

# Python MeCab( 日本語形態素解析 )

[Python][ 自然言語処理 ]

- ・ <http://taku910.github.io/mecab/>

## インストール

- ・ <http://d.hatena.ne.jp/atree/20090512/1242122536>

### ■ janome( 辞書内包 / Pure Python 実装の形態素解析器 )

- ・ <http://mocobeta-backup.tumblr.com/post/115843098157/pure-python-janome>
- ・ <http://mocobeta.github.io/janome/>
- ・ MeCab の辞書を利用

### ■ MeCab

```
# wget http://mecab.googlecode.com/files/mecab-0.996.tar.gz
# tar -xvf http://mecab.googlecode.com/files/mecab-0.996.tar.gz
# cd mecab-0.996
# ./configure --with-charset=utf8
# make
# make install
```

./configure で、Your compiler is not powerful enough to compile MeCab エラーとなったら、gcc-c++ をインストール

```
# yum install gcc-c++
```

### ■ 辞書のインストール

```
# wget http://mecab.googlecode.com/files/mecab-ipadic-2.7.0-20070801.tar.gz
# tar -xvf http://mecab.googlecode.com/files/mecab-ipadic-2.7.0-20070801.tar.gz
# cd mecab-ipadic-2.7.0-20070801
# ./configure --with-charset=utf8
# make
# make install
```

### ■ Python バインディングのインストール

```
# wget http://mecab.googlecode.com/files/mecab-python-0.996.tar.gz
# tar -xvf mecab-python-0.996.tar.gz
# cd mecab-python-0.996
# python2.7 setup.py install
```

- ・ python2.7 としているのは、[Python2.7 をインストールしているため](#)

### ■ ImportError: libmecab.so.2

以下のエラーがでる場合

```
# python2.7
>>> import MeCab
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
  File "MeCab.py", line 26, in <module>
```

```

    MeCab = swig_import_helper()
    File "MeCab.py", line 18, in swig_import_helper
    import MeCab
ImportError: libmecab.so.2: cannot open shared object file: No such file or directory

```

## ライブラリを読み込ませるように設定

```

# cd /etc/ld.so.conf.d
# vi mecab.conf

```

- ・以下の内容を mecab.conf に記述

```

/usr/local/lib

```

- ・読み込み

```

# ldconfig

```

## 実行してみる

```

# python2.7
>>> import MeCab
>>> mecab = MeCab.Tagger('-Ochasen')
>>> print mecab.parse('泣き叫ぶ鬼の串刺し。何かに憑かれた餓鬼の群れ')
泣き叫ぶ ナキサケブ 泣き叫ぶ 動詞 - 自立 五段・バ行 基本形
鬼 オニ 鬼 名詞 - 一般
の ノ の 助詞 - 連体化
串刺し クシザシ 串刺し 名詞 - 一般
。 。 記号 - 句点
何 ナニ 何 名詞 - 代名詞 - 一般
か カ か 助詞 - 副助詞 / 並立助詞 / 終助詞
に ニ に 助詞 - 格助詞 - 一般
憑かれ ツカレ 憑かれる 動詞 - 自立 一段 連用形
た タ た 助動詞 特殊・タ 基本形
餓鬼 ガキ 餓鬼 名詞 - 一般
の ノ の 助詞 - 連体化
群れ ムレ 群れ 名詞 - 一般
EOS

```