

NER (Named Entity Recognition) Tags across Packages

BERT, spaCy, Stanford CoreNLP, Stanza

Contents

BERT	1
spaCy.....	1
Stanford CoreNLP	2
Stanza.....	2

BERT

The model currently available in the NLP Suite for NER with BERT only processes the English language.

geo for geographical entity

org for organization entity

per for person entity

gpe for geopolitical entity

tim for time indicator entity

art for artifact entity

eve for event entity

nat for natural phenomenon entity

O is assigned if a word doesn't belong to any entity.

spaCy

spaCy includes a list of 18 entity types:

Entity Type	Type Description
CARDINAL	Numerals that do not fall under another type
DATE	Absolute or relative dates or periods
EVENT	Named hurricanes, battles, wars, sports events, etc.
FAC	Buildings, airports, highways, bridges, etc.
GPE	Countries, cities, states
LANGUAGE	Any named language
LAW	Named documents made into law
LOC	Non-GPE locations, mountain ranges, bodies of water
MONEY	Monetary values, including unit
NORP	Nationalities or religious or political groups
ORDINAL	“first”, “second”, etc.
ORG	Companies, agencies, institutions etc.
PERCENT	Percentage, including “%”
PERSON	People, including fictional

Entity Type	Type Description
PRODUCT	Objects, vehicles, foods, etc. (not services)
QUANTITY	Measurements, as of weight or distance
TIME	Times smaller than a day
WORK_OF_ART	Titles of books, songs, etc.

Stanford CoreNLP

PERSON
 ORGANIZATION
 MISC
 MONEY
 NUMBER
 ORDINAL
 PERCENT
 DATE
 TIME
 DURATION
 SETE
 MAIL
 URL
 CITY
 STATE_OR_PROVINCE
 COUNTRY
 LOCATION
 NATIONALITY
 RELIGION
 TITLE
 IDEOLOGY
 CRIMINAL_CHARGE
 CAUSE_OF_DEATH

Stanza

```

"fr": [
  "LOC",
  "MISC",
  "ORG",
  "PER"
],
"en": [
  "CARDINAL",
  "DATE",
  "EVENT",

```

```
"FAC",
"GPE",
"LANGUAGE",
"LAW",
"LOC",
"MONEY",
"NORP",
"ORDINAL",
"ORG",
"PERCENT",
"PERSON",
"PRODUCT",
"QUANTITY",
"TIME",
"WORK_OF_ART"
],
"zh-hans": [
  "CARDINAL",
  "DATE",
  "EVENT",
  "FAC",
  "GPE",
  "LANGUAGE",
  "LAW",
  "LOC",
  "MONEY",
  "NORP",
  "ORDINAL",
  "ORG",
  "PERCENT",
  "PERSON",
  "PRODUCT",
  "QUANTITY",
  "TIME",
  "WORK_OF_ART"
],
"ru": [
  "LOC",
  "MISC",
  "ORG",
  "PER"
],
"uk": [
  "LOC",
```

```
    "MISC",
    "ORG",
    "PERS"
  ],
  "ar": [
    "LOC",
    "MISC",
    "ORG",
    "PER"
  ],
  "hu": [
    "LOC",
    "MISC",
    "ORG",
    "PER"
  ],
  "af": [
    "LOC",
    "MISC",
    "ORG",
    "PERS"
  ],
  "bg": [
    "EVT",
    "LOC",
    "ORG",
    "PER",
    "PRO"
  ],
  "fi": [
    "DATE",
    "EVENT",
    "LOC",
    "ORG",
    "PER",
    "PRO"
  ],
  "my": [
    "LOC",
    "NE",
    "NUM",
    "ORG",
    "PNAME",
    "RACE",
```

```
    "TIME"
  ],
  "it": [
    "LOC",
    "ORG",
    "PER"
  ],
  "de": [
    "LOC",
    "MISC",
    "ORG",
    "PER"
  ],
  "nl": [
    "LOC",
    "MISC",
    "ORG",
    "PER"
  ],
  "vi": [
    "LOCATION",
    "MISCELLANEOUS",
    "ORGANIZATION",
    "PERSON"
  ],
  "es": [
    "LOC",
    "MISC",
    "ORG",
    "PER"
  ]
]
```