

In [2]:

```
#from stanfordcorenlp import StanfordCoreNLP
from nltk.tag import StanfordPOSTagger
```

In [3]:

```
# jar = "/Users/philip/Desktop/Filipino-Stanford-Part-of-Speech-Tagger/stanford-pos
tagger.jar"
jar = "/Users/philip/Desktop/intro-to-nlp/[HW4] POS TAGGER/stanford-postagger.jar"
model = "/Users/philip/Desktop/intro-to-nlp/[HW4] POS TAGGER/filipino-left5words-ow
lqn2-distsim-pref6-inf2.tagger"
```

In [4]:

```
pos_tagger = StanfordPOSTagger(model, jar, encoding = "UTF-8")
#, encoding='utf8'
```

model

In [5]:

```
pos_tagger?
```

```
Type:          StanfordPOSTagger
String form:   <nltk.tag.stanford.StanfordPOSTagger object at 0x7fec1d4c3
450>
File:          ~/opt/anaconda3/lib/python3.7/site-packages/nltk/tag/stanf
ord.py
Docstring:
A class for pos tagging with Stanford Tagger. The input is the paths t
o:
- a model trained on training data
- (optionally) the path to the stanford tagger jar file. If not specif
ied here,
  then this jar file must be specified in the CLASSPATH environment va
riable.
- (optionally) the encoding of the training data (default: UTF-8)
```

Example:

```
>>> from nltk.tag import StanfordPOSTagger
>>> st = StanfordPOSTagger('english-bidirectional-distsim.tagger')
>>> st.tag('What is the airspeed of an unladen swallow ?'.split())
[('What', 'WP'), ('is', 'VBZ'), ('the', 'DT'), ('airspeed', 'NN'),
('of', 'IN'), ('an', 'DT'), ('unladen', 'JJ'), ('swallow', 'VB'),
('?', '.')]

```

In [6]:

```
pos_tagger.tag('Kahit maputi na ang buhok ko'.split())
```

Out[6]:

```
[('', 'Kahit|CCT'),  
 ('', 'maputi|JJD'),  
 ('', 'na|CCP'),  
 ('', 'ang|DTC'),  
 ('', 'buhok|NNC'),  
 ('', 'ko|PRS')]
```

Demo Day 08/31/2021

In [10]:

```
pos_tagger.tag("Pumunta sa palengke ang Hapon.".split())
```

Out[10]:

```
[('', 'Pumunta|VBAF'),  
 ('', 'sa|CCT'),  
 ('', 'palengke|NNC'),  
 ('', 'ang|DTC'),  
 ('', 'Hapon.|NNPA')]
```

In [13]:

```
pos_tagger.tag("Pumunta sa palengke ang Hapon kaninang hapon.".split())
```

Out[13]:

```
[('', 'Pumunta|VBAF'),  
 ('', 'sa|CCT'),  
 ('', 'palengke|NNC'),  
 ('', 'ang|DTC'),  
 ('', 'Hapon|NNP'),  
 ('kaninang|NNC', 'CCP'),  
 ('', 'hapon.|NNC')]
```

In [12]:

```
pos_tagger.tag("Hapong-hapo ang Hapon na pumunta sa palengke kaninang hapon.".split  
())
```

Out[12]:

```
[(' ', 'Hapong-hapo|JJD'),  
 (' ', 'ang|DTC'),  
 (' ', 'Hapon|NNP'),  
 (' ', 'na|CCP'),  
 (' ', 'pumunta|VBAF'),  
 (' ', 'sa|CCT'),  
 (' ', 'palengke|NNC'),  
 (' ', 'kaninang|NNC'),  
 (' ', 'hapon.|NNC')]
```

In []: