

IBM Watson

Speech to Text ツールマニュアル

日本アイ・ビー・エム株式会社

IBM



目次

1. 辞書のカスタマイズ

1. 辞書の作成と削除
2. 単語の新規登録と削除
3. csvファイルのフォーマット
4. コーパスの新規登録と削除
5. txtファイルのフォーマット
6. 辞書のトレーニング

2. 音声認識

1. wavファイルの音声認識
2. マイク入力の音声認識
3. カスタム辞書の使用

3. トラブルシューティング

1. FAQ
2. リンク集

ツール概要：カスタマイズされた辞書を使って音声認識が可能

Speech To Text Customization

endPoint

username password

Manage custom language models

Name

Description

List Models

Delete

Delete All

Name: Test, ID: ea9ae880-f4d9-11e6-b51c-5df425e9097c
Name: Test, ID: 19937c50-f4db-11e6-a3e6-0f8cc95f9cb2
Name: Test, ID: 37d853e0-f4de-11e6-bbca-09398e65745e
Name: Test, ID: 3c3bd510-f4de-11e6-b5a7-9ff011559050
Name: Test, ID: e0487a60-f5b9-11e6-b5a7-9ff011559050

Select Model

csv File

Add Words

Get Words

Delete Words

Corpus txt File

Corpus Name

Add Corpus

List Corpora

Delete Corpus

Training

Status
カスタマイズまたはトレーニングを行う辞書が設定されました。

Speech To Text

endpoint

username password

model

☒ Use Customization ☐ Output text file (only wav file)

wav File

[0.61]金沢 工業 大学 は と 呼ばれ ます

Status
音声認識が完了しました。

Upload

Stop mic input

[0.61]金沢 工業 大学 は と 呼ばれ ます

Status
音声認識が完了しました。

1. 辞書のカスタマイズ



1-1. 辞書の新規作成と削除

1. インターネットへ接続されていることを確認し、クレデンシャル情報が正しいものか確認
2. Name, Descriptionを入力後、Createをクリックして新規作成
3. List Modelsで新しく作成された辞書を確認
4. 任意の辞書を選択してDeleteをクリックすると削除可能
5. Delete Allで全ての辞書を削除

The screenshot shows the 'Speech To Text Customization' window. It includes fields for 'endPoint', 'username', and 'password'. Below these is a section titled 'Manage custom language models' with input fields for 'Name' and 'Description', and a 'Create' button. A list of models is displayed, with one model selected and highlighted in blue. At the bottom, there are buttons for 'List Models', 'Delete', 'Delete All', 'Select Model', and 'Get Status'.

① endPoint
username password

② Manage custom language models
Name
Description

③ Name: Test, ID: ea9ae880-f4d9-11e6-b51c-5df425e9097c
Name: Test, ID: 19937c50-f4db-11e6-a3e6-0f8cc95f9cb2
Name: Test, ID: 37d853e0-f4de-11e6-bbca-09398e65745e
Name: Test, ID: 3c3bd510-f4de-11e6-b5a7-9ff011559050
Name: Test, ID: e0487a60-f5b9-11e6-b5a7-9ff011559050

④

⑤

1-2. 単語の登録と削除

1. カスタマイズを行いたい辞書をリストから選択しSelect Modelをクリック
2. csv Fileをクリックし、単語登録に使うcsvファイルを選択（次項参照）
3. Add Wordsをクリックして辞書に単語を登録
4. Get Wordsで登録された単語の確認、Delete Wordsで削除が可能

The screenshot shows the IBM Watson interface for word registration and deletion. The interface is divided into several sections:

- List Models:** A table listing models with columns for Name and ID. The second model is selected. **Annotations:** 1 points to the 'List Models' button, 2 points to the 'Delete' button, 3 points to the 'Delete All' button, and 6 points to the selected model row.
- Select Model:** A dropdown menu showing the selected model ID: 19937c50-f4db-11e6-a3e6-0f8cc95f9cb2. **Annotation:** 4 points to the 'Select Model' button.
- csv File:** A text input field containing the path: :orpora/addWord_test_utf.csv. **Annotation:** 5 points to the 'csv File' button.
- Add Words:** A text input field containing the words: BKT,[ボッケチョウ], IBM,[アイビーエム], 金工大,[キンケイダイ, キンコウタ]. **Annotation:** 3 points to the 'Add Words' button.
- Get Words:** A button to retrieve the words from the corpus.
- Delete Words:** A button to delete words from the corpus.
- Corpus txt File:** A text input field for the corpus file path.
- Corpus Name:** A text input field for the corpus name.
- Add Corpus:** A button to add a new corpus.
- List Corpora:** A button to list the existing corpora.
- Delete Corpus:** A button to delete a corpus.
- Training:** A button to start training the model.
- Status:** A section showing the current status: 現在 3 個の単語が登録されています。 (Currently, 3 words are registered).

1-3. 単語登録用csvファイルのフォーマット

- 文字コードはUTF-8のみ可能
 - エクセルまたは任意のテキストエディタで作成
- 列は下記の順番を厳守
 - 一列目(Word)
 - > 表示させたい文字列を入力
 - 二列目以降(SoundsLike)
 - > 認識させたい「ヨミ」を全角カタカナで入力
 - > 2種類以上の場合、右に不定長で並べる

	A	B	C
1	金工大	キンコウダイ	
2	IBM Japan	ニホンアイピーエム	アイピーエムジャパン
3			
4			
5			

```
addWord_test_utf.csv *
1 金工大,キンコウダイ,キンケイダイ
2 IBM,アイピーエム,
3 BKT,ボッケチョウ,
```

1-4. コーパスの登録と削除

1. カスタマイズを行いたい辞書をリストから選択しSelect Modelをクリック
2. Corpus txt Fileをクリックし、単語登録に使うtxtファイルを選択（次項参照）
3. Corpus Nameに任意の名前を入力し、Add Corpusをクリックして辞書に単語を登録
4. List Corporaで登録されたコーパスの名前と状態を確認、Delete Corpusで任意コーパスの削除が可能
5. List Corporaでコーパスの状態を確認し、BEING_PROCESSEDからANALYZEDになれば辞書のトレーニングが可能

LIST MODELS

Name: Test , ID: 19937c50-f4db-11e6-a3e6-0f8cc95f9cb2
Name: Test , ID: 37d853e0-f4de-11e6-bbca-09398e65745e
Name: Test , ID: 3c3bd510-f4de-11e6-b5a7-9ff011559050
Name: Test , ID: e0487a60-f5b9-11e6-b5a7-9ff011559050

Delete

Delete All

① Select Model 19937c50-f4db-11e6-a3e6-0f8cc95f9cb2 Get Status TRAINING

csv File :orpora/addWord_test_utf.csv

Add Words BKT,[ボッケチョウ]
IBM,[アイビーエム]
金工大,[キンケイダイ, キンコウタ]

Get Words

Delete Words

② Corpus txt File _and_corpora/corpas_kit1_utf.txt

Corpus Name kit1

③ Add Corpus kit1, BEING_PROCESSED

④ List Corpora

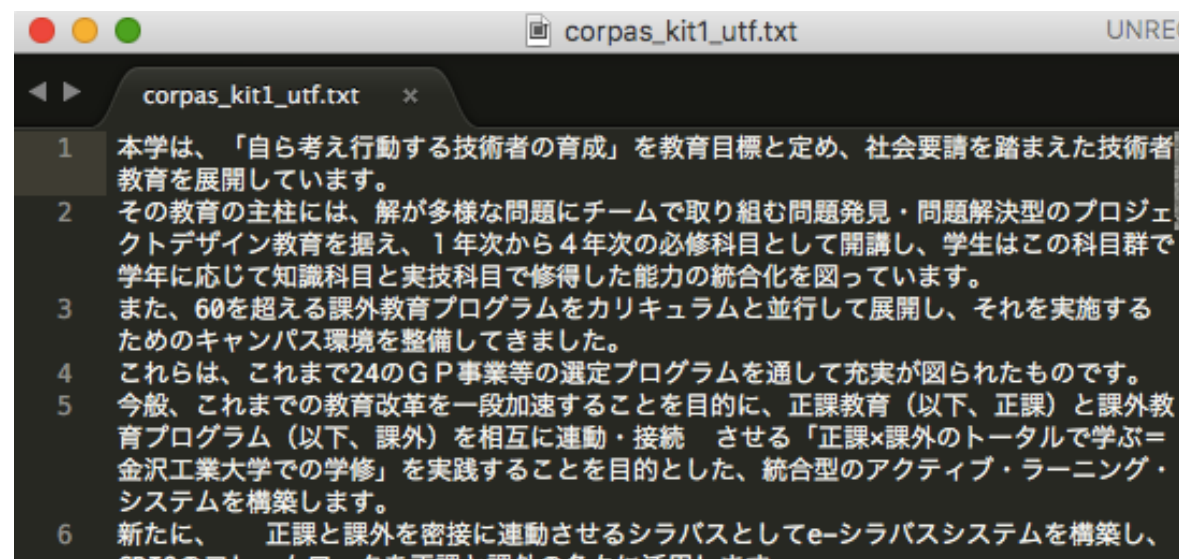
Delete Corpus

Status
現在 1 個のコーパスが登録されています。

Training

1-5. コーパス登録用txtファイルのフォーマット

- 文字コードはUTF-8のみ可能
 - 任意のテキストエディタで作成
- 一つの行に一つの文章とすると精度が上がる
(必須ではない)



1-6. 辞書のトレーニング

1. 単語、コーパスを新たに登録したり削除した場合、Trainingをクリックすることで辞書をトレーニングする（リクエストに数十秒かかる場合もある）
2. トレーニング後にGet Statusでモデルの状態を確認し、AVAILABLEになっていれば音声認識に利用可能

List Models

Delete

Delete All

Name: Test , ID: ea9ae880-f4d9-11e6-b51c-5df425e9097c
Name: Test , ID: 19937c50-f4db-11e6-a3e6-0f8cc95f9cb2
Name: Test , ID: 37d853e0-f4de-11e6-bbca-09398e65745e
Name: Test , ID: 3c3bd510-f4de-11e6-b5a7-9ff011559050
Name: Test , ID: e0487a60-f5b9-11e6-b5a7-9ff011559050

Select Model

19937c50-f4db-11e6-a3e6-0f8cc95f9cb2

Get Status

csv File

:orpora/addWord_test_utf.csv

Add Words

BKT,[ボッケチョウ]
IBM,[アイビーエム]
金工大,[キンケイダイ, キンコウタ]

Get Words

Delete Words

Corpus txt File

Corpus Name

Add Corpus

List Corpora

Delete Corpus

Training

Status

現在 3 個の単語が登録されています。

2. 音声認識



2-1. wavファイルの音声認識

1. インターネットへ接続されていることを確認し、endpoint, username, passwordを入力
2. 認識言語をModelから選択（日英）
3. wav Fileをクリックし、音声認識に使うwavファイルを選択
4. Uploadをクリック
5. 辞書を使う場合はUse customizationにチェックを入れて辞書IDを入力
6. 書き起こしの結果を自動で保存したい場合はOutput text fileにチェックを入れる（wavファイル名の拡張子がtxtになったファイルが同じディレクトリに作成される）

The screenshot shows the 'Speech To Text' configuration window. Red circles with numbers 1 through 6 are placed over specific fields and buttons to correspond with the numbered steps in the list on the left.

- ① points to the **endpoint** field containing `https://stream.watson-jp/speech-to-text/api`.
- ② points to the **model** dropdown menu, which is currently set to `ja-JP_BroadbandModel`.
- ③ points to the **wav File** button.
- ④ points to the **Upload** button.
- ⑤ points to the **Use Customization** checkbox, which is checked. Next to it is a field containing the customization ID `19937c50-f4db-11e6-a3e6-0f8cc95f9`.
- ⑥ points to the **Output text file (only wav file)** checkbox, which is unchecked.

Below the configuration fields, there are two buttons: **Start mic input** and **Stop mic input**. At the bottom, a text area displays the transcription result: `[0.61]金沢 工業 大学 は 錦江台 と 呼ばれ ます`.

2-2. マイク入力の音声認識

1. インターネットへ接続されていることを確認し、endpoint, username, passwordを入力
2. 認識言語をModelから選択（日英）
3. Start mic inputをクリックし、音声認識を開始
4. 認識が終了したらStop mic inputをクリックして認識を終了する
5. 辞書を使う場合はUse customizationにチェックを入れて辞書IDを入力

注意点

- 4. の後に次の動作を行うと書き起こされた文字列が失われる
- マイク入力の場合、テキストファイルの自動保存ができない
- 書き起こした文字列を保存したい場合は、「メモ帳」などの別ソフトに貼り付けて保存する

Speech To Text

① endpoint

② username password

⑤ model

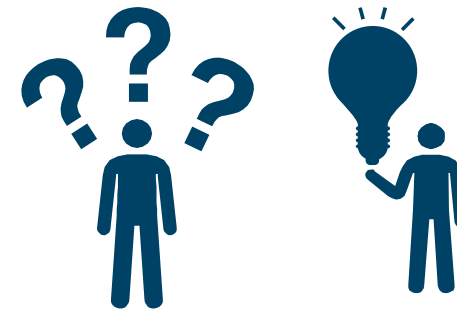
☒ Use Customization ☐ Output text file (only wav file)

wav File ③

④

[2.16]日本 IBM
[8.04]金沢 工業 大学
[11.83]金工大 と 呼ばれ ます
[15.14]マイク 入力 の テスト を 終了 します

3. トラブルシューティング

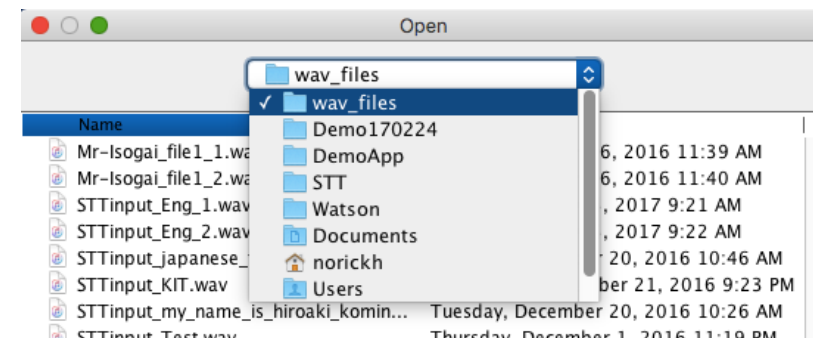


3-1. FAQ (1/2)

- 辞書は最大何個まで登録できますか？
 - 制限はありませんが、一度に使えるのは一つの辞書だけです。
- 単語は最大何個まで登録できますか？
 - 最大3000個です。
- コーパスは最大何個まで登録できますか？
 - 個数に制限はないですが、全てのコーパスを合わせて1000万字を超えないようにしてください。詳細は次をご覧ください。<https://www.ibm.com/watson/developercloud/doc/speech-to-text/custom.shtml#prepareCorpus>
- Add Wordsでsynchronized errorが表示されます
 - 通信エラーです。できるまで繰り返してください。
- 単語を個別に削除したいのですが
 - その機能はサポートしていないので、一度全ての単語を削除してから新しくcsvを登録してください。

3-1. FAQ (2/2)

- wav音声ファイルを途中で認識しなくなっていました。
 - 一度にアップロードできるのは100MBまでなので、分割して読み込ませてください。
- ファイルを選択するときのウィンドウ上で、上位のディレクトリに行くにはどうしたらいいですか。
 - 現在のディレクトリが表示されている部分をクリックしてみてください。（右図）
- その他動作がおかしくなった時
 - インターネットの接続を確認してください。
 - アプリを再起動してください。



3-2. リンク集

- 単語登録の仕様書
 - <https://www.ibm.com/watson/developercloud/doc/speech-to-text/custom.shtml#addWords>
- コーパス登録の仕様書
 - <https://www.ibm.com/watson/developercloud/doc/speech-to-text/custom.shtml#addCorpora>
- APIリファレンス
 - APIのリクエストとレスポンスの仕様書です
 - <https://www.ibm.com/watson/developercloud/speech-to-text/api/v1/>
- API Explorer
 - GUIでcURLを使ったリクエストとレスポンスを観測できるツールです
 - <https://watson-api-explorer.mybluemix.net/apis/speech-to-text-v1>
- Watson Developer Tools
 - 上記を含むリンク集で、SDKやコミュニティを参照することができます
 - <https://www.ibm.com/watson/developercloud/developer-tools.html>