语音识别

由于一个人的声音特征与他的说话内容无关，这份api可以实现从声音中提取相关信息，记录独特的声纹，并识别声纹。

每个音频应在5秒至5分钟内。可上传多个音频。有效的语音时间累积达到30秒时，才会登记成功（可以手动取消这种限制）。

一个档案（porfile）内包含一个人的一系列音频（audio）。

* 向档案添加语音（create enrollment）

1. Request URL

[https://westus.api.cognitive.microsoft.com/spid/v1.0/identificationProfiles/{identificationProfileId}/enroll[?shortAudio](https://westus.api.cognitive.microsoft.com/spid/v1.0/identificationProfiles/%7bidentificationProfileId%7d/enroll%5b?shortAudio)]

1. Request parameters
   1. identificationProfileId (string)：档案的Id号码。由“Create Profile API”（见下）建立档案。每个id唯一。
   2. shortAudio (bool，optional)：是否取消“有效时间必须达30s”这种限制。为true时则取消限制。
2. request body
   1. 你需要在request body内包含视频文件。见网页示例代码。
3. 音频必须满足以下要求
   1. Container：WAV
   2. Encoding：PCM
   3. Rate：16K
   4. Sample Format：16 bit
   5. Channels：Mono
4. 调用成功将返回
   1. Accepted字样
   2. 类似于以下的链接，以提供更多信息

<https://api.projectoxford.ai/spid/v1.0/operations/EF217D0C-9085-45D7-AAE0-2B36471B89B5>

* 创建档案（Create Profile）

1. Request URL

<https://westus.api.cognitive.microsoft.com/spid/v1.0/identificationProfiles>

1. Request body
   1. locale（string）：指定语言，可能的取值如下
      1. en-US (English US)：美式英语
      2. zh-CN (Chinese Mandarin)：普通话
2. 返回值
   1. identificationProfileId（string）：档案的Id。用于其他接口。

* 删除档案

1. Request URL

[https://westus.api.cognitive.microsoft.com/spid/v1.0/identificationProfiles/{identificationProfileId}](https://westus.api.cognitive.microsoft.com/spid/v1.0/identificationProfiles/%7bidentificationProfileId%7d)

1. Request parameters
   1. identificationProfileId （string）：被删除的档案id

* 得到（你创建的）所有档案（get all profile）

1. Request URL

<https://westus.api.cognitive.microsoft.com/spid/v1.0/identificationProfiles>

1. 返回值（JSON）：一个大数组，元素是object，包含以下元素。
   1. identificationProfileId（string）：档案的id
   2. locale（string）：对应的语言
   3. enrollmentSpeechTime（number）：音频的有效语音时间
   4. createdDateTime（string）：创建时间
   5. lastActionDateTime（string）：最后使用这份档案的时间
   6. EnrollmentStatus（string）：包含以下取值
      1. Enrolling：有效语音时长不足，需要更多数据资料。（见上）
      2. Training：正在训练模型，请等待一段时间
      3. Enrolled：可以进行工作

* 得到一个档案（get profile）

1. Request URL

[https://westus.api.cognitive.microsoft.com/spid/v1.0/identificationProfiles/{identificationProfileId}](https://westus.api.cognitive.microsoft.com/spid/v1.0/identificationProfiles/%7bidentificationProfileId%7d)

1. identificationProfileId （string）：（同上）
2. 返回值（同get all profile）

* 档案初始化（reset enrollment）

将会清空从服务器上永久清空档案内的数据。

1. Request URL

[https://westus.api.cognitive.microsoft.com/spid/v1.0/identificationProfiles/{identificationProfileId}/reset](https://westus.api.cognitive.microsoft.com/spid/v1.0/identificationProfiles/%7bidentificationProfileId%7d/reset)

1. equest parameters
   1. identificationProfileId （string）：（同上）

* 语音识别（Identification）

自动识别一份音频内是谁在讲话。

视频要求见上文。

1. Request URL

[https://westus.api.cognitive.microsoft.com/spid/v1.0/identify?identificationProfileIds={identificationProfileIds}[&shortAudio](https://westus.api.cognitive.microsoft.com/spid/v1.0/identify?identificationProfileIds=%7bidentificationProfileIds%7d%5b&shortAudio)]

1. 参数
   1. identificationProfileIds （string）：字符串包含一组由逗号隔开的profileId。数目不超过10个。识别这些人中的哪位在视频中讲话了。
   2. shortAudio（bool，optional）：（见上同名变量）
2. request body
   1. 你需要在request body内包含视频文件。见网页示例代码。
3. 返回值
   1. 类似于以下的链接，以提供更多信息https://api.projectoxford.ai/spid/v1.0/operations/EF217D0C-9085-45D7-AAE0-2B36471B89B5

* 获得识别信息（接上一条目）

1. Request URL
   1. [https://westus.api.cognitive.microsoft.com/spid/v1.0/operations/{operationId}](https://westus.api.cognitive.microsoft.com/spid/v1.0/operations/%7boperationId%7d)
2. Request parameters
   1. operationId（ string）：由identification或create enrollment的相关信息得到
3. 返回值（JSON）
   1. status（string）：包含以下状态
      1. notstarted：未开始运行
      2. running：正在运行
      3. failed：执行完毕，失败
      4. succeeded：执行完毕，成功
   2. createdDateTime（string）：见上同名变量
   3. lastActionDateTime（string）：见上同名变量
   4. message（string）：当status为failed时，该变量包含与错误有关的信息。
   5. processingResult（object）：当status为success时，该object储存有关信息
   6. EnrollmentStatus（string）：见上同名变量
   7. enrollmentSpeechTime（number）：语音总时长（单位秒）
   8. identifiedProfileId（string）：匹配到的人的id。若该值为” 00000000-0000-0000-0000-000000000000”则说明匹配不到候选的任何一个人。
   9. confidence（string）：准确度，有以下取值
      1. Low
      2. Normal
      3. High