

##Assignment No.03##

#Name: Radhika Santosh Darode

##Roll No: 16

#Batch: B1

#Title: Assignment on Named Entity Recognition (NER) in Python with Spacy library

```
import spacy
```

```
from spacy import displacy
```

```
NER = spacy.load("en_core_web_sm")
```

```
#raw_text="The Indian Space Research Organisation or is the national space agency of India, headquartered in Bengaluru. It operates under Department of Space which is directly overseen by the Prime Minister of India while Chairman of ISRO acts as executive of DOS as well."
```

```
raw_text="Bitcoin and all major top cryptocurrencies were trading in red at 3:45 pm on Saturday, June 12. In line with its recent trends, overall global crypto market was down by over 15 per cent on the weekend, showed CoinSwitch Kuber data. World number one cryptocurrency Bitcoin was down by 6% and was trading at Rs 27,28,815 after hitting day's high of Rs 29,00,208. "
```

```
text1= NER(raw_text)
```

```
for word in text1.ents:
```

```
    print(word.text,word.label_)
```

```
spacy.explain("ORG")
```

```
spacy.explain("GPE")
```

```
#displacy.render(text1,style="ent",jupyter=True)
```

```
'''
```

output:

3:45 pm TIME

Saturday, June 12 DATE

over 15 per cent MONEY

the weekend DATE

CoinSwitch Kuber PRODUCT

one CARDINAL

Bitcoin PERSON

6% PERCENT

day DATE

Rs 29,00,208 ORG

```
'''
```

```
##Assignment No.03##
```

```
#Name: Radhika Santosh Darode
```

```
##Roll No: 16
```

```
#Batch: B1
```

```
#Title: Assignment on Named Entity Recognition (NER) in Python with Spacy library
```

```
import nltk
```

```
from nltk.tokenize import word_tokenize
```

```
nltk.download('punkt')
```

```
nltk.download('maxent_ne_chunker')
```

```
nltk.download('words')
```

```
text = "Sundar Pichai, the CEO of Google Inc. is walking in the streets of California."
```

```
#text="Google's CEO sundar pichai is from india"
```

```
tokenized_word = word_tokenize(text)
```

```
nltk.download('averaged_perceptron_tagger')
```

```
nltk.download('universal_tagset')
```

```
tags = nltk.pos_tag(tokenized_word, tagset='universal')
```

```
entities = nltk.chunk.ne_chunk(tags, binary=False)
```

```
print(entities)
```

```
'''
```

```
(S
```

```
  Sundar/NOUN
```

```
  Pichai/NOUN
```

```
  ,/.
```

```
  the/DET
```

```
  CEO/NOUN
```

```
  of/ADP
```

```
  Google/NOUN
```

```
  Inc./NOUN
```

```
  is/VERB
```

```
  walking/VERB
```

```
  in/ADP
```

```
  the/DET
```

```
  streets/NOUN
```

of/ADP

(GPE California/NOUN)

./.)

'''