

## 1.功能包介绍

### 1.1 离线语音合成

本功能包中提供了离线语音合成功能,合成的引擎在 SDK 中,可以在无网络的情况下使用,其提供的接口如 2.1 所示,本功能包目录如下:

```
├── audio
├── CMakeLists.txt
├── include
├── launch
├── lib
├── package.xml
├── src
└── srv
```

#### ☒ include

存放语音合成相关的头文件.

#### ☒ lib

存放所有动态库文件,用户需要根据自己预使用的 linux 平台进行动态库的选择.现支持 arm64(支持英伟达 TX1 TX2 NANO 等 JETSON 系列)以及 x64(通用 PC).

#### ☒ launch

包含离线语音合成接口功能包启动文件,在 launch 文件中加载了 appid,所谓 appid 是在讯飞开发平台创办应用时得到的,该参数非常重要,只有有效的 appid 且该 appid 对应的应用中有可用的离线语音合成引擎及发音人时才可保证正常使用(见下 config 中介绍.),若出现离线语音合成报错,则可考虑更换 appid 以及下述 config 中的资源文件.

#### ☒ config

存放离线语音合成相关的资源文件 msc,该 msc/res/tts 文件夹中存放合成引擎"common.jet"以及说话人引擎"xiaoyan.jet"和"xiaofeng.jet"(默认提供的是这 xiaoyan、xiaofeng 两种发音人,如果需要其他发言人需单独购买).这两个引擎与 appid 对应.更换时需要一并更换.更换方式见第 4 节.此外,在 msc 中会保存 log 记录,可定期清除.

#### ☒ src

包含离线语音合成接口源程序.

#### ☒ srv

包含自定义的服务类型.

☒ **audio**

存放合成好的 wav 文件,下次合成将替换上一合成结果.

## 1.2 在线语音合成

讯飞在线语音合成目前支持中文普通话、部分中文方言以及英文播报,音色涉及男声、女声、童声、中年等.在官网有发音人列表,建议在人机交互界面先在线体验发音人效果,然后再调本功能包中提供的发音人接口以满足用户期望.本功能包提供的发音人列表为如下,用户可更换我自己的 **appid** 然后在网站上选择开通发音人来增加或修改发音人列表.此外,需要注意的是在线语音合成单次最多支持 **8192** 个字节,超过限制的话需要先分段合成,而后请自行进行拼接处理.特别地,讯飞在线语音合成已上线语音合成 **2.0**,声色更加好听.表中后三行是推荐的 **2.0** 发音人,你可在你的控制台中添加和试用.

讯飞小燕	普通话	xiaoyan
讯飞许久	普通话	aisjiuxu
讯飞小萍	普通话	aisxping
讯飞小婧	普通话	aisjinger
讯飞许小宝	普通话	aisbabyxu
讯飞小丸子	普通话	xiaowanzi
讯飞蜡笔小新	普通话	xiaoxin
讯飞晓峰	普通话	xiaofeng
讯飞小奇	普通话	xiaoqi
讯飞小露 (2.0)	普通话	x2_yezi
讯飞一菲 (2.0)	普通话	x2_yifei
讯飞小鹏 (2.0)	普通话	x2_pengfei

本功能包的文件目录如下:

```
└─ audio
└─ CMakeLists.txt
└─ include
└─ launch
└─ lib
└─ package.xml
└─ src
└─ srv
```

☒ **include**

存放语音合成相关的头文件.

#### ☒ lib

存放所有动态库文件,用户需要根据自己预使用的 linux 平台进行动态库的选择.现支持 arm64(支持英伟达 TX1 TX2 NANO 等 JETSON 系列)以及 x64(通用 64 位 PC).

#### ☒ launch

包含离线语音合成接口功能包启动文件,在 launch 文件中加载了 **appid**,所谓 **appid** 是在讯飞开发平台创办应用时得到的,该参数非常重要,只有在使用期限范围内的 **appid** 才可正常使用,使用时需要与动态库 **libmsc.so** 配合使用若出现离线语音合成报错,则可考虑更换 **appid** 以及相应的动态库.

此外 launch 文件也加载了一些接口包中使用的默认参数,如说话人语速 **voice\_speed**,说话人名字"xiaoyan",服务接口名字等.在 launch 文件中启动是为了简化调试过程.

#### ☒ src

包含在线语音合成接口源程序.

#### ☒ srv

包含自定义的服务类型.

#### ☒ audio

存放合成好的 wav 文件,下次合成将替换上一合成结果.

#### ☒ config

存放 msc 的 log 文件.

## 2 语音合成接口描述

### 2.1 离线语音合成

☐ 类 型: 服务

☐ 服务名称: /xf\_mic\_tts\_offline\_node/play\_txt\_wav

☐ 服务类型: xf\_mic\_tts\_offline::Play\_TTS\_srv

```
'''
string text
---
string result
string fail_reason
'''
```

□ 内 容: text 为要播放的文本内容,中英文均可。

□ 描 述: 若一切正常,则开始合成并播放,result 返回为"ok",否则返回为"fail",且已知返回类型为:

- 1) "text is null\_error",表示请求的文本为空。
- 2) "login\_error",表示登录失败,可能的原因是资源路径无效
- 3) "11212\_license\_expired\_error",表示语音合成引擎授权过期.若出现该错误可根据第 4.3 节中的方法进行更换。

□ 调用示例:

//调用一次离线语音合成,合成文本为"hello,iflytek".请注意标点符号应为英文。

```
$ rosservice call /xf_mic_tts_offline_node/play_txt_wav "hello,iflytek" ""
```

## 2.2 在线语音合成

□ 类 型: 服务

□ 服务名称: /xf\_mic\_tts\_online\_node/play\_txt\_wav

□ 服务类型: xf\_mic\_tts\_online::Play\_TTS\_Online\_srv

```
""  
  
string text  
string appID  
string speakerName  
---  
string result  
string fail_reason  
""
```

□ 内 容: text 为要播放的文本内容,中英文均可;appID 为开放平台的应用账号,有默认值,若服务失效,需要更换,注意,若更换 appid 则要更换镜像中 lib 中的动态库,可参考 4.2 节.speakerName 为发音人名字,若不指定有默认值"xiaoyan",可选的发音人为: "xiaoyan"、"xiaoxin"、"xiaoqi"、"xiaofeng"、"xiaowanzi"、"aisjiuxu"、"aisxping"、"aisjinger"、"aisbabyxu"。

□ 描 述: 若一切正常,则开始合成并播放,result 返回为"ok",否则返回为"fail",且已知返回类型为:

- 1) "text is null\_error",表示请求的文本为空。
- 2) "login\_error",表示登录失败。
- 3) "unfit\_appid\_and\_lib\_error",账户与资源不匹配。(1040\*)
- 4) "speaker\_name\_error",发音人不存在。

□ 调用示例:

//调用一次在线语音合成,合成文本为"hello,iflytek",appid 和发音人为默认.请注意标点符号应为英文。

```
$ rosservice call /xf_mic_tts_online_node/play_txt_wav "hello,iflytek" "" ""
```

//调用一次在线语音合成,合成文本为"hello,iflytek",appid 为"12345678"和发音人为"aisjiuxu"

```
$rosservice call /xf_mic_tts_online_node/play_txt_wav "hello,iflytek" "12345678" "aisjiuxu"
```

### 3 启动方式

在使用本功能包前,请确保已安装音频播放的库,如下:

```
$ sudo apt-get install sox
$ sudo apt-get install mplayer
```

#### 1) 启动离线语音合成功能包

离线语音合成的接口,启动方式为:

```
'''
$ roslaunch xf_mic_tts_offline xf_mic_tts_offline.launch
'''
```

#### 2) 启动在线语音合成功能包

离线语音合成的接口,启动方式为:

```
'''
$ roslaunch xf_mic_tts_online xf_mic_tts_online.launch
'''
```

## 4.讯飞开放平台应用创建与集成

首先到[讯飞开放平台](https://console.xfyun.cn/app/mysapp)注册账户,然后点击右上角“控制台”,进入自己的控制平台.如果是首次使用讯飞开发平台,则控制台中应用是空的,点击创建新应用.

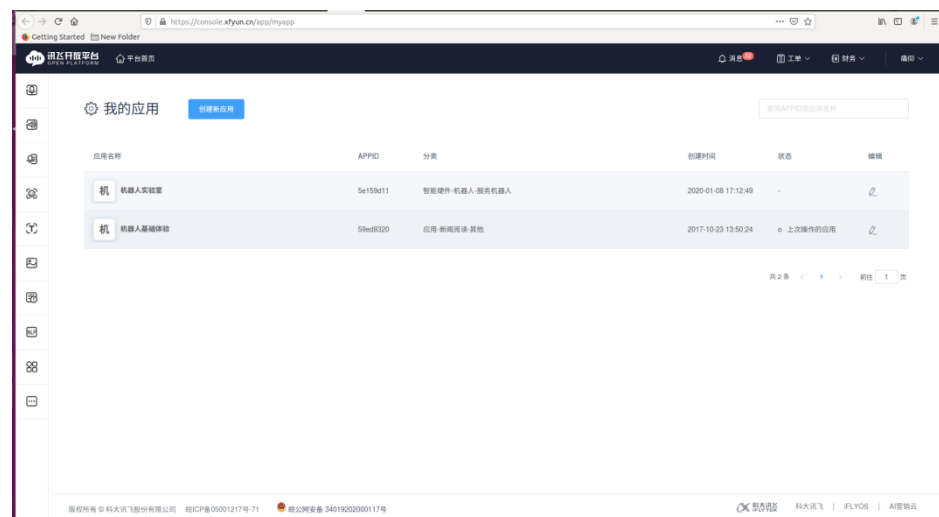


图 8 进入控制台

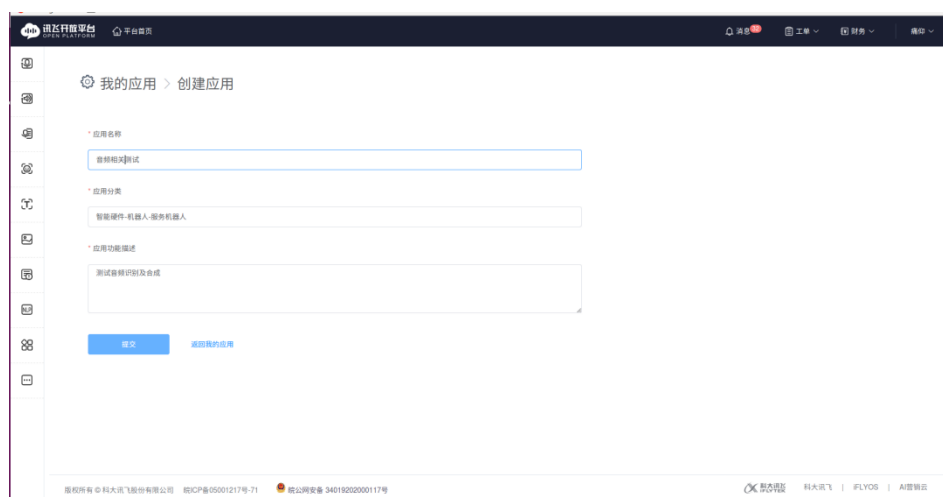


图 9 创建新应用

点击提交后,可看到用户的应用列表,选择刚刚创建的应用,则进行该应用配置界面,在左侧会有“语音识别”、“语音合成”、“语音扩展”、“人脸识别”、“图像识别”等菜单.创建完成之后就可以下载 sdk 使用了,两种方式:

1) 在开放平台主页,在菜单中选择 SDK 下载,然后选择应用名称,平台以及要使用的能力,此时下载的 SDK 会将所选的能力打包下载,其中的 lib 下的 libmsc.so 动态库也可看做是打包在一起的,其适用于你勾选的所有能力.如 4.5 所示.

2) 选择单个能力(如 4.1-4.4 介绍),然后下载单个能力的 SDK 包,此时 SDK 里的动态库仅仅适用于该能力,若下载多个能力,则会有多个 SDK 包,对应多个 libmsc.so 动态库.

## 4.1 集成离线命令词识别

### 4.1.1 添加离线命令词识别应用

选择左侧菜单中“语音识别”中的“离线命令词识别”,然后选择图中右侧的“Linux MSC”,点击“下载”来下载 sdk.

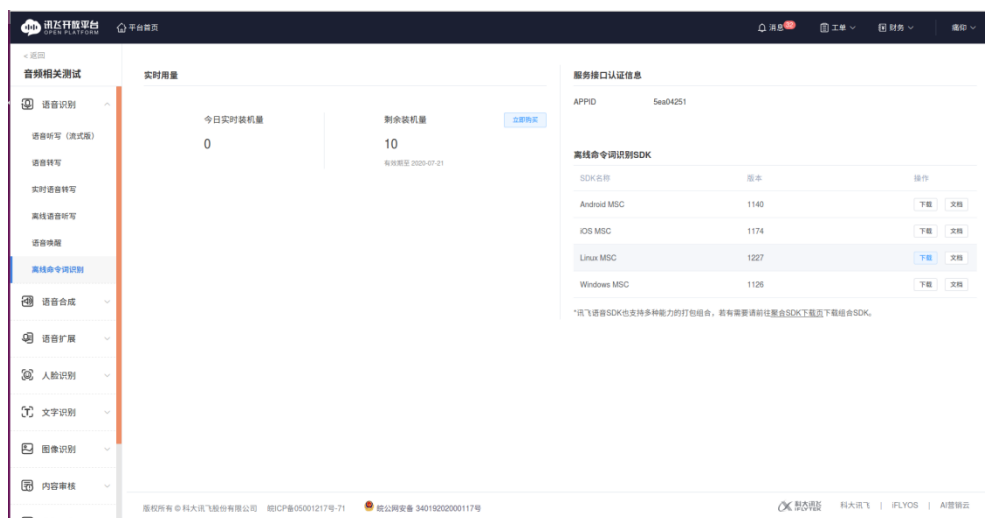


图 下载离线命令词识别 SDK

下载好的离线命令词识别的 sdk 包名称是以“标识+appid”形式命名的解压后的目录结构为:

```
└─ bin
|   └─ call.bnf
|   └─ msc
|   └─ wav
└─ doc
|   └─ BNF Grammar Development Manual.pdf
|   └─ readme.txt
└─ include
|   └─ msp_cmh.h
|   └─ msp_errors.h
|   └─ msp_types.h
|   └─ qisr.h
└─ libs
|   └─ x64
|   └─ x86
└─ README
```

```
└─ release.txt
└─ samples
    └─ asr_offline_record_sample
    └─ asr_offline_sample
```

## 4.1.2 集成方法

若在使用离线命令词识别时出现 10102 等错误时,需要替换为自己的 APPID 和离线识别引擎,appid 我们在创建应用的时候就已得到,离线识别引擎 common.jet 在 bin/msc 中,找到 common.jet 后就可以进行替换了.

1) **替换 sdk**:将用户自己的 common.jet 文件替换 vvui/config/msc/res/asr 路径下的 common.jet.将用户自己的 appid 替换 vvui/include 中 user\_interface.h 文件中的"char \*APPID".然后再次编译.

2) **替换 ros 接口包**:将用户自己的 common.jet 文件替换 vvui\_ros-master/xf\_mic\_asr\_offline/config/msc/res/asr 路径下的 common.jet,将 appid 替换 vvui\_ros-master/xf\_mic\_asr\_offline/config 中 mic\_offline\_params.yaml 文件中的"appid".然后再次编译.

## 4.2 集成在线语音合成

### 4.2.1 添加在线语音合成应用

选择左侧菜单中"语音合成"中的"在线语音合成(流式版)",然后选择图中右侧的"Linux MSC",点击"下载"来下载 sdk.



图 添加在线语音合成应用

下载好的 sdk 包名称是以"标识+appid"形式命名的解压后的目录结构为:

```
└─ bin
└─ gm_continuous_digit.abnf
```



- |   └─ ise\_cn
- |   └─ ise\_en
- |   └─ msc
- |   └─ source.txt
- |   └─ userwords.txt
- |   └─ wav
- └─ doc
- |   └─ Grammar Development Guidelines.chm
- |   └─ Open Semantic Platform API Documents.pdf
- |   └─ readme.txt
- |   └─ Recording API Documents.doc
- └─ include
- |   └─ msp\_cmh.h
- |   └─ msp\_errors.h
- |   └─ msp\_types.h
- |   └─ qise.h
- |   └─ qisr.h
- |   └─ qtts.h
- └─ libs
- |   └─ x64
- |   └─ x86
- └─ README
- └─ release.txt
- └─ samples
  - └─ iat\_online\_record\_sample

```

└─ iat_online_sample

└─ ise_online_sample

└─ sch_translate_sample

└─ tts_online_sample

```

## 4.2.2 集成方法

创建在线语音合成后,系统默认会赠送免费的发音人,见图 11 中左侧发音人授权管理.若在使用在线语音合成接口时出现错误,需要替换 appid 以及动态库,更换步骤为:

- 1) 将用户 appid 替换 xf\_mic\_tts\_online.launch 文件中的“app\_id”参数.
- 2) 将刚下载的 sdk 中 x64 中的 libmsc.so 替换 vvui\_ros-master/xf\_mic\_tts\_online/lib/x64 中的 libmsc\_online\_tts.so 文件,并重命名为 libmsc\_online\_tts.so.但该动态库只可保证你在 x64 中使用,若你使用的是 arm 平台,则需要对 libmsc.so 进行交叉编译后才可使用.

## 4.3 集成离线语音合成

### 4.3.1 添加离线语音合成应用

选择左侧菜单中“语音合成”中的“离线语音合成(普通版)”,然后选择图中右侧的“Linux MSC”,点击“下载”来下载 sdk.

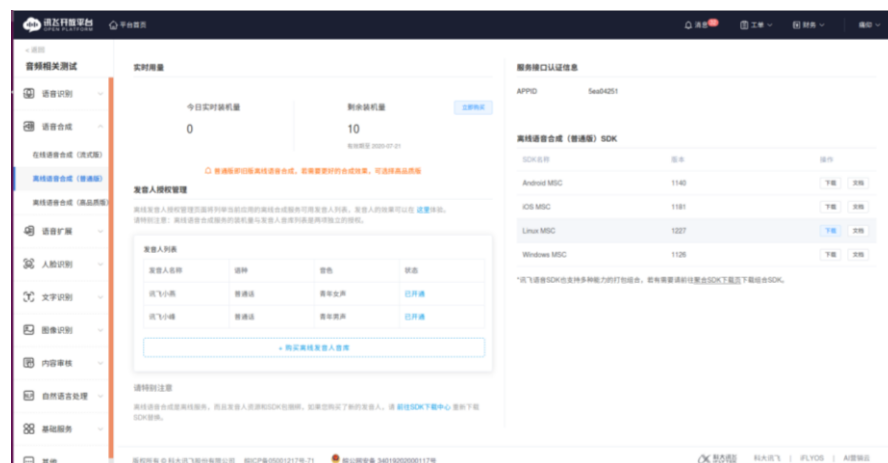


图 添加离线语音合成应用

下载好的离线命令词识别的 sdk 包名称是以“标识+appid”形式命名的解压后的目录结构为:

```

└─ bin
|   └─ msc
|
|       └─ msc.cfg

```

```

|   └─ res
├─ doc
|   └─ readme.txt
├─ include
|   └─ msp_cmn.h
|   └─ msp_errors.h
|   └─ msp_types.h
|   └─ qtts.h
├─ libs
|   └─ x64
|       └─ libmsc.so
|   └─ x86
|       └─ libmsc.so
├─ README
├─ release.txt
└─ samples
    └─ tts_offline_sample
        └─ 32bit_make.sh
        └─ 64bit_make.sh
        └─ Makefile
        └─ tts_offline_sample.c

```

### 4.3.2 集成方法

若在使用离线语音合成接口时出现错误,需要替换 **appid** 以及离线合成引擎文件,更换步骤为:

- 1) 将用户 **appid** 替换 `xf_mic_tts_offline.launch` 文件中的“appid”参数.
- 2) 将用户下载的 **sdk** 中 `bin/msc/res/tts` 下的“common.jet”,“xiaoyan.jet”,“xiaofeng.jet”替换 `vvui_ros-master/xf_mic_tts_offline/config/msc/res/tts` 路径中的相应文件.
- 3) 再次编译后运行.

注：若替换后仍出现同样报错,则可考虑更换动态库,编译后将刚下载的 sdk 中 x64 中的 libmssc.so 替换 vvui\_ros-master/xf\_mic\_tts\_online/lib/x64 中的 libmssc\_offline\_tts.so 文件,并重命名为 libmssc\_offline\_tts.so.但该动态库只可保证你在 x64 中使用,若你使用的是 arm 平台,则需要对 libmssc.so 进行交叉编译后才可使用。

## 4.4 集成 AIUI

### 4.4.1 添加 AIUI 应用

选择左侧菜单中的“其他”,会看到右侧的列表,如图所示。

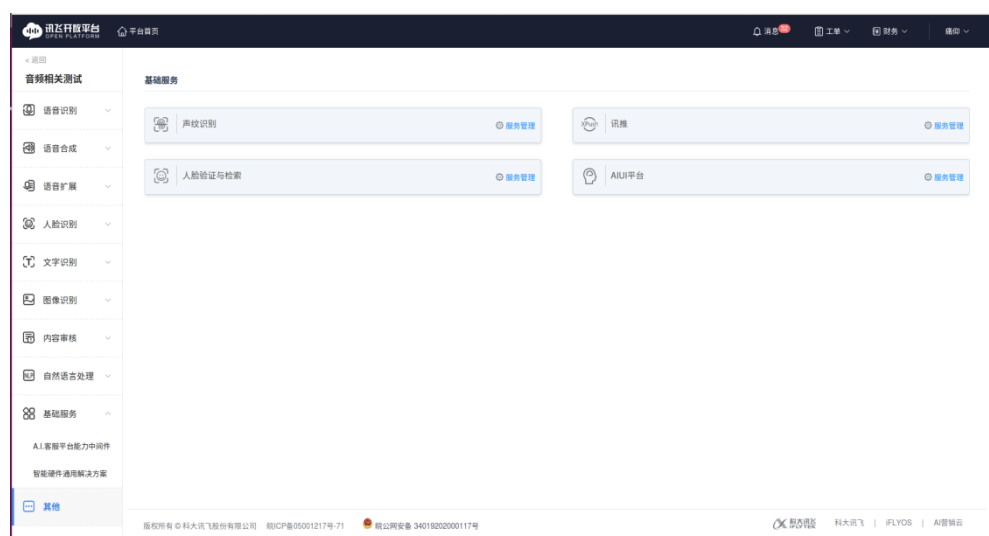


图 13 添加 AIUI 应用

选择“AIUI 平台”右侧的“服务管理”可以进入 AIUI 开放平台,该平台左侧有“应用信息”,“应用配置”,“开发工具”,“审核上线”,“服务统计”,“用户统计”菜单栏.若该 appid 是第一次添加 AIUI,系统会自动跳转到“应用信息”菜单下让用户设置应用平台,选择“Linux”即可.若已添加过则默认进入“应用配置”菜单中,如图 14 所示。

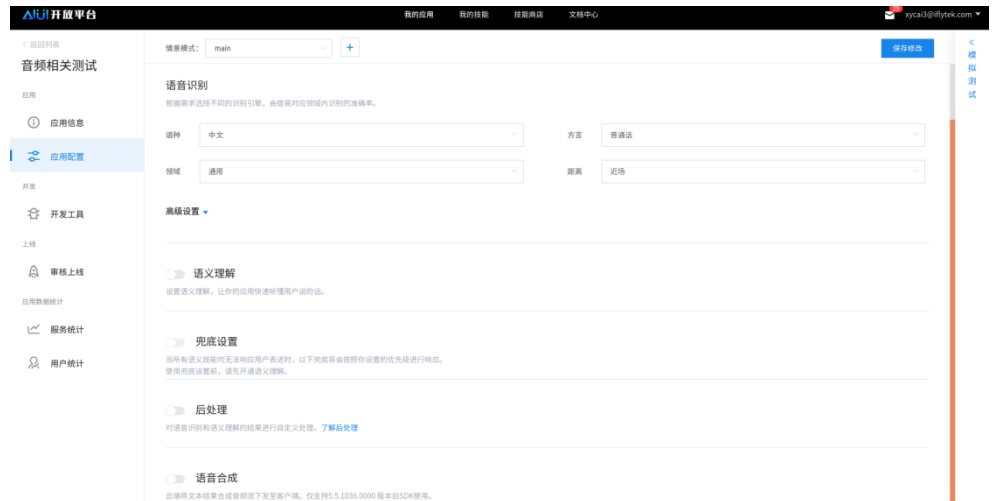


图 14 AIUI 应用配置

其中情景模式默认为“main”,可以结合用户场景需求根据使用说明来设置其他情景模式.在这里用户需要将“语义理解”、“兜底设置”、“语音合成”三个高级设置勾选,在语义理解中用户关注“语义技能”,首先选择“商店技能”,然后“添加商店技能”,你可以在弹出的技能商店中选择你需要的技能,如“天气”、“星座”等,如图 16 所示.点击确定即可.添加完后要选择页面右上角的“保存修改”才可生效.这个时候,就可以使用你刚添加的技能进行对话了.

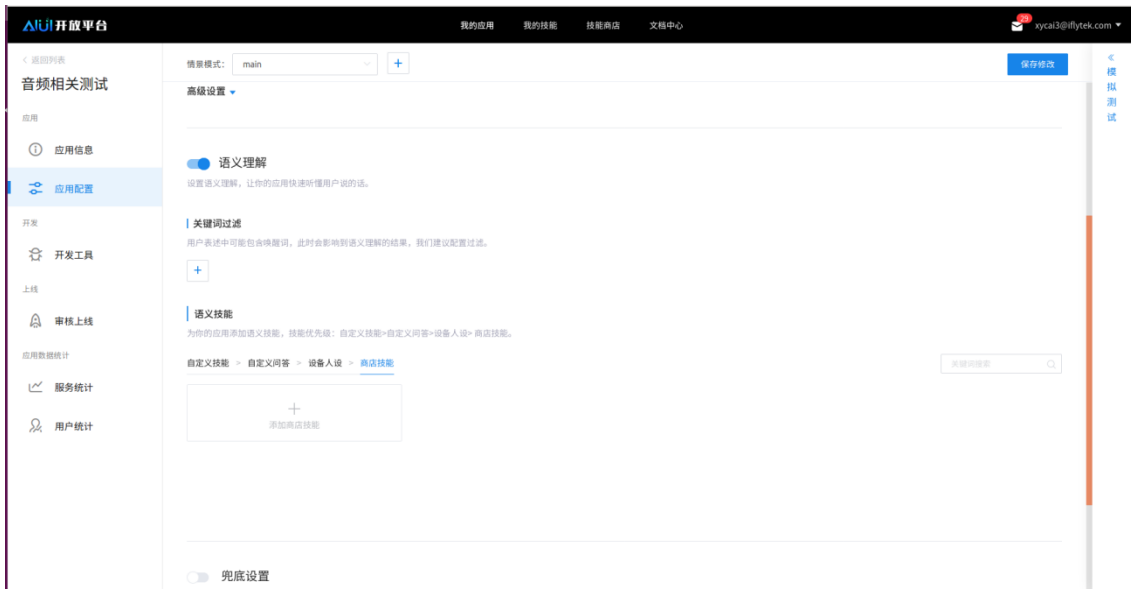


图 15 AIUI 配置语义理解



图 16 AIUI 添加商店技能

在“兜底测试”中将“图灵机器人”、“讯飞闲聊”等开启后,就可以避免在一些语料未开通时,机器人依然可以答复诸如“这些我还没学会”等等的兜底回复.除此,aiui 还提供了一些自定义的技能便于用户使用,选择“我的技能”菜单,系统会跳转到[技能控制台](#),然后根据[创建方法](#)在该平台创建完技能并发布后,再回到图中 AIUI 应用配置界面,在语音理解-语义技能中的自定义技能中添加刚创建的技能然后“保存修改”就可以了.之后就可以根据自己创建的技能以及技能标识进行交互判断了.当然,关于 AIUI 和其相关的技能应用还有很多,用户可根据官网的文档进行系统的学习,以充分使用科大讯飞开放平台的能力.

完成“应用配置”后,用户就可以下载 AIUI 的 sdk 了,选择“开发工具”,可以看到如图 17 所示的界面:

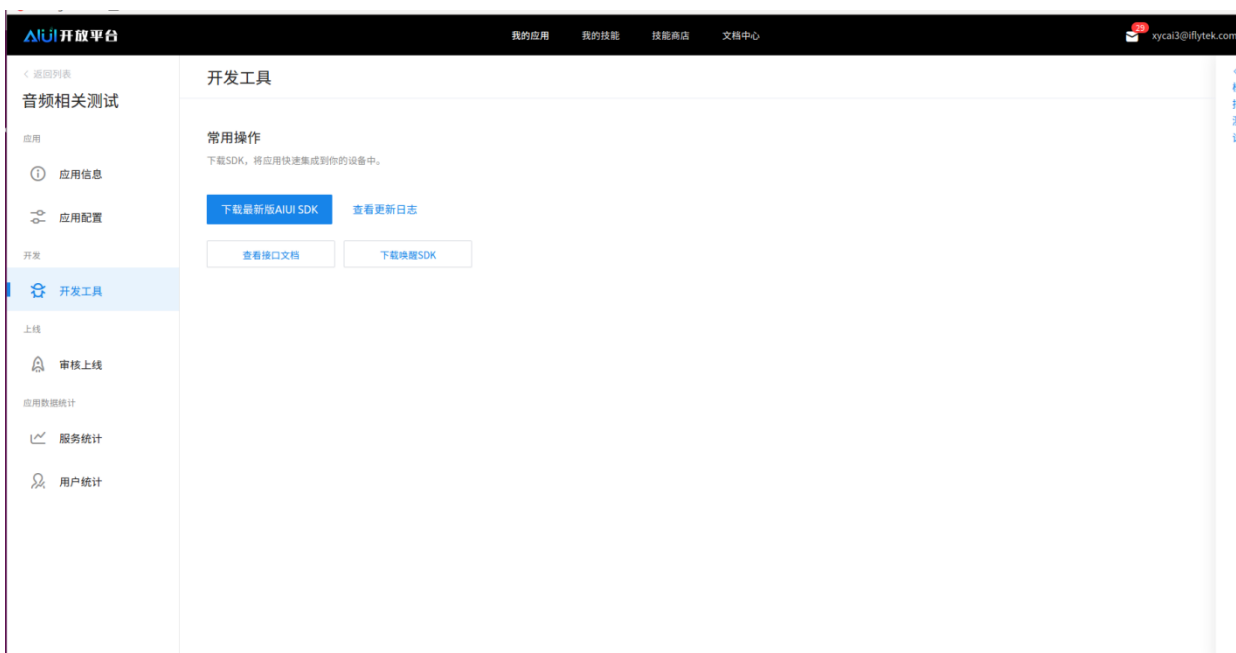


图 16 下载 AIUI SDK

点击“下载最新版 AIUISDK”即可.下载好的 sdk 包名称是以“标识+appid”形式命名的解压后的目录结构为

```
├── bin
|   ├── gm_continuous_digit.abnf
|   ├── ise_cn
|   ├── ise_en
|   ├── msc
|   ├── output.txt
|   ├── source.txt
|   ├── tts_sample.wav
|   ├── userwords.txt
|   └── wav
├── include
|   ├── aiui
|   ├── msp_cmh.h
|   ├── msp_errors.h
|   ├── msp_types.h
|   ├── qise.h
|   ├── qisr.h
|   └── qtts.h
├── libs
|   ├── x64
|   └── x86
└── README
```

```
└─ samples
    ├── aiui_sample
    ├── asr_sample
    ├── iat_record_sample
    ├── iat_sample
    ├── ise_sample
    └─ tts_sample
```

## 4.4.2 集成方法

若在使用 2.5 节中在线人机交互案例时出现“”的错误,是因为 `appid` 选择的 AIUI 应用是免费的,每天交互次数受限,你可根据如下方法进行修改,来使用你创建的“aiui”应用,方法是在“AIUI 开放平台”上找到“APPID”和“APPKEY”,在进行替换的时候只用这两个参数就可以了,替换方法将这两个参数分别替换 `AIUITester.cpp` 文件中的 `AIUITester::createAgent` 函数里面的字符串 `appid` 以及 `key`,

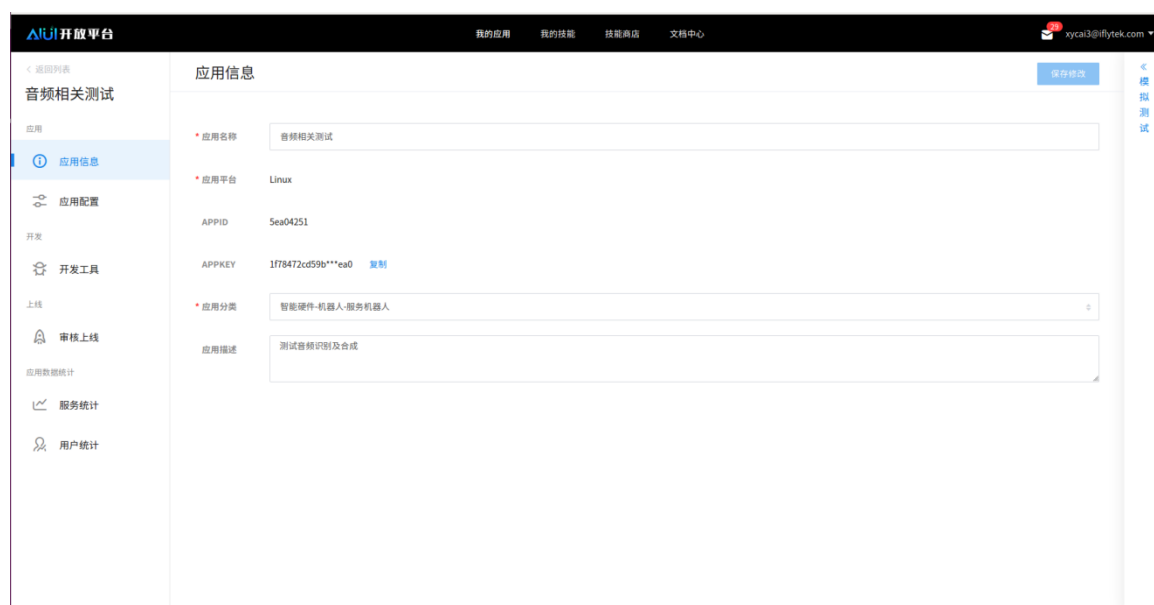


图 17 查看 APIID 与 APPKEY



## 4.5 合并下载 SDK

进入讯飞开发平台主页,上部菜单栏里找到“资料库”,然后选择“SDK”下载.

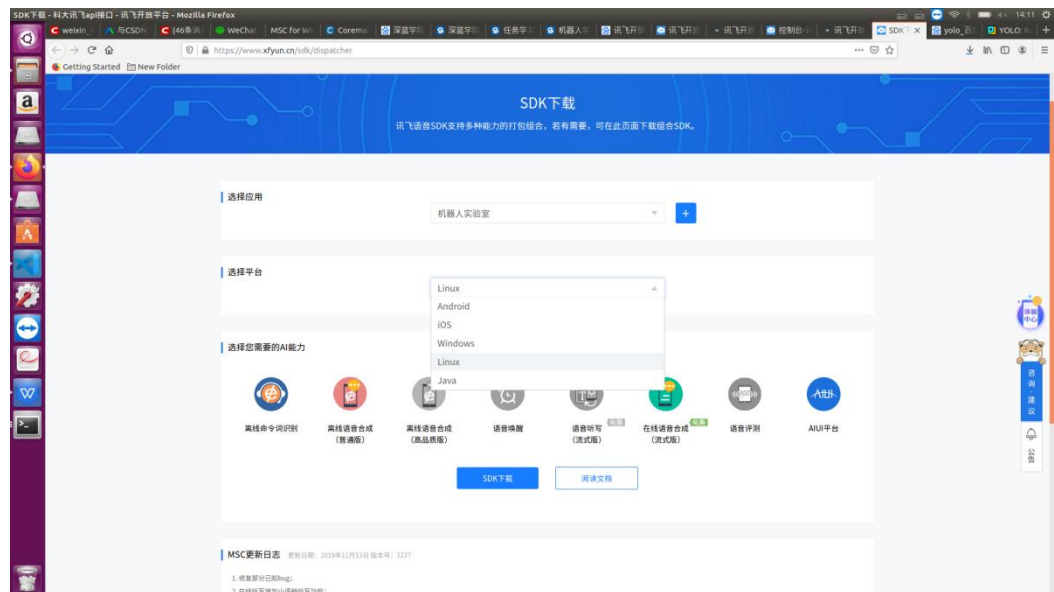


图 合并下载 SDK

## 5. 常见错误

- ☑ “10407” 或 “unfit\_appid\_and\_lib\_error”  
appid 与动态库 libmsc.so 不匹配, 在修改后一定要编译后再运行;
- ☑ “11210”  
appid 与离线资源即 common.jet 不匹配.
- ☑ “11201”  
在线业务超次数或离线资源装机量超限. 如离线命令词识别中, 一个试用版的 appid 和资源在超过 10 台机器上使用.
- ☑ 其他能力应用错误  
可参照 <https://shimo.im/sheet/w3yUy39uNKs0J7DT> 进行修改.