

```
from google.colab import drive
drive.mount("gdrive")
```

Mounted at gdrive

pwd

'/content'

```
cd drive/MyDrive/coven_nlp_sept
```

/content/drive/MyDrive/coven_nlp_sept

```
!python -m spacy download en_core_web_sm
```

```
➡ Collecting en_core_web_sm==2.2.5
```

Downloading https://github.com/explosion/spacy-models/releases/download/en_core_web_sm-2.2.5/en_core_web_sm-2.2.5.tar.gz (12

12.0 MB 24.7 MB/s

```
Requirement already satisfied: spacy>=2.2.2 in /usr/local/lib/python3.7/dist-packages (from en_core_web_sm==2.2.5) (2.2.4)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in /usr/local/lib/python3.7/dist-packages (from spacy>=2.2.2->en_core_web
Requirement already satisfied: thinc==7.4.0 in /usr/local/lib/python3.7/dist-packages (from spacy>=2.2.2->en_core_web_sm==2.2.
Requirement already satisfied: blis<0.5.0,>=0.4.0 in /usr/local/lib/python3.7/dist-packages (from spacy>=2.2.2->en_core_web_sm
Requirement already satisfied: srsly<1.1.0,>=1.0.2 in /usr/local/lib/python3.7/dist-packages (from spacy>=2.2.2->en_core_web_s
Requirement already satisfied: plac<1.2.0,>=0.9.6 in /usr/local/lib/python3.7/dist-packages (from spacy>=2.2.2->en_core_web_sm
Requirement already satisfied: requests<3.0.0,>=2.13.0 in /usr/local/lib/python3.7/dist-packages (from spacy>=2.2.2->en_core_w
Requirement already satisfied: setuptools in /usr/local/lib/python3.7/dist-packages (from spacy>=2.2.2->en_core_web_sm==2.2.5)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in /usr/local/lib/python3.7/dist-packages (from spacy>=2.2.2->en_core
Requirement already satisfied: wasabi<1.1.0,>=0.4.0 in /usr/local/lib/python3.7/dist-packages (from spacy>=2.2.2->en_core_web_
Requirement already satisfied: numpy>=1.15.0 in /usr/local/lib/python3.7/dist-packages (from spacy>=2.2.2->en_core_web_sm==2.2
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in /usr/local/lib/python3.7/dist-packages (from spacy>=2.2.2->en_core_web_s
Requirement already satisfied: catalogue<1.1.0,>=0.0.7 in /usr/local/lib/python3.7/dist-packages (from spacy>=2.2.2->en_core_w
Requirement already satisfied: tqdm<5.0.0,>=4.38.0 in /usr/local/lib/python3.7/dist-packages (from spacy>=2.2.2->en_core_web_s
Requirement already satisfied: importlib-metadata>=0.20 in /usr/local/lib/python3.7/dist-packages (from catalogue<1.1.0,>=0.0.
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-packages (from importlib-metadata>=0.20->catalogue<1
Requirement already satisfied: typing-extensions>=3.6.4 in /usr/local/lib/python3.7/dist-packages (from importlib-metadata>=0.
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-packages (from requests<3.0.0,>=2.13.0->spacy>=2.
```

```
Requirement already satisfied: chardet<4,>=3.0.2 in /usr/local/lib/python3.7/dist-packages (from requests<3.0.0,>=2.13.0->spac
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.7/dist-packages (from requests<3.0.0,>=2.13.0->spa
Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in /usr/local/lib/python3.7/dist-packages (from request
✓ Download and installation successful
You can now load the model via spacy.load('en_core_web_sm')
```

```
import spacy
```

```
# pip install docx2txt
```

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

```
import docx2txt
```

```
project2 = docx2txt.process("projecttext.docx")
```

```
print(project2)
```

Why President Obama didn't release footage of the raid that killed Osama bin Laden

As he announced the death of Islamic State leader Abu Bakr al-Baghdadi on Sunday, President Trump described the event in detai

Now, it turns out the public may get to watch part of this movie themselves. President Trump said Monday that the White House

At a news conference two days later, Gen. Frank McKenzie, the head of the United States Central Command, displayed aerial foot

It is not yet clear if more footage could ultimately be released. But the move marks another key difference in Trump's handling

In that case, Barrack Obama refused to release images of Bin Laden's body, citing national security concerns. No video footage

What video footage might exist of bin Laden's death, if any, is unclear. CBS News reported in 2011 that 25 helmet cams had recorded

More unambiguous, though no less contentious, were the photos of bin Laden's body. On May 4, two days after the raid that killed

```
doc = nlp(project)
```

```
for ent in doc.ents:
```

```
    print(ent.text, ent.label_, ent.start_char, ent.end_char)
```

```
Obama PERSON 14 19
```

```
Osama bin Laden PERSON 67 82
```

```
Islamic NORP 114 121
```

```
Abu Bakr al-Baghdadi PERSON 135 155
```

```
Sunday DATE 159 165
```

```
Trump PERSON 177 182
```

```
U.S. GPE 366 370
```

```
Trump PERSON 477 482
```

```
Monday DATE 488 494
```

```
the White House ORG 500 515
```

```
Trump PERSON 646 651
```

```
Washington GPE 687 697
```

```
Chicago GPE 701 708
```

```
two days later DATE 732 746
```

```
Frank McKenzie PERSON 753 767
```

```
the United States Central Command ORG 781 814
```

```
Abu Bakr al-Baghdadi PERSON 900 920
```

Baghdadi ORG 962 970
Trump ORG 1151 1156
Abu Bakr al-Baghdadi PERSON 1171 1191
Barack Obama PERSON 1215 1227
2011 DATE 1230 1234
Osama bin Laden PERSON 1270 1285
al-Qaeda ORG 1318 1326
Barrack Obama PERSON 1343 1356
Bin Laden PERSON 1386 1395
U.S. GPE 1498 1502
bin Laden PERSON 1550 1559
CBS News ORG 1589 1597
2011 DATE 1610 1614
25 CARDINAL 1620 1622
bin Laden PERSON 1659 1668
al-Qaeda ORG 1702 1710
New Yorker ORG 1747 1757
bin Laden PERSON 1935 1944
May 4 DATE 1956 1961
two days DATE 1963 1971
Obama PERSON 2021 2026
the United States GPE 2051 2068
bin Laden PERSON 2100 2109

REDACT ENTITY PROJECT

```
import spacy
nlp = spacy.load("en_core_web_md")
project = open("project.txt").read()

def redact_entity(rawtext):
    untouched = rawtext
    doc = nlp(rawtext)
    for ent in doc.ents:
        untouched = untouched[:ent.start_char] + " [ " + ent.label_ + " ] " + untouched[ent.end_char:]
    return untouched
```

```
redacted_text = redact_entity(project)
```

```
print(redacted_text)
```

Why President [PERSON] didn't release footage of the raid that [PERSON] n Laden

As he announced the death [NORP] amic State [PERSON] aghdadi on S [DATE] Preside [PERSON] ump described the event in

Now, it turns out the public may get to watch part of this movie themsel [PERSON] [DATE] de [ORG] onday that the White H

It is not yet clear if more footage could ultimately be released. But the move marks another key difference in Trump's h [ORG]
[ORG] case, Barrack O [PERSON] to release images of Bin Laden' [PERSON] iting national security concerns. No video foo

What video footage might exist of bi [PERSON] death, if any, is unclear. [ORG] s reported i [DATE] [CARDINAL] that

More unambiguous, though no less contentio [PERSON] the photo [DATE [DATE] aden's body. On May 4, two days after the raid

