

“Webとクルマのハッカソン2018”

mimi® ASR API ご紹介

mimi® ASR HTTP API 仕様

- ・このAPIは、サーバーに音声データを送信すると、音声データ中の人間の声（日本語）を、単語ごとに認識する音声認識機能を提供するバッチAPIです。
- ・認識途中の結果を得ることはできません。
- ・音声認識時間処理に要する時間は、入力音声長の約半分となります。つまり、10秒の音声を入力した場合、約5秒の認識時間が掛かることになります。
- ・音声ファイル等の認識処理に好適です。
- ・マイク入力等から逐次的に認識したい場合はmimi® ASR WebSocket API（リアルタイムAPI）をご利用ください。
- ・詳細は利用ガイドをご一読ください。

出力JSON形式

```
{
  "response": [
    {"pronunciation": "ヨミカナ1", "result": "認識単語1", "time": [開始時刻1, 終了時刻1]},
    {"pronunciation": "ヨミカナ2", "result": "認識単語2", "time": [開始時刻2, 終了時刻2]},
    ... 以下文章の終わりまで単語毎の認識結果が続く ...
  ],
  "session_id": "結果のユニークID",
  "status": "recog-finished",
  "type": "asr#mimilvcsr"
}
```

実行結果例

```
{
  "response": [
    {"pronunciation": "ローニャク", "result": "老若", "time": [ 300, 810 ] },
    {"pronunciation": "ナンニョ", "result": "男女", "time": [ 810, 1180 ] },
    {"pronunciation": "ガ", "result": "が", "time": [ 1180, 1390 ] },
    {"pronunciation": "ヒ", "result": "火", "time": [ 1650, 1890 ] },
    {"pronunciation": "ヲ", "result": "を", "time": [ 1890, 1960 ] },
    {"pronunciation": "カコン", "result": "困ん", "time": [ 1960, 2350 ] },
    {"pronunciation": "デ", "result": "で", "time": [ 2350, 2460 ] },
    {"pronunciation": "ノミ", "result": "のみ", "time": [ 2460, 2840 ] },
    {"pronunciation": "テ", "result": "手", "time": [ 3100, 3250 ] },
    {"pronunciation": "ヲ", "result": "を", "time": [ 3250, 3350 ] },
    {"pronunciation": "ツナイ", "result": "つない", "time": [ 3350, 3700 ] },
    {"pronunciation": "デ", "result": "で", "time": [ 3700, 3850 ] },
    {"pronunciation": "ウタウ", "result": "歌う", "time": [ 3850, 4310 ] },
    "session_id": "8c0f9fdc-00da-11e8-b96e-42010af00006",
    "status": "recog-finished",
    "type": "asr#mimilvcsr"
  ]
}
```