

用語抽出

用語抽出について

用語抽出は文章から用語を抽出するツールです。

用語抽出結果に対して用語集や自動翻訳を用いて訳語を提示することも可能です。

用語抽出の種類について

用語抽出には、「単言語」、「対訳」の2種類があります。

単言語は、原テキストから用語を抽出します。

対訳は、対訳テキストから用語を抽出します。

単言語用語抽出について

単言語用語抽出は、単言語のデータファイル(例えば、日本語の文章だけのファイル)に対して用語抽出を行います。

また登録時に、用語集と自動翻訳を指定することで、抽出結果に訳テキストを追加することができます。

対訳用語抽出について

対訳言語用語抽出は、対訳データファイル(例えば、日本語と英語の文章のファイル)に対して用語抽出を行います。

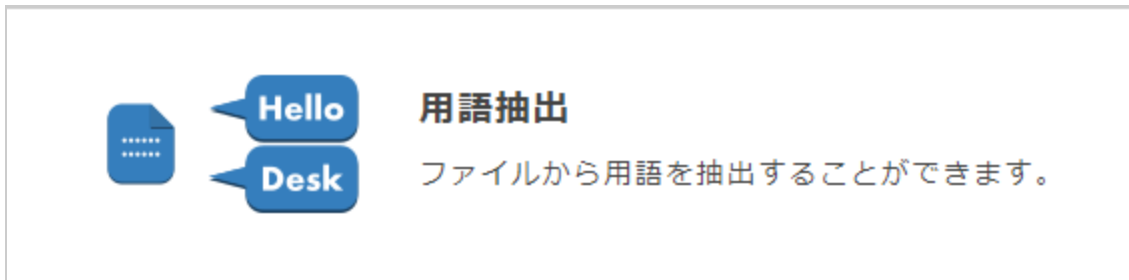
また登録時に、用語集と自動翻訳を指定することで、抽出結果に訳テキストが存在しない場合、訳テキストを追加することもできます。

用語抽出の登録

用語抽出の登録を行います。

(1) メニューから[便利アプリ]を選択します。

(2) [用語抽出]を選択します。



(3) [新規登録]ボタンを押します。



(4) フォームの各項目を入力していきます。

Step.1 基本情報を入力してください。

項目	説明
名前	用語抽出の名前を設定します。
言語	用語抽出の言語を設定します。 ※単言語用語抽出は「言語」、対訳用語抽出は「言語方向」を選択します。
ファイル	用語抽出で使用する対象ファイルを選択します。

Step.2 オプションを選択してください。

項目	説明
----	----

訳語提示用語集	抽出結果の原テキストに対して用語集と比較し、マッチした場合、結果に訳テキストを追加します。
---------	---

訳語提示自動翻訳	抽出結果の原テキストに対して自動翻訳を行い、結果に訳テキストを追加します。
----------	---------------------------------------

(5) 全て入力し終わったら[OK]ボタンを押します。

登録が成功すると、「待機中」もしくは「処理中」状態になります。

※登録には時間が掛かる場合があります。

ヒント

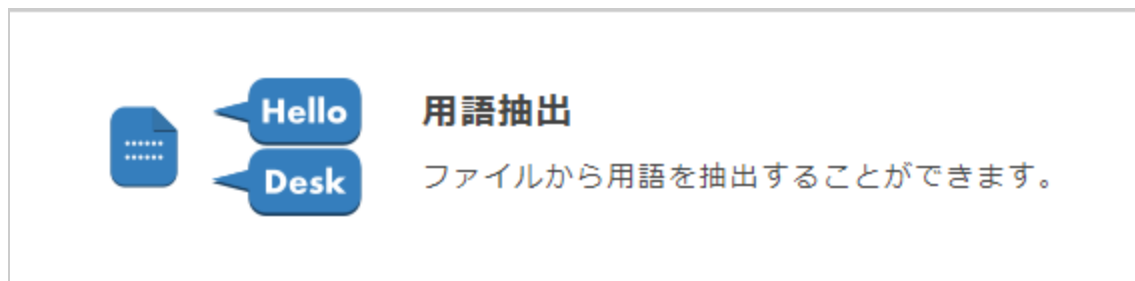
ファイル形式は、単言語か対訳かによって異なります。

用語抽出の削除

(1) メニューから[便利アプリ]を選択します。



(2) [用語抽出]を選択します。



(3) [削除]ボタンを押します。

操作



(4) 確認ダイアログが表示されますので、[OK]ボタンを押します。

成功すると、成功画面が表示されます。

⚠ 注意

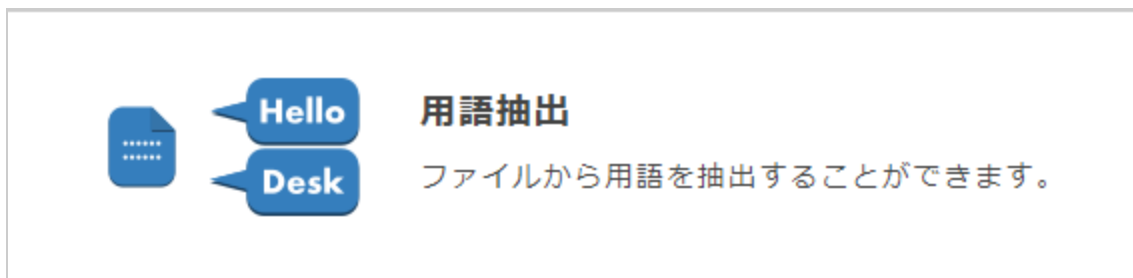
一度削除すると、元に戻すことはできません。

用語抽出のダウンロード

(1) メニューから[便利アプリ]を選択します。



(2) [用語抽出]を選択します。



(3) [ダウンロード]ボタンを押します。



成功するとTSV形式で出力されます。

※ダウンロードファイルの内容は、「用語抽出の結果について」で説明します。

用語抽出の結果について

用語抽出の結果は以下のTSV形式で出力されます。

項目

説明

出現頻度	用語が使われている出現頻度を表します。
------	---------------------

原テキスト	対象となる文章の原言語が出力されます。
-------	---------------------

訳テキスト	対象となる文章の訳言語が出力されます。
-------	---------------------

注意

文字コードはUTF-8です。

用語抽出のエラーについて

用語抽出の結果がエラーと表示される場合、出力データが空の場合は、以下項目を確認してみてください。

- 入力したファイルの文字コードがUTF-8であるか？
 - 言語方向を設定しているが、どちらかのテキストの入力が空データではないか？
 - 言語方向を設定しているが、テキストデータの方向が逆ではないか？
-