

```
In [1]: import spacy
nlp = spacy.load("en_core_web_sm")
```

```
In [2]: nlp.pipe_names
```

```
Out[2]: ['tok2vec', 'tagger', 'parser', 'attribute_ruler', 'lemmatizer', 'ner']
```

```
In [6]: # suppose we find to entity in this doc

doc = nlp ("Tesla Inc is going to acquire Twitter.inc for $45 billion")

# ent.text will give us the labels
for ent in doc.ents:
    print(ent.text, "|", ent.label_, "|", spacy.explain(ent.label_))
```

Tesla Inc | ORG | Companies, agencies, institutions, etc.  
 Twitter.inc | ORG | Companies, agencies, institutions, etc.  
 \$45 billion | MONEY | Monetary values, including unit

```
In [7]: # we will use displacy library to visualize render in a better way
```

```
from spacy import displacy
displacy.render(doc, style="ent")
```

Tesla Inc **ORG** is going to acquire Twitter.inc **ORG** for \$45 billion **MONEY**

## List down all the entities which NER supports

```
In [9]: nlp.pipe_labels["ner"]
```

```
Out[9]: ['CARDINAL',
        'DATE',
        'EVENT',
        'FAC',
        'GPE',
        'LANGUAGE',
        'LAW',
        'LOC',
        'MONEY',
        'NORP',
        'ORDINAL',
        'ORG',
        'PERCENT',
        'PERSON',
        'PRODUCT',
        'QUANTITY',
        'TIME',
        'WORK_OF_ART']
```

```
In [10]: doc = nlp("Michael Bloomberg founded Bloomberg in 1982")
for ent in doc.ents:
    print (ent.text, "|", ent.label_, "|", spacy.explain(ent.label_))
```

Michael Bloomberg | PERSON | People, including fictional  
 Bloomberg | PERSON | People, including fictional  
 1982 | DATE | Absolute or relative dates or periods

Since it is pre trained model, it works but it didn't a perfect as we can see from above example, another Bloomberg is a company

We can use hugging face for some of the library in spacy, if we want to manually add some model in pre trained model present in spacy- we can use Span and it is a class in Spacy

```
In [11]: # when you look at single token
```

```
doc[0]
```

```
Out[11]: Michael
```

```
In [12]: type(doc[0])
```

```
Out[12]: spacy.tokens.token.Token
```

```
In [13]: doc[2:5]
```

```
Out[13]: founded Bloomberg in
```

```
In [14]: type(doc[2:5])
```

```
Out[14]: spacy.tokens.span.Span
```

It works as span - with a normal python type code

## Setting custom entities

```
In [32]: from spacy.tokens import span
s1 = Span(doc, 3,4, label = "org")
doc.set_ents([s1], default = "unmodified")
```

```
In [33]: for ent in doc.ents:
    print (ent.text, "|", ent.label_)
```

```
Michael Bloomberg | PERSON
Bloomberg | org
1982 | DATE
```