**智能机器人语音解决方案**

1. **讯飞语音服务**



**1、语音听写：**将连续语音快速识别为文字

1）支持Linux平台开发；

2）性能指标：输入速度：180字/min；响应时间：< 500ms；

3）支持中、英、粤、藏、维；

4）支持个性化词条定制的语言模型；

5）仅有在线SDK（免费）；

**2、命令词识别：**精确识别特定的词语

1）支持Linux平台开发；

2）识别开发者自定义的命令词；

3）命令词最大长度为31个字，个数为215个；需要自己构建语法；

4）分为**在线版**和**离线版**

A：在线命令词识别：SDK免费

B：离线命令词识别：收费标准如下

【离线命令词识别按照装机量授权数收费。根据开发者购买的装机量授权数收费，可以一次购买的装机量授权数为2千、5千、1万、5万，对应的价格分别为8000元、17500元、30000元、125000元。另外，由于linux平台SDK包的制作比其它平台复杂，价格较其它平台稍高。成功付款后，即可下载SDK包。】

**3、语音合成：**将文字信息转化为语音

1）支持Linux平台开发；

2）支持语音个性定制（中英粤多语种、川豫多方言、男女声多风格的选择）

3）分为**在线**版和**离线**版

A：开发包大小：离线SDK包比在线SDK包大；

B：响应效率：离线服务核心计算在本地，减少了网络交互的时间，相比在线服务响应快；

C：效果方面：由于本地计算能力和存储空间的限制，离线资源相对小点，但保证离线服务效果

D：在线版免费，离线版收费如下：

【离线合成按照装机量授权数收费。根据开发者购买的装机量授权数收费，可以一次购买的装机量授权数为2千、5千、1万、5万，对应的价格分别为8000元、17500元、30000元、125000元。另外，由于linux平台SDK包的制作比其它平台复杂，价格较其它平台稍高。成功付款后，即可下载SDK包。

离线语音合成购买的SDK包中默认提供xiaoyan、xiaofeng两种发音人，如果想购买其他发音人资源，价格计算方法为单机成本加1，即 价格=（单机成本+1\*新增的发音人个数）\*装机量 ，比如安卓平台的离线合成基础包（装机量=2000，单机成本=4，价格8000=2000\*4），如果还需要两个发音人（比如粤语、英文），那么价格=（4+2）\*2000=12000】**2000台装机量起！**

**4、开放语义：**进行语音理解的，实现语义提取和智能问答(集语义解析、智能问答、垂直搜索和业务接入于一体，提供一站式人机智能交互技术接入平台。)

1）支持Linux平台开发；

2）必须是在线的；

**5、语音唤醒：**个性化定制唤醒词，无需触碰，直接进去

1）支持Linux；

2）应用场景举例：通过唤醒词直接操控智能家居（开启、关闭、换台等）

3）可以为应用一次定制唤醒词的个数为1个、2个、3个，对应的价格分别为2999元、3499元、3799元；

4）唤醒词支持4—6个汉字或2—6个英文字母；

**6、语音云：**云端中文自然语音处理服务平台，提供分词、词性标注、命名实体识别、依存句法分析、语义角色标注等自然语言处理服务。

备注：不需要，忽略；

**7、语音测评：**自动对发音水平进行评价

备注：不需要，忽略；

**8、声纹识别：**依据独一无二的生理特征建立精确的声纹模型，实现快速精准的身份认证

1）不支持Linux平台（只支持Android和IOS）；

2）支持三种类型声纹密码：

1、文本密码（密码为指定文本内容）

2、 自由说（密码内容不限）

3、数字密码（密码为随机数字串）。

**9、人脸识别：**提供人脸验证、人脸检测、人脸关键点检测等功能

1）不支持linux平台（只支持Android和IOS）；

1. **智能机器人语音解决方案**

**+**

**+**

**1. 离线命令词识别**

**2．在线语义**

**3. 离线合成**

**2.1语音听写 + 2.2开放语义**

**识别指令**

**语音反馈**

**语义理解+智能问答**

**1）模块介绍**

**1、“离线命令词识别”—【耳朵】**

1）预期功能：识别用户语音指令

2）可备选方案：“离线语音听写”

**离线命令词识别** *VS* **离线语音听写**

|  |  |  |  |
| --- | --- | --- | --- |
|  | 识别内容 | 识别目的 | 效果 |
| 离线命令词识别 | 开发者自定义的命令词 | 精确识别特定的词语 | 待对比 |
| 离线语音听写 | 任意输入的词语或者句子 | 将语音识别为文字 |

**2、“在线语义”—【耳朵+大脑】**

**2.1 语音听写**

1）预期功能：将连续语音转化给文本

**2.2 开放语义**

1）必须是在线的

2）预期功能：语义提取+智能回答

a语义提取示例：“打电话给张三”

{http://www.xfyun.cn/img/services/Expanded.gif

"semantic": {http://www.xfyun.cn/img/services/Expanded.gif

"slots": {http://www.xfyun.cn/img/services/Expanded.gif

"name": "张三"

}

},

"rc": 0,

"operation": "CALL",

"service": "telephone",

"text": "打电话给张三"

}

b智能回答示例：

——“你最最好了”

——“哪里啊，你过奖了。”

**3、“离线合成”—【嘴巴】**

1）预期功能：将文字转化为语音，形成语音反馈。

**2）备注说明**

1、“在线语义”已经包含了“离线命令词识别”的语音识别的功能，为什么还要“离线命令词识别”模块？

**答：**防备没有网络的时候，也能执行一些用户命令。另，有网情况下，使用语义可以实现更多的功能。

**2、待补充**

**……**

**三、存在的问题和困难**

**1、关于本方案的可行性**

上述方案所需的4个独立模块能否集成到一起待检验。这是因为不同功能的SDK里面的libmsc.so 存在不能共用的情况，有些组合可以，有些组合不可以。

**2、linux平台的SDK，没有录音模块，需要自己编写。**

**3、离线版本SDK和在线SDK版本不能同时使用。**

**4、对于长音频识别，目前讯飞语音只支持Android平台。**