```
# pip install -U spacy

# !python -m spacy download en_core_web_sm

import spacy

from spacy import displacy

nlp = spacy.load('en_core_web_sm')

some_text = "Peter is a great guy. He works with Google. He is a good citizen of Nigeria in the western part of Africa. He likes sir

some_text
```

```
'Peter is a great guy. He works with Google. He is a good citizen of Nigeria in the western part of Africa. He likes singing.
He studied Biology in Oxford University and he takes his studies very serious'
```

```
document = nlp(some_text)

document
```

```
Peter is a great guy. He works with Google. He is a good citizen of Nigeria in the western part of Africa. He likes singing. H
```

Sentence Detection

```
list(document.sents)
```

```
[Peter is a great guy.,
```

```
 He works with Google.,
 He is a good citizen of Nigeria in the western part of Africa.,
 He likes singing.,
 He studied Biology in Oxford University and he takes his studies very serious]
```

```
list(document.sents)[1]
```

```
 He works with Google.
```

```
list(document.sents)[2:4]
```

```
 [He is a good citizen of Nigeria in the western part of Africa.,
  He likes singing.]
```

```
for k in document.sents:
  print(k)
  # print("\n")
```

```
 Peter is a great guy.
 He works with Google.
 He is a good citizen of Nigeria in the western part of Africa.
 He likes singing.
 He studied Biology in Oxford University and he takes his studies very serious
```

Tokenization

```
for token in document:
  print(token.text)
```

```
 Peter
 is
 a
 great
 guy
 .
```

```
He
works
with
Google
.
He
is
a
good
citizen
of
Nigeria
in
the
western
part
of
Africa
.
He
likes
singing
.
He
studied
Biology
in
Oxford
University
and
he
takes
his
studies
very
serious
```

Part of Speech

```
for token in document:
  print(token.text, token.pos_)
```

```
    Peter PROPN
    is AUX
    a DET
    great ADJ
    guy NOUN
    . PUNCT
    He PRON
    works VERB
    with ADP
    Google PROPN
    . PUNCT
    He PRON
    is AUX
    a DET
    good ADJ
    citizen NOUN
    of ADP
    Nigeria PROPN
    in ADP
    the DET
    western ADJ
    part NOUN
    of ADP
    Africa PROPN
    . PUNCT
    He PRON
    likes VERB
    singing VERB
    . PUNCT
    He PRON
    studied VERB
    Biology PROPN
    in ADP
    Oxford PROPN
    University PROPN
    and CCONJ
    he PRON
    takes VERB
```

```
his PRON
studies NOUN
very ADV
serious ADJ
```

Dependency Tagging

```
for token in document:
  print(token.text, token.dep_)
```

```
Peter nsubj
is ROOT
a det
great amod
guy attr
. punct
He nsubj
works ROOT
with prep
Google pobj
. punct
He nsubj
is ROOT
a det
good amod
citizen attr
of prep
Nigeria pobj
in prep
the det
western amod
part pobj
of prep
Africa pobj
. punct
He nsubj
likes ROOT
singing xcomp
. punct
He nsubj
```

```
studied ROOT
Biology dobj
in prep
Oxford compound
University pobj
and cc
he nsubj
takes conj
his poss
studies dobj
very advmod
serious oprd
```

Lemmatization and Stemming

```
for token in document:
  print(token.text, token.lemma_)
```

```
Peter Peter
is be
a a
great great
guy guy
. .
He he
works work
with with
Google Google
. .
He he
is be
a a
good good
citizen citizen
of of
Nigeria Nigeria
in in
the the
western western
```

```
part part
of of
Africa Africa
. .
He he
likes like
singing singe
. .
He he
studied study
Biology Biology
in in
Oxford Oxford
University University
and and
he he
takes take
his his
studies study
very very
serious serious
```

Stop Words

```
for token in document:
  print(token.text, token.is_stop)
```

```
Peter False
is True
a True
great False
guy False
. False
He True
works False
with True
Google False
. False
He True
```

```
     is True
     a True
     good False
     citizen False
     of True
     Nigeria False
     in True
     the True
     western False
     part True
     of True
     Africa False
     . False
     He True
     likes False
     singing False
     . False
     He True
     studied False
     Biology False
     in True
     Oxford False
     University False
     and True
     he True
     takes False
     his True
     studies False
     very True
     serious True


for token in document:
  if (token.is_stop == False):
    print(token.text, token.is_stop)

     Peter False
     great False
     guy False
     . False
     works False
```

```
Google False
. False
good False
citizen False
Nigeria False
western False
Africa False
. False
likes False
singing False
. False
studied False
Biology False
Oxford False
University False
takes False
studies False
```

Shape

```
for token in document:
  print(token.text, token.shape_)
```

```
Peter Xxxxx
is xx
a x
great xxxx
guy xxx
. .
He Xx
works xxxx
with xxxx
Google Xxxxx
. .
He Xx
is xx
a x
good xxxx
citizen xxxx
```

```
of xx
Nigeria Xxxxx
in xx
the xxx
western xxxx
part xxxx
of xx
Africa Xxxxx
. .
He Xx
likes xxxx
singing xxxx
. .
He Xx
studied xxxx
Biology Xxxxx
in xx
Oxford Xxxxx
University Xxxxx
and xxx
he xx
takes xxxx
his xxx
studies xxxx
very xxxx
serious xxxx
```

Named Entity Recognition (NER)

```
for ent in document.ents:
  print(ent.text, ent.label_)
```

```
Peter PERSON
Google ORG
Nigeria GPE
Africa LOC
Biology ORG
Oxford University ORG
```

```
for ent in document.ents:
  print(ent.text, ent.label_, ent.start_char, ent.end_char)
```

```
Peter PERSON 0 5
Nigeria GPE 68 75
Africa LOC 99 105
Biology PERSON 136 143
Oxford University ORG 147 164
```

Word Similarities

```
some_text1 = nlp("Peter is great")
some_text2 = nlp("Kunle loves egg")
some_text3 = nlp("cat and dog")
```

```
some_text1.similarity(some_text2)
```

```
/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:1: UserWarning: [W007] The model you're using has no word vectors
  """Entry point for launching an IPython kernel.
0.3702053396464464
```

```
some_text2.similarity(some_text3)
```

```
/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:1: UserWarning: [W007] The model you're using has no word vectors
  """Entry point for launching an IPython kernel.
0.19990374791874288
```

```
displacy.serve(document, style="dep") # ent, dep
```

```
Using the 'ent' visualizer
```

```
Serving on http://0.0.0.0:5000 ...

Shutting down server on port 5000.
```