16k采样率、单通道、pcm编码格式。

5.3 APIG.0101 报错

访问的API不存在或尚未在环境中发布。The API does not exist or has not been published in the environment.

- 请检查API的URL是否拼写正确，例如，URL中是否缺少project_id。
- 检查URI中的endpoint是否填写正确，例如，调用SIS服务api，endpoint错填为其他服务。

  https://ocr.ap-southeast-3.myhuaweicloud.com/v1/{project_id}/asr/short-audio
- HTTP请求方法（POST，GET等）是否正确。
- 具体的URI信息请在各接口的API页面查看。

5.4 APIG.0201 报错

如果调用API时返回错误信息与错误码："error_msg":"Backend timeout.","error_code":"APIG.0201"，表示请求超时。

您可以尝试以下方案进行解决。

使用Postman等工具直接调用服务，查看是否成功，如果调用成功表示服务接口没有问题，请执行如下步骤1和步骤2进行检查：

1. 请检查原调用请求是否过于频繁，如果是并发过大，可以通过重试机制解决，在代码里检查返回值，碰到这个并发错误可以延时一小段时间（如2-5s）重试请求；也可以后端检查上一个请求结果，上一个请求返回之后再发送下一个请求，避免请求过于频繁。
2. 请检查音频是否过大，或网络延时过长。如果音频过大，可利用mp3格式进行压缩；如果网络延时较长，可以适当提升网络传输速度。

您可以直接使用SIS服务提供的SDK调用服务，避免出现该问题，SDK获取地址请参见**获取SDK**。

如以上方案无法解决问题，请**联系我们**。

5.5 APIG.0301 报错

如果在调用API时返回错误信息与错误码：

- "error_msg":"Incorrect IAM authentication information: decrypt token fail","error_code":"APIG.0301"，是因为token解密失败。

  解决方案：（1）请检查token是否过期；（2）获取token的方法，请求体信息是否填写正确，token是否正确完整；（3）获取token的环境与调用的环境是否一致是否跨区；（4）检查账号权限，是否欠费，被冻结等。

版权所有 © 华为技术有限公司

- "error_msg":"Incorrect IAM authentication information: verify aksk signature fail","error_code":"APIG.0301"，是因为AK/SK认证失败。

  解决方案：（1）请检查AK/SK是否正确，AK对应的SK错误，不匹配；AK/SK中多填了空格；（2）AK/SK频繁出现鉴权出错，连续错误5次以上，被锁定5分钟（5分钟内，则一直认为其是异常的鉴权请求），5分钟后解锁重新认证；（3）检查账号权限，是否欠费，被冻结等账号。

5.6 Token 与 Project ID 不匹配报错

获取Token时，出现the parameters of requesting token are wrong or project_id doesn't match the token报错。

检查获取Token过程中，请求url中的区域信息与获取token请求体中的project name是否匹配。

```
POST https://iam.ap-southeast-3.myhuaweicloud.com/v3/auth/tokens
Content-Type: application/json

{
    "auth": {
        "identity": {
            "methods": [
                "password"
            ],
            "password": {
                "user": {
                    "name": "username",
                    "password": "********",
                    "domain": {
                        "name": "domainname"
                    }
                }
            }
        },
        "scope": {
            "project": {
                "name": "projectname"
            }
        }
    }
}
```

5.7 calc ak sk signature fail:signature expired 报错

调用API时，出现Incorrect IAM authentication information: calc ak sk signature fail:signature expired报错。表示的是AK、SK凭证过期。

请先检测您使用的是临时AK、SK，还是永久性AK、SK。临时安全凭证通过接口**临时访问密钥AK/SK**获取；永久安全凭证通过**我的凭证**界面控制台获取。

版权所有 © 华为技术有限公司

如果使用的是永久性AK、SK出现此报错。请检查您的服务器时间和网络时间是否一致，时间同步后，可解决该报错。

 版权所有 © 华为技术有限公司

6 部署类

6.1 语音交互服务是否支持私有化部署或离线部署

目前语音交互服务暂不支持私有化部署和离线部署。

 版权所有 © 华为技术有限公司

7 区域与可用区

7.1 什么是区域与可用区

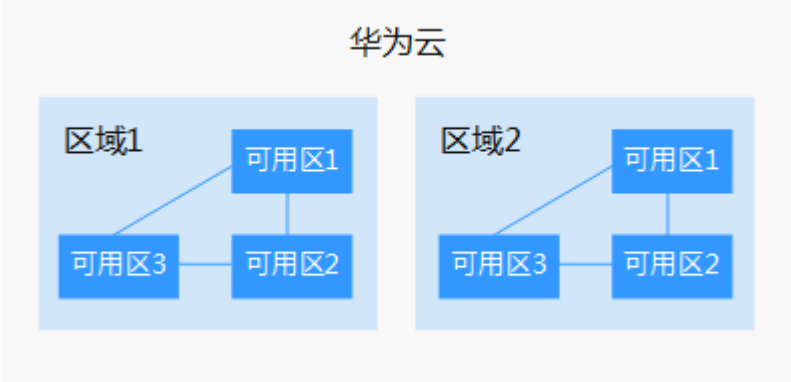
什么是区域、可用区

我们用区域和可用区来描述数据中心的位置，您可以在特定的区域、可用区创建资源。

- 区域（Region）：从地理位置和网络时延维度划分，同一个Region内共享弹性计算、块存储、对象存储、VPC网络、弹性公网IP、镜像等公共服务。Region分为通用Region和专属Region，通用Region指面向公共租户提供通用云服务的Region；专属Region指只承载同一类业务或只面向特定租户提供业务服务的专用Region。
- 可用区（AZ，Availability Zone）：一个AZ是一个或多个物理数据中心的集合，有独立的风火水电，AZ内逻辑上再将计算、网络、存储等资源划分成多个集群。一个Region中的多个AZ间通过高速光纤相连，以满足用户跨AZ构建高可用性系统的需求。

图7-1阐明了区域和可用区之间的关系。

图 7-1 区域和可用区



目前，华为云已在全球多个地域开放云服务，您可以根据需求选择适合自己的区域和可用区。

 版权所有 © 华为技术有限公司

如何选择区域?

选择区域时，您需要考虑以下几个因素：

- 地理位置

  一般情况下，建议就近选择靠近您或者您的目标用户的区域，这样可以减少网络时延，提高访问速度。不过，在基础设施、BGP网络品质、资源的操作与配置等方面，中国大陆各个区域间区别不大，如果您或者您的目标用户在中国大陆，可以不用考虑不同区域造成的网络时延问题。

  - 在除中国大陆以外的亚太地区有业务的用户，可以选择“中国-香港”、“亚太-曼谷”或“亚太-新加坡”区域。
  - 在非洲地区有业务的用户，可以选择“南非-约翰内斯堡”区域。
  - 在欧洲地区有业务的用户，可以选择“欧洲-巴黎”区域。

- 云服务之间的关系

  如果多个云服务一起搭配使用，需要注意：

  - 不同区域的弹性云服务器、关系型数据库、对象存储服务内网不互通。
  - 不同区域的弹性云服务器不支持跨区域部署在同一负载均衡器下。

- 资源的价格

  不同区域的资源价格可能有差异。

如何选择可用区

是否将资源放在同一可用区内，主要取决于您对容灾能力和网络时延的要求。

- 如果您的应用需要较高的容灾能力，建议您将资源部署在同一区域的不同可用区内。
- 如果您的应用要求实例之间的网络延时较低，则建议您将资源创建在同一可用区内。

区域和终端节点

当您通过API使用资源时，您必须指定其区域终端节点。有关区域和终端节点的更多信息，请参阅语音交互服务的**地区和终端节点**。

 版权所有 © 华为技术有限公司

8 权限类

8.1 语音交互服务需要申请什么权限

语音交互服务本身无权限限制，用户不需要额外添加任何权限，即可对语音交互服务执行操作。

8.2 如何通过主账号创建一个子账号

语音交互服务服务不支持IAM细粒度划分策略，IAM账号与主账号具备相同的操作权限。

您可以使用IAM的用户管理功能，给员工或应用程序创建IAM用户，可避免分享自己的账号密码。详细创建步骤请参见**创建IAM用户**。

版权所有 © 华为技术有限公司