

```
In [7]: import pandas as pd
from google.cloud import bigquery

import warnings
warnings.filterwarnings("ignore")
```

```
In [8]: # Mount google drive
from google.colab import drive
drive.mount('/content/drive')
```

Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).

Fetching and Filtering Data of noteevents

```
In [52]: from google.colab import auth
auth.authenticate_user()
print('Authenticated')

!gcloud projects list

from google.cloud import bigquery

# Construct a BigQuery client object.
client = bigquery.Client(project='clinical-entity-extraction')

"""
ICD codes related to Hypertension:
4010 - Malignant essential hypertension
4011 - Benign essential hypertension
4019 - Unspecified essential hypertension
"""

# Fetch notes only for ICD-9 code 4010(Malignant essential hypertension)
query = """
SELECT SUBJECT_ID, TEXT, CATEGORY
FROM `physionet-data.mimiciii_notes.noteevents`
WHERE SUBJECT_ID IN (
    SELECT d.SUBJECT_ID
    FROM `physionet-data.mimiciii_clinical.diagnoses_icd` d
    WHERE d.ICD9_CODE = '4010' -- Hypertension code
    AND d.SEQ_NUM = 1 -- Assuming 1 indicates primary diagnosis
)
AND CATEGORY LIKE 'Discharge summary';
"""

# Run the query
query_job = client.query(query)

# Print the results
noteevents_df = query_job.to_dataframe()

len(noteevents_df)
```

Authenticated

PROJECT_ID	NAME	PROJECT_NUMBER
clinical-entity-extraction	clinical-entity-extraction	321960627270

Out[52]: 162

```
In [53]: patients_dict = {"SUBJECT_ID":[], "CATEGORY":[], "TEXT":[]};
for i in range(0, len(noteevents_df)):
    patients_dict["SUBJECT_ID"].append(noteevents_df.loc[i, 'SUBJECT_ID'])
    patients_dict["CATEGORY"].append(noteevents_df.loc[i, 'CATEGORY'])
    patients_dict["TEXT"].append(noteevents_df.loc[i, 'TEXT'])
```

```
In [54]: patients_df = pd.DataFrame(patients_dict)
len(patients_df)
patients_df = patients_df.iloc[2:] # Cleanup
#patients_df = patients_df.iloc[1:] #Cleanup
```

```
In [162... #print first few records
patients_df.head(1)
```

```
Out[162... SUBJECT_ID      CATEGORY      TEXT
2          109  Discharge summary  Admission Date: [**2140-1-19**] ...
```

```
In [56]: # Download the patients_df dataframe in .csv and excel format
patients_df.to_csv('Patient_Summary.csv', index = False)
patients_df.to_excel("Patient_Summary.xlsx")
```

```
In [57]: # Copy Processed data to google drive
!cp 'Patient_Summary.csv' '/content/drive/MyDrive/Colab Notebooks/AIH/Patient_Summa
!cp 'Patient_Summary.xlsx' '/content/drive/MyDrive/Colab Notebooks/AIH/Patient_Summa
```

Functions to clean text and extract tokens and entities.

```
In [ ]: ! pip install -U spacy
```

Requirement already satisfied: spacy in /usr/local/lib/python3.11/dist-packages (3.7.5)
Collecting spacy
 Downloading spacy-3.8.4-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (27 kB)
Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.11 in /usr/local/lib/python3.11/dist-packages (from spacy) (3.0.12)
Requirement already satisfied: spacy-loggers<2.0.0,>=1.0.0 in /usr/local/lib/python3.11/dist-packages (from spacy) (1.0.5)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in /usr/local/lib/python3.11/dist-packages (from spacy) (1.0.12)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in /usr/local/lib/python3.11/dist-packages (from spacy) (2.0.11)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in /usr/local/lib/python3.11/dist-packages (from spacy) (3.0.9)
Collecting thinc<8.4.0,>=8.3.4 (from spacy)
 Downloading thinc-8.3.4-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (15 kB)
Requirement already satisfied: wasabi<1.2.0,>=0.9.1 in /usr/local/lib/python3.11/dist-packages (from spacy) (1.1.3)
Requirement already satisfied: srsly<3.0.0,>=2.4.3 in /usr/local/lib/python3.11/dist-packages (from spacy) (2.5.1)
Requirement already satisfied: catalogue<2.1.0,>=2.0.6 in /usr/local/lib/python3.11/dist-packages (from spacy) (2.0.10)
Requirement already satisfied: weasel<0.5.0,>=0.1.0 in /usr/local/lib/python3.11/dist-packages (from spacy) (0.4.1)
Requirement already satisfied: typer<1.0.0,>=0.3.0 in /usr/local/lib/python3.11/dist-packages (from spacy) (0.15.1)
Requirement already satisfied: tqdm<5.0.0,>=4.38.0 in /usr/local/lib/python3.11/dist-packages (from spacy) (4.67.1)
Requirement already satisfied: numpy>=1.19.0 in /usr/local/lib/python3.11/dist-packages (from spacy) (1.26.4)
Requirement already satisfied: requests<3.0.0,>=2.13.0 in /usr/local/lib/python3.11/dist-packages (from spacy) (2.32.3)
Requirement already satisfied: pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4 in /usr/local/lib/python3.11/dist-packages (from spacy) (2.10.6)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.11/dist-packages (from spacy) (3.1.5)
Requirement already satisfied: setuptools in /usr/local/lib/python3.11/dist-packages (from spacy) (75.1.0)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.11/dist-packages (from spacy) (24.2)
Requirement already satisfied: langcodes<4.0.0,>=3.2.0 in /usr/local/lib/python3.11/dist-packages (from spacy) (3.5.0)
Requirement already satisfied: language-data>=1.2 in /usr/local/lib/python3.11/dist-packages (from langcodes<4.0.0,>=3.2.0->spacy) (1.3.0)
Requirement already satisfied: annotated-types>=0.6.0 in /usr/local/lib/python3.11/dist-packages (from pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4->spacy) (0.7.0)
Requirement already satisfied: pydantic-core==2.27.2 in /usr/local/lib/python3.11/dist-packages (from pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4->spacy) (2.27.2)
Requirement already satisfied: typing-extensions>=4.12.2 in /usr/local/lib/python3.11/dist-packages (from pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4->spacy) (4.12.2)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests<3.0.0,>=2.13.0->spacy) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-packages (from requests<3.0.0,>=2.13.0->spacy) (3.10)

Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-packages (from requests<3.0.0,>=2.13.0->spacy) (2.3.0)

Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-packages (from requests<3.0.0,>=2.13.0->spacy) (2025.1.31)

Collecting blis<1.3.0,>=1.2.0 (from thinc<8.4.0,>=8.3.4->spacy)

Downloading blis-1.2.0-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (7.7 kB)

Requirement already satisfied: confection<1.0.0,>=0.0.1 in /usr/local/lib/python3.11/dist-packages (from thinc<8.4.0,>=8.3.4->spacy) (0.1.5)

Requirement already satisfied: click>=8.0.0 in /usr/local/lib/python3.11/dist-packages (from typer<1.0.0,>=0.3.0->spacy) (8.1.8)

Requirement already satisfied: shellingham>=1.3.0 in /usr/local/lib/python3.11/dist-packages (from typer<1.0.0,>=0.3.0->spacy) (1.5.4)

Requirement already satisfied: rich>=10.11.0 in /usr/local/lib/python3.11/dist-packages (from typer<1.0.0,>=0.3.0->spacy) (13.9.4)

Requirement already satisfied: cloudpathlib<1.0.0,>=0.7.0 in /usr/local/lib/python3.11/dist-packages (from weasel<0.5.0,>=0.1.0->spacy) (0.20.0)

Requirement already satisfied: smart-open<8.0.0,>=5.2.1 in /usr/local/lib/python3.11/dist-packages (from weasel<0.5.0,>=0.1.0->spacy) (7.1.0)

Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.11/dist-packages (from jinja2->spacy) (3.0.2)

Requirement already satisfied: marisa-trie>=1.1.0 in /usr/local/lib/python3.11/dist-packages (from language-data>=1.2->langcodes<4.0.0,>=3.2.0->spacy) (1.2.1)

Requirement already satisfied: markdown-it-py>=2.2.0 in /usr/local/lib/python3.11/dist-packages (from rich>=10.11.0->typer<1.0.0,>=0.3.0->spacy) (3.0.0)

Requirement already satisfied: pygments<3.0.0,>=2.13.0 in /usr/local/lib/python3.11/dist-packages (from rich>=10.11.0->typer<1.0.0,>=0.3.0->spacy) (2.18.0)

Requirement already satisfied: wrapt in /usr/local/lib/python3.11/dist-packages (from smart-open<8.0.0,>=5.2.1->weasel<0.5.0,>=0.1.0->spacy) (1.17.2)

Requirement already satisfied: mdurl~=0.1 in /usr/local/lib/python3.11/dist-packages (from markdown-it-py>=2.2.0->rich>=10.11.0->typer<1.0.0,>=0.3.0->spacy) (0.1.2)

Downloading spacy-3.8.4-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (30.6 MB)

30.6/30.6 MB 56.3 MB/s eta 0:00:00

Downloading thinc-8.3.4-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (3.9 MB)

3.9/3.9 MB 99.9 MB/s eta 0:00:00

Downloading blis-1.2.0-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (1.7 MB)

11.7/11.7 MB 90.2 MB/s eta 0:00:00

Installing collected packages: blis, thinc, spacy

Attempting uninstall: blis

Found existing installation: blis 0.7.11

Uninstalling blis-0.7.11:

Successfully uninstalled blis-0.7.11

Attempting uninstall: thinc

Found existing installation: thinc 8.2.5

Uninstalling thinc-8.2.5:

Successfully uninstalled thinc-8.2.5

Attempting uninstall: spacy

Found existing installation: spacy 3.7.5

Uninstalling spacy-3.7.5:

Successfully uninstalled spacy-3.7.5

ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behaviour is the source of the following dependency conflicts.

en-core-web-sm 3.7.1 requires spacy<3.8.0,>=3.7.2, but you have spacy 3.8.4 which is incompatible.

Successfully installed blis-1.2.0 spacy-3.8.4 thinc-8.3.4

In []: ! python -m spacy download en_core_web_sm

```

2025-02-13 21:02:59.719284: E external/local_xla/xla/stream_executor/cuda/cuda_fft.c
c:477] Unable to register cuFFT factory: Attempting to register factory for plugin c
uFFT when one has already been registered
WARNING: All log messages before absl::InitializeLog() is called are written to STDERR
E0000 00:00:1739480579.740664    20463 cuda_dnn.cc:8310] Unable to register cuDNN fac
tory: Attempting to register factory for plugin cuDNN when one has already been regi
stered
E0000 00:00:1739480579.747289    20463 cuda_blas.cc:1418] Unable to register cuBLAS f
actory: Attempting to register factory for plugin cuBLAS when one has already been r
egistered
Collecting en-core-web-sm==3.6.0
  Downloading https://github.com/explosion/spacy-models/releases/download/en_core_we
b_sm-3.6.0/en_core_web_sm-3.6.0-py3-none-any.whl (12.8 MB)
      12.8/12.8 MB 58.3 MB/s eta 0:00:00
Requirement already satisfied: spacy<3.7.0,>=3.6.0 in /usr/local/lib/python3.11/dist
-packages (from en-core-web-sm==3.6.0) (3.6.1)
Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.11 in /usr/local/lib/python
3.11/dist-packages (from spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (3.0.12)
Requirement already satisfied: spacy-loggers<2.0.0,>=1.0.0 in /usr/local/lib/python
3.11/dist-packages (from spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (1.0.5)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in /usr/local/lib/python3.1
1/dist-packages (from spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (1.0.12)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in /usr/local/lib/python3.11/dist
-packages (from spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (2.0.11)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in /usr/local/lib/python3.11/di
st-packages (from spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (3.0.9)
Requirement already satisfied: thinc<8.2.0,>=8.1.8 in /usr/local/lib/python3.11/dist
-packages (from spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (8.1.12)
Requirement already satisfied: wasabi<1.2.0,>=0.9.1 in /usr/local/lib/python3.11/dis
t-packages (from spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (1.1.3)
Requirement already satisfied: srsly<3.0.0,>=2.4.3 in /usr/local/lib/python3.11/dist
-packages (from spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (2.5.1)
Requirement already satisfied: catalogue<2.1.0,>=2.0.6 in /usr/local/lib/python3.11/
dist-packages (from spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (2.0.10)
Requirement already satisfied: typer<0.10.0,>=0.3.0 in /usr/local/lib/python3.11/dis
t-packages (from spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (0.9.4)
Requirement already satisfied: pathy>=0.10.0 in /usr/local/lib/python3.11/dist-packa
ges (from spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (0.11.0)
Requirement already satisfied: smart-open<7.0.0,>=5.2.1 in /usr/local/lib/python3.1
1/dist-packages (from spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (6.4.0)
Requirement already satisfied: tqdm<5.0.0,>=4.38.0 in /usr/local/lib/python3.11/dist
-packages (from spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (4.67.1)
Requirement already satisfied: numpy>=1.15.0 in /usr/local/lib/python3.11/dist-packa
ges (from spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (1.26.4)
Requirement already satisfied: requests<3.0.0,>=2.13.0 in /usr/local/lib/python3.11/
dist-packages (from spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (2.32.3)
Requirement already satisfied: pydantic!=1.8,!<1.8.1,<3.0.0,>=1.7.4 in /usr/local/li
b/python3.11/dist-packages (from spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (2.10.
6)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.11/dist-packages (fr
om spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (3.1.5)
Requirement already satisfied: setuptools in /usr/local/lib/python3.11/dist-packages
 (from spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (75.1.0)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.11/dist-pac
kages (from spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (24.2)

```

Requirement already satisfied: langcodes<4.0.0,>=3.2.0 in /usr/local/lib/python3.11/dist-packages (from spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (3.5.0)

Requirement already satisfied: language-data>=1.2 in /usr/local/lib/python3.11/dist-packages (from langcodes<4.0.0,>=3.2.0->spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (1.3.0)

Requirement already satisfied: pathlib-abc==0.1.1 in /usr/local/lib/python3.11/dist-packages (from pathy>=0.10.0->spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (0.1.1)

Requirement already satisfied: annotated-types>=0.6.0 in /usr/local/lib/python3.11/dist-packages (from pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4->spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (0.7.0)

Requirement already satisfied: pydantic-core==2.27.2 in /usr/local/lib/python3.11/dist-packages (from pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4->spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (2.27.2)

Requirement already satisfied: typing-extensions>=4.12.2 in /usr/local/lib/python3.11/dist-packages (from pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4->spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (4.12.2)

Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests<3.0.0,>=2.13.0->spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (3.4.1)

Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-packages (from requests<3.0.0,>=2.13.0->spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (3.10)

Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-packages (from requests<3.0.0,>=2.13.0->spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (2.3.0)

Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-packages (from requests<3.0.0,>=2.13.0->spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (2025.1.31)

Requirement already satisfied: blis<0.8.0,>=0.7.8 in /usr/local/lib/python3.11/dist-packages (from thinc<8.2.0,>=8.1.8->spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (0.7.11)

Requirement already satisfied: confection<1.0.0,>=0.0.1 in /usr/local/lib/python3.11/dist-packages (from thinc<8.2.0,>=8.1.8->spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (0.1.5)

Requirement already satisfied: click<9.0.0,>=7.1.1 in /usr/local/lib/python3.11/dist-packages (from typer<0.10.0,>=0.3.0->spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (8.1.8)

Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.11/dist-packages (from jinja2->spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (3.0.2)

Requirement already satisfied: marisa-trie>=1.1.0 in /usr/local/lib/python3.11/dist-packages (from language-data>=1.2->langcodes<4.0.0,>=3.2.0->spacy<3.7.0,>=3.6.0->en-core-web-sm==3.6.0) (1.2.1)

Installing collected packages: en-core-web-sm

Attempting uninstall: en-core-web-sm

Found existing installation: en-core-web-sm 3.7.1

Uninstalling en-core-web-sm-3.7.1:

Successfully uninstalled en-core-web-sm-3.7.1

Successfully installed en-core-web-sm-3.6.0

✓ Download and installation successful

You can now load the package via `spacy.load('en_core_web_sm')`

```
In [41]: import spacy

# Function to clean and extract tokens (removing punctuation and spaces)
def extract_cleaned_text(text, nlp_model):
    doc = nlp_model(str(text)) # Convert to string and process text
```

```
tokens = [token.text for token in doc if not token.is_punct and not token.is_sp
return " ".join(tokens) # Return cleaned text as a string
```

```
In [96]: import numpy as np
from sklearn.manifold import TSNE
import matplotlib.pyplot as plt
%matplotlib inline

def tsne_plot(model, words, preTrained=False):
    """
    Creates and displays two t-SNE plots:
    1. Simple scatter plot with labels.
    2. Scatter plot with distance-based coloring.

    Parameters:
    - model: The Word2Vec model or pre-trained model.
    - words: List of words to visualize.
    - preTrained: Boolean flag to choose between Word2Vec or pre-trained model.
    """
    labels = []
    tokens = []

    # Apply t-SNE for dimensionality reduction
    tsne_model = TSNE(perplexity=30, early_exaggeration=12, n_components=2, init='p

    # Prepare tokens and labels
    for word in words:
        if preTrained:
            tokens.append(model[word]) # Pre-trained word vectors
        else:
            tokens.append(model.wv[word]) # Word2Vec model vectors
            labels.append(word)

    tokens = np.array(tokens)
    new_values = tsne_model.fit_transform(tokens)

    x = new_values[:, 0]
    y = new_values[:, 1]

    # First plot: Scatter plot with annotations
    plt.figure(figsize=(10, 8))
    for i in range(len(x)):
        plt.scatter(x[i], y[i])
        plt.annotate(labels[i],
                     xy=(x[i], y[i]),
                     xytext=(5, 2),
                     textcoords='offset points',
                     ha='right',
                     va='bottom')
    plt.title("t-SNE Visualization (Labels)")
    plt.show()

    # Second plot: Scatter plot with distance-based coloring
    plt.figure(figsize=(10, 8))
    distances = np.sqrt(new_values[:, 0]**2 + new_values[:, 1]**2)
    plt.scatter(new_values[:, 0], new_values[:, 1], c=distances, cmap='plasma')
```



```
plt.colorbar(label="Distance from Origin")
plt.title("t-SNE Visualization with Distance-Based Coloring")
plt.show()
```

```
In [43]: import spacy

def build_corpus(df, model="en_core_web_sm"):
    """
    Extracts named entities from the specified text column in a DataFrame using a spaCy model and builds a corpus.

    Parameters:
    - df (pd.DataFrame): DataFrame containing text data.
    - text_column (str): Column name containing processed text.
    - model (str): spaCy model to use (default: "en_core_web_sm").

    Returns:
    - corpus (list of lists): Extracted entities per document.
    """
    nlp = model
    corpus = []

    for _, row in df.iterrows():
        tokens = [ent.text for ent in nlp(row["Processed_Text"]).ents]
        corpus.append(tokens)

    # Calculate word counts
    word_counts = [len(doc) for doc in corpus]

    print(f"Minimum word count: {min(word_counts)}, Maximum word count: {max(word_counts)}")

    return corpus
```

Using Spacy

```
In [58]: import pandas as pd

# Load Patient summary
patients_df_scapy = pd.read_csv("/content/drive/MyDrive/Colab Notebooks/AIH/Patient_Summary.csv")

# Load the spacy model
nlp_spacy = spacy.load('en_core_web_sm')

# Apply token extraction
patients_df_scapy["Processed_Text"] = patients_df_scapy["TEXT"].apply(lambda text: nlp_spacy(text).ents)

# Copy Processed data to google drive
patients_df_scapy.to_csv(r'Patient_Summary_Spacy.csv', index = False)
!cp 'Patient_Summary_Spacy.csv' '/content/drive/MyDrive/Colab Notebooks/AIH/Patient_Summary_Spacy.csv'
```

```
In [59]: import pandas as pd

# Load the processed patient data with spacy
patients_df_scapy = pd.read_csv("/content/drive/MyDrive/Colab Notebooks/AIH/Patient_Summary_Spacy.csv")
```

```
nlp_spacy = spacy.load('en_core_web_sm') # Load the specified NLP model
```

```
In [60]: doc = nlp_spacy( patients_df_scapy['Processed_Text'][0])  
for ent in doc.ents:  
    print(ent.text, ent.start_char, ent.end_char, ent.label_)
```

2297 146 150 DATE
Known 257 262 PERSON
22 year old 277 288 DATE
ESRD 321 325 ORG
TTP 350 353 ORG
HOcm 358 362 ORG
HA 381 383 ORG
Awoke 409 414 ORG
this a.m. 415 424 TIME
8/10 430 434 CARDINAL
HA 454 456 ORG
Monday 521 527 DATE
SOB + Diarrhea 650 664 ORG
1 day 667 672 DATE
217/140 691 698 CARDINAL
254/152 715 722 CARDINAL
IV 30 744 749 DATE
1 755 756 CARDINAL
4 766 767 CARDINAL
208 801 804 CARDINAL
50 mg x 843 850 QUANTITY
1 851 852 CARDINAL
193/134 904 911 CARDINAL
CXR 1006 1009 ORG
ROS 1023 1026 ORG
past week 1036 1045 DATE
SOB N V + 1066 1075 ORG
BP 1122 1124 GPE
191/126 1129 1136 CARDINAL
2 1408 1409 CARDINAL
ESRD 1410 1414 ORG
SLE 1427 1430 ORG
1 1461 1462 CARDINAL
every 3 months 1468 1482 DATE
2 years 1487 1494 DATE
3 1516 1517 CARDINAL
2137 1534 1538 DATE
3 1593 1594 CARDINAL
2137 1599 1603 DATE
1 1640 1641 CARDINAL
Echo 1744 1748 GPE
2137 1752 1756 DATE
2139 1776 1780 DATE
20 1783 1785 CARDINAL
7 1786 1787 CARDINAL
8 1828 1829 CARDINAL
9 1837 1838 CARDINAL
Staph 1848 1853 PERSON
6 1887 1888 CARDINAL
10 1892 1894 CARDINAL
16 year old 1990 2001 DATE
Graduated Name2 NI School 2010 2035 ORG
Denies 2102 2108 WORK_OF_ART
SLE -Grandfather 2150 2166 ORG
DM 2195 2197 GPE
Physical Exam Vitals 2279 2299 ORG

98.0 2300 2304 CARDINAL
173/51 2305 2311 CARDINAL
86 2312 2314 CARDINAL
100 2318 2321 CARDINAL
RA HEENT 2322 2330 PERSON
w/ EOMI 2381 2388 ORG
LAD 2434 2437 GPE
Cardiac RRR NL 2453 2467 ORG
S1 2468 2470 PRODUCT
w/ Valsalva 2561 2572 PERSON
Lungs 2580 2585 WORK_OF_ART
Abd 2619 2622 PERSON
HSM 2641 2644 ORG
GU 2668 2670 ORG
2 + DP 2688 2694 DATE
Neuro AOx3 2739 2749 PERSON
Results UA 2811 2821 PERSON
100 2830 2833 CARDINAL
CXR 2873 2876 NORP
EKG NSR nml 2901 2912 ORG
LAE LVH 2943 2950 ORG
V5 2988 2990 CARDINAL
2139 3030 3034 DATE
26 3038 3040 CARDINAL
Brief Hospital Course 3076 3097 WORK_OF_ART
22 year old 3115 3126 DATE
ESRD 3159 3163 ORG
Unclear 3230 3237 ORG
Compliant 3300 3309 PERSON
ED 3401 3403 GPE
PO 3463 3465 GPE
ICU 3488 3491 ORG
150s-170s 3527 3536 CARDINAL
40 mg 3625 3630 QUANTITY
40 mg po 3643 3651 QUANTITY
BP 3675 3677 ORG
UA 3727 3729 ORG
CRF CE 's 3767 3776 LOC
1 3779 3780 CARDINAL
Headache 3790 3798 FAC
CT 3814 3816 ORG
Headaches 3840 3849 PERSON
Uveitis Followed 3931 3947 ORG
Optho 3979 3984 PERSON
ESRD 4022 4026 ORG
Patient 4075 4082 PERSON
500 ml 4119 4125 QUANTITY
45 kg 4179 4184 QUANTITY
Began Sevalamer 4197 4212 ORG
TID 4217 4220 ORG
Admission Lisinopril 40 4404 4427 PRODUCT
PO QD 4431 4436 FAC
600 4447 4450 CARDINAL
PO 4451 4453 GPE
320 4468 4471 CARDINAL
PO QD Clonidine 4475 4490 ORG

0.3 4491 4494 CARDINAL
PO QD 4530 4535 FAC
Prednisolone Acetate 1 4557 4579 ORG
Lorazepam 1 4611 4622 LAW
PO Q4 4626 4631 FAC
PRN Discharge 4635 4648 ORG
1 4661 4662 CARDINAL
Three 4691 4696 CARDINAL
3 4697 4698 CARDINAL
3 4713 4714 CARDINAL
0.3 mg/24 4748 4757 QUANTITY
Patch Weekly Sig 4761 4777 ORG
One 4778 4781 CARDINAL
Thursday 3 4821 4831 DATE
One 4853 4856 CARDINAL
2 4885 4886 CARDINAL
One 4927 4930 CARDINAL
4 to 6 hours 4955 4967 TIME
5 4978 4979 CARDINAL
160 4990 4993 CARDINAL
Tablet 4997 5003 ORG
Two 5008 5011 CARDINAL
2 5012 5013 CARDINAL
PO DAILY Daily 5021 5035 ORG
6 5036 5037 CARDINAL
Prednisolone Acetate 1 Drops 5038 5066 ORG
One 5082 5085 CARDINAL
1 5086 5087 CARDINAL
7 Lisinopril 40 5119 5134 TIME
Tablet 5138 5144 GPE
One 5149 5152 CARDINAL
Tablet(s Refills:*2 5186 5205 PERSON
Tablet 5225 5231 PERSON
One 5236 5239 CARDINAL
3 5256 5257 CARDINAL
20 5312 5314 CARDINAL
Tablet 5318 5324 PERSON
Two 5329 5332 CARDINAL
2 5333 5334 CARDINAL
Kit Kit Sig 5374 5385 PERSON
One 5386 5389 CARDINAL
Kit Refills:*0 Discharge 5429 5453 PERSON
daily 5830 5835 DATE
180 5952 5955 CARDINAL
Followup Instructions Please 6067 6095 WORK_OF_ART
First Name8 6180 6191 PERSON
4883 6228 6232 DATE
the Division of Nephrology 6236 6262 ORG
Wednesday 2 3 6266 6279 DATE
9:30 AM 6283 6290 TIME
First Name4 6436 6447 ORG
2423 6484 6488 DATE
Tuesday 1 26 6492 6504 DATE
3:30 PM 6508 6515 TIME
1 250 6542 6547 CARDINAL
2539 6656 6660 DATE

the Division of Hematology 6664 6690 ORG
2 9 6763 6766 DATE
3 p.m. 6770 6776 TIME
the Location un of the Hospital Ward Name 6802 6843 ORG
23 6844 6846 CARDINAL
the Hospital1 18 Hospital Ward Name 6859 6894 FAC
516 6895 6898 CARDINAL
44536 6925 6930 CARDINAL
First Name 8982 6963 6978 DATE
Telephone Fax 1 32192 6982 7003 ORG

```
In [61]: from spacy import displacy  
  
# Visualize named entities using displacy  
displacy.render(nlp_spacy(patients_df_scapy['Processed_Text'][0]), style="ent")
```

Admission Date 2140 1 19 Discharge Date 2140 1 21 Date of Birth 2117 8 7 Sex F Service
 MEDICINE Allergies Penicillins Attending:[**First Name3 LF 2297 DATE Chief Complaint
 headache Major Surgical or Invasive Procedure Hemodialysis History of Present Illness Ms.
 Known PERSON lastname is a 22 year old DATE female with SLE lupus nephritis ESRD
 ORG on HD malignant HTN h o TTP ORG and HOCM ORG who presents with HA
 ORG and hypertensive urgency Awoke ORG this a.m. TIME with 8/10 CARDINAL
 left sided frontal HA ORG was n't sure if it was d t flare of uveitis that had started on
 Monday DATE or d t HTN Decided to skip HD and come to ED for evaluation No vision
 changes numbness weakness change in gait chest pain SOB + Diarrhea ORG x 1 day
 DATE In ED patient was 217/140 CARDINAL but elevated to 254/152 CARDINAL >
 received labetalol IV 30 DATE mg x 1 CARDINAL and MSO4 4 CARDINAL mg and
 pressures dropped to SBPs 208 CARDINAL and HA improved Repeat labetalol with 50 mg
 x QUANTITY 1 CARDINAL and repeated dose of morphine dropped pressures to
 193/134 CARDINAL > labetalol gtt started asa given and HA resolved Head CT negative for
 intracranial bleed and CXR ORG unremarkable ROS ORG cold for past week DATE
 no fevers chills CP SOB N V + ORG diarrhea Upon arrival to the floor patient's BP GPE
 was 191/126 CARDINAL labetalol gtt was not started No sxs no HA She states that she is
 compliant with all her meds and her mother cooks with no salt and she has been adherent to
 diet Past Medical History 1 Lupus 2134 Diagnosed after she began to have swollen fingers a
 rash and painful joints 2 CARDINAL ESRD ORG secondary to SLE ORG 2135 Was
 initially on cytoxan 1 CARDINAL dose every 3 months DATE for 2 years DATE until
 began dialysis 3 CARDINAL times a week in 2137 DATE T Th Sat Awaiting living donor
 transplant from mother 3 CARDINAL HTN 2137 DATE Normal BPs run 180's/120 's Has
 had 1 CARDINAL hypertensive crisis that precipitated seizures in the past 4 Uveitis
 secondary to SLE 4 15 5 HOCM per Echo GPE in 2137 DATE 6 Vaginal bleeding 2139
 DATE 9 20 CARDINAL 7 CARDINAL Multiple episodes of dialysis reactions 8
 CARDINAL Anemia 9 CARDINAL Coag neg Staph PERSON bacteremia and HD line

infection 6 **CARDINAL** 15 10 **CARDINAL** H O UE clot was on coumadin but no longer
 Social History Lives in Location 669 with mother and 16 year old **DATE** brother Graduated
 Name2 NI School **ORG** and then got sick so currently is not working or attending school
 Denies **WORK_OF_ART** any T E D. Family History -No history of SLE -Grandfather **ORG**
 has HTN -Distant history of DM **GPE** -No history of clotting disorders -No other history of
 other autoimmune diseases Physical Exam Vitals **ORG** 98.0 **CARDINAL** 173/51
CARDINAL 86 **CARDINAL** 15 100 **CARDINAL** RA HEENT **PERSON** L eye injected w
 periorbital edema R eye reactive w/ EOMI **ORG** anicteric sclera MMM OP clear Neck supple
 no LAD **GPE** no thyromegaly Cardiac RRR NL **ORG** S1 **PRODUCT** and S2 + S4 III VI
 systolic ejection murmur LUSB radiating to apex and axilla intensifies w/ Valsalva **PERSON**
 no rub Lungs **WORK_OF_ART** CTAB no wheezes rhonchi crackles Abd **PERSON** soft NTND
 NABS no HSM **ORG** no rebound or guarding GU **ORG** no CVAT Ext warm 2 + DP
DATE pulses no C C E L femoral dialysis catheter Neuro AOx3 **PERSON** CN II XII intact
 strength sensation grossly intact Pertinent Results UA **PERSON** mod bld 100 **CARDINAL**
 protein present on prior UAs Radiology CXR **NORP** No acute CP abnormality EKG NSR
 nml **ORG** axis nml intervals borderline LAE LVH **ORG** J point elevation in V2,V3 TWI I aVL
 V5 **CARDINAL** V6 No change when compared to prior on 2139 **DATE** 11 26 **CARDINAL**
 CT HEAD No intracranial hemorrhage Brief Hospital Course **WORK_OF_ART** A P Patient is a
 22 year old **DATE** female with SLE lupus nephritis ESRD **ORG** on HD who presents with
 hypertensive urgency Hypertensive urgency Unclear **ORG** precipitant Possibly secondary to
 pain from worsening uveitis Compliant **PERSON** with meds Denies illicit and tox screen
 negative Patient was started on labetalol drip in ED **GPE** with good BP response and was
 subsequently transitioned to PO **GPE** anti hypertensives in ICU **ORG** with maintenance
 of stable SBPs in 150s-170s **CARDINAL** baseline 170s-190s Per nephrologist 's
 recommendations home lisinopril was increased to 40 mg **QUANTITY** po bid from 40 mg
 po **QUANTITY** qd for better baseline BP **ORG** control No clinical evidence of end organ
 damage UA **ORG** difficult ro interpret in setting of CRF CE 's **LOC** x 1 **CARDINAL**

negative Headache FAC No evidence by CT ORG for intracranial bleed Headaches PERSON were well controlled with morphine sulfate and had resolved by time of discharge Uveitis Followed ORG by outpatient optho specialist Optho PERSON not consulted per patient 's request ESRD ORG Secondary to lupus nephritis On transplant list Patient PERSON received hemodialysis in house with 500 ml QUANTITY ultrafiltrate without complications At dry weight of 45 kg QUANTITY per patient Began Sevalamer ORG 800 TID ORG with meals Given difficulty in interpreting renin and aldosterone levels in acutely ill patients these were not drawn and will need to be drawn at outpatient follow up Medications on Admission Lisinopril 40 PRODUCT mg PO QD FAC Labetalol 600 CARDINAL PO GPE TID Valsartan 320 CARDINAL mg PO QD Clonidine ORG 0.3 CARDINAL mg transdermal QW Prednisone 40 mg PO QD FAC Atropine 1 Hospital1 Prednisolone Acetate 1 ORG Q1H Moxifloxacin eye drops qid Lorazepam 1 LAW mg PO Q4 FAC 6H PRN Discharge ORG Medications 1 CARDINAL Labetalol 200 mg Tablet Sig Three CARDINAL 3 CARDINAL Tablet PO TID 3 CARDINAL times a day Tablet(s 2 Clonidine 0.3 mg/24 QUANTITY hr Patch Weekly Sig ORG One CARDINAL 1 Patch Weekly Transdermal QTHUR every Thursday 3 DATE Atropine 1 Drops Sig One CARDINAL 1 Drop Ophthalmic Hospital1 2 CARDINAL times a day 4 Lorazepam 1 mg Tablet Sig One CARDINAL 1 Tablet PO Q4 6H every 4 to 6 hours TIME as needed 5 CARDINAL Valsartan 160 CARDINAL mg Tablet ORG Sig Two CARDINAL 2 CARDINAL Tablet PO DAILY Daily ORG 6 CARDINAL Prednisolone Acetate 1 Drops ORG Suspension Sig One CARDINAL 1 CARDINAL Drop Ophthalmic Q1H every hour 7 Lisinopril 40 TIME mg Tablet GPE Sig One CARDINAL 1 Tablet PO twice a day Disp:*60 Tablet(s Refills:*2 PERSON 8 Sevelamer 800 mg Tablet PERSON Sig One CARDINAL 1 Tablet PO TID 3 CARDINAL times a day Disp:*90 Tablet(s Refills:*2 9 Prednisone 20 CARDINAL mg Tablet PERSON Sig Two CARDINAL 2 CARDINAL Tablet PO once a day 10 Blood Pressure Kit Kit Sig PERSON One CARDINAL 1 Kit Miscellaneous once a day Disp:*1 Kit Refills:*0 Discharge PERSON Disposition Home Discharge Diagnosis Hypertensive urgency

Discharge Condition Good Discharge Instructions Please take all of your blood pressure medications as prescribed You should adhere to a low salt diet as increased levels of sodium can drive your blood pressure up You are being discharged with a prescription for a home blood pressure monitor which you can use to take **daily DATE** measurements You should call your primary care physician for Name Initial PRE systolic blood pressures greater than **180 CARDINAL** or if you experience headaches nausea vomiting chest pain shortness of breath or any other concerning symptoms **Followup Instructions Please WORK_OF_ART** resume hemodialysis according to your regular schedule You are scheduled to see Dr. **First Name8 PERSON** NamePattern2 Last Name NamePattern1 **4883 DATE** in **the Division of Nephrology ORG** on **Wednesday 2 3 DATE** at **9:30 AM TIME** Please call Telephone Fax 1 435 if you need to reschedule You are scheduled to follow up with your primary care physician Last Name NamePattern4 **First Name4 ORG** NamePattern1 Last Name NamePattern1 **2423 DATE** on **Tuesday 1 26 DATE** at **3:30 PM TIME** Please call Telephone Fax **1 250 CARDINAL** if you need to reschedule You have been referred to see Dr. **First Name4 NamePattern1 Last Name NamePattern1 2539 DATE** in **the Division of Hematology ORG** for further evaluation of your anemia This appointment is scheduled for **2 9 DATE** at **3 p.m. TIME** His office is located on **the Location un of the Hospital Ward Name ORG 23 CARDINAL** Building on **the Hospital1 18 Hospital Ward Name FAC 516 CARDINAL** Please call Dr.[**Name NI **44536 CARDINAL** administrative assistant Doctor **First Name 8982 DATE** at **Telephone Fax 1 32192 ORG** if you need to confirm or reschedule

Word2Vec and t-SNE Visualization Using SpaCy-Processed Data

```
In [81]: from gensim.models import Word2Vec
```

```
#Build corpus
corpus = build_corpus(patients_df_scapy, nlp_spacy)
```

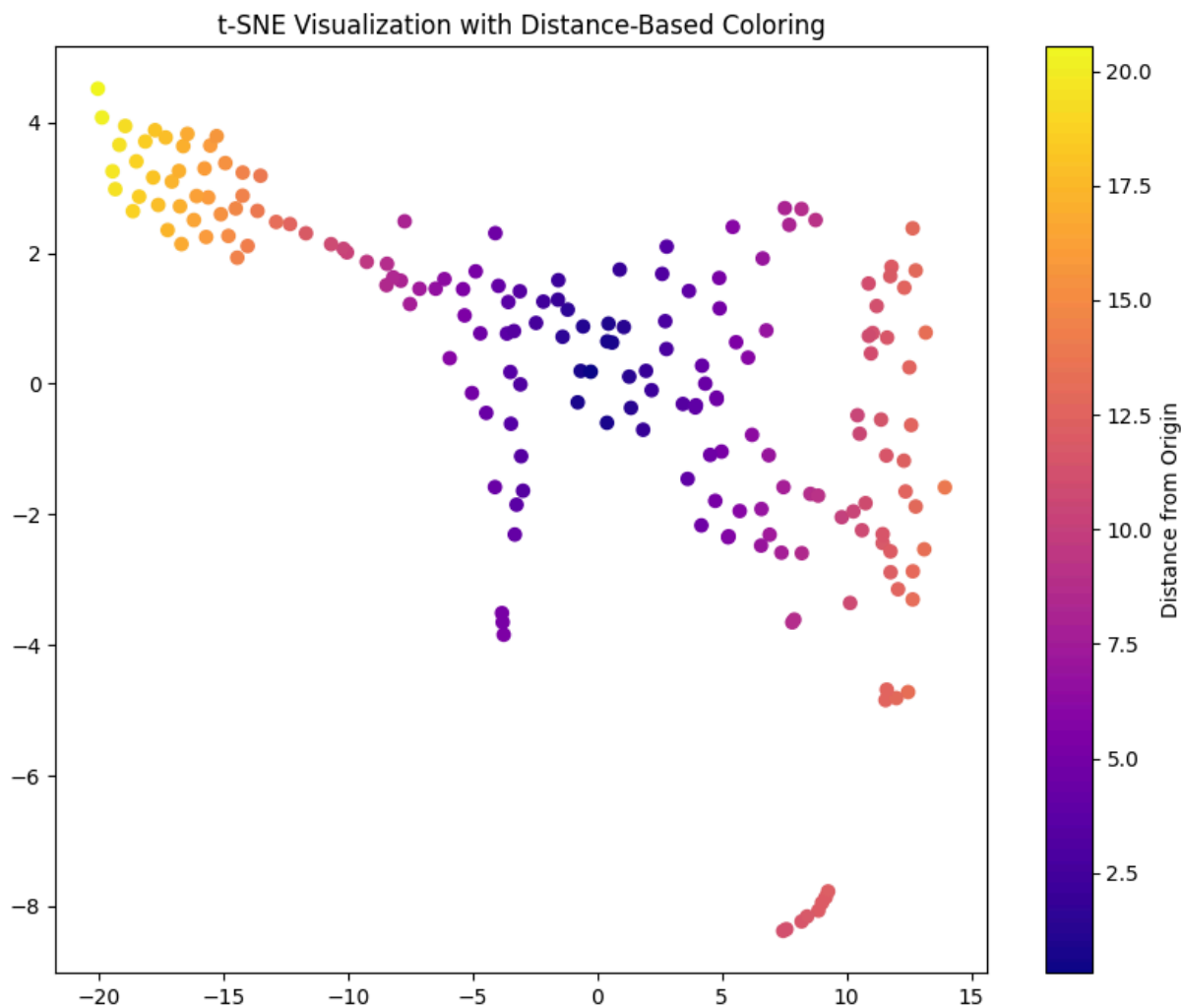
Minimum word count: 12, Maximum word count: 493, Median word count: 213

```
In [137... model_word2vec = Word2Vec(corpus, min_count=50, window=3, vector_size=100)
```

```
In [138... model_word2vec.wv.similar_by_key("Clonidine"), model_word2vec.wv.similar_by_key("BP
```

```
Out[138... ([('Patch Weekly', 0.9990012049674988),  
      ('Q8H', 0.9982399940490723),  
      ('Lisinopril', 0.9979649782180786),  
      ('Patch Weekly Sig', 0.9978789687156677),  
      ('Labetalol', 0.9974281787872314),  
      ('Hydralazine', 0.9969014525413513),  
      ('HCl', 0.9962852597236633),  
      ('hours', 0.9962539076805115),  
      ('Metoprolol', 0.9959376454353333),  
      ('Aspirin', 0.9956783652305603)],  
      [('MICU', 0.9974619150161743),  
      ('transferred', 0.9973183870315552),  
      ('RA', 0.997249186038971),  
      ('Pt', 0.9971693754196167),  
      ('baseline', 0.9967537522315979),  
      ('floor', 0.9967240691184998),  
      ('admitted', 0.9966705441474915),  
      ('patient 's', 0.9966193437576294),  
      ('EKG', 0.9965095520019531),  
      ('hypertension', 0.9962207674980164)])]
```

```
In [140... vocabs = model_word2vec.wv.key_to_index.keys()  
new_v = np.array(list(vocabs))  
tsne_plot(model_word2vec, new_v)
```

```
In [ ]: import gensim
import gensim.downloader as api

info = api.info() # show info about available models/datasets
print(info)
pretrained_model= api.load("glove-wiki-gigaword-50")
vocabs = pretrained_model.key_to_index.keys()
new_v = np.array(list(vocabs))
tsne_plot(pretrained_model,new_v, True)
```

```
{'corpora': {'semeval-2016-2017-task3-subtaskBC': {'num_records': -1, 'record_format': 'dict', 'file_size': 6344358, 'reader_code': 'https://github.com/RaRe-Technologies/gensim-data/releases/download/semeval-2016-2017-task3-subtaskB-eng/__init__.py', 'license': 'All files released for the task are free for general research use', 'fields': {'2016-train': ['...'], '2016-dev': ['...'], '2017-test': ['...'], '2016-test': ['...']}, 'description': 'SemEval 2016 / 2017 Task 3 Subtask B and C datasets contain train+development (317 original questions, 3,169 related questions, and 31,690 comments), and test datasets in English. The description of the tasks and the collected data is given in sections 3 and 4.1 of the task paper http://alt.qcri.org/semeval2016/task3/data/uploads/semeval2016-task3-report.pdf linked in section "Papers" of https://github.com/RaRe-Technologies/gensim-data/issues/18.', 'checksum': '701ea67acd82e75f95e1d8e62fb0ad29', 'file_name': 'semeval-2016-2017-task3-subtaskBC.gz', 'read_more': ['http://alt.qcri.org/semeval2017/task3/', 'http://alt.qcri.org/semeval2017/task3/data/uploads/semeval2017-task3.pdf', 'https://github.com/RaRe-Technologies/gensim-data/issues/18', 'https://github.com/Witiko/semeval-2016\_2017-task3-subtaskB-english'], 'parts': 1}, 'semeval-2016-2017-task3-subtaskA-unannotated': {'num_records': 189941, 'record_format': 'dict', 'file_size': 234373151, 'reader_code': 'https://github.com/RaRe-Technologies/gensim-data/releases/download/semeval-2016-2017-task3-subtaskA-unannotated-eng/__init__.py', 'license': 'These datasets are free for general research use.', 'fields': {'THREAD_SEQUENCE': '', 'RelQuestion': {'RELQ_CATEGORY': 'question category, according to the Qatar Living taxonomy', 'RELQ_DATE': 'date of posting', 'RELQ_ID': 'question identifier', 'RELQ_USERID': 'identifier of the user asking the question', 'RELQ_USERNAME': 'name of the user asking the question', 'RelQBody': 'body of question', 'RelQSubject': 'subject of question'}, 'RelComments': [{'RelCText': 'text of answer', 'RELC_USERID': 'identifier of the user posting the comment', 'RELC_ID': 'comment identifier', 'RELC_USERNAME': 'name of the user posting the comment', 'RELC_DATE': 'date of posting'}]}, 'description': 'SemEval 2016 / 2017 Task 3 Subtask A unannotated dataset contains 189,941 questions and 1,894,456 comments in English collected from the Community Question Answering (CQA) web forum of Qatar Living. These can be used as a corpus for language modelling.', 'checksum': '2de0e2f2c4f91c66ae4fcf58d50ba816', 'file_name': 'semeval-2016-2017-task3-subtaskA-unannotated.gz', 'read_more': ['http://alt.qcri.org/semeval2016/task3/', 'http://alt.qcri.org/semeval2016/task3/data/uploads/semeval2016-task3-report.pdf', 'https://github.com/RaRe-Technologies/gensim-data/issues/18', 'https://github.com/Witiko/semeval-2016\_2017-task3-subtaskA-unannotated-english'], 'parts': 1}, 'patent-2017': {'num_records': 353197, 'record_format': 'dict', 'file_size': 3087262469, 'reader_code': 'https://github.com/RaRe-Technologies/gensim-data/releases/download/patent-2017/__init__.py', 'license': 'not found', 'description': 'Patent Grant Full Text. Contains the full text including tables, sequence data and in-line mathematical expressions of each patent grant issued in 2017.', 'checksum-0': '818501f0b9af62d3b88294d86d509f8f', 'checksum-1': '66c05635c1d3c7a19b4a335829d09ffa', 'file_name': 'patent-2017.gz', 'read_more': ['http://patents.reedtech.com/pgrbft.php'], 'parts': 2}, 'quora-duplicate-questions': {'num_records': 404290, 'record_format': 'dict', 'file_size': 21684784, 'reader_code': 'https://github.com/RaRe-Technologies/gensim-data/releases/download/quora-duplicate-questions/__init__.py', 'license': 'probably https://www.quora.com/about/tos', 'fields': {'question1': 'the full text of each question', 'question2': 'the full text of each question', 'qid1': 'unique ids of each question', 'qid2': 'unique ids of each question', 'id': 'the id of a training set question pair', 'is_duplicate': 'the target variable, set to 1 if question1 and question2 have essentially the same meaning, and 0 otherwise'}, 'description': 'Over 400,000 lines of potential question duplicate pairs. Each line contains IDs for each question in the pair, the full text for each question, and a binary value that indicates whether the line contains a duplicate pair or not.', 'checksum': 'd7cfa7fbc6e2ec71ab74c495586c6365', 'file_name': 'quora-duplicate-questions.gz', 'read_more': ['https://data.quora.com/First-Quora-Dataset-Release-Question-Pairs'], 'parts': 1}, 'wiki-english-20171001': {'num_records': 4924894, 'record_format': 'dict', 'file_size': 6516051717, 'reader_code': 'http
```

```
s://github.com/RaRe-Technologies/gensim-data/releases/download/wiki-english-20171001/
__init__.py', 'license': 'https://dumps.wikimedia.org/legal.html', 'fields': {'section_texts': 'list of body of sections', 'section_titles': 'list of titles of sections', 'title': 'Title of wiki article'}, 'description': 'Extracted Wikipedia dump from October 2017. Produced by `python -m gensim.scripts.segment_wiki -f enwiki-20171001-pages-articles.xml.bz2 -o wiki-en.gz`, 'checksum-0': 'a7d7d7fd41ea7e2d7fa32ec1bb640d71', 'checksum-1': 'b2683e3356ffbc3b6c2dca6e9801f9f', 'checksum-2': 'c5cde2a9ae77b3c4ebce804f6df542c2', 'checksum-3': '00b71144ed5e3aeeb885de84f7452b81', 'file_name': 'wiki-english-20171001.gz', 'read_more': ['https://dumps.wikimedia.org/enwiki/20171001/'], 'parts': 4}, 'text8': {'num_records': 1701, 'record_format': 'list of str (tokens)', 'file_size': 33182058, 'reader_code': 'https://github.com/RaRe-Technologies/gensim-data/releases/download/text8/__init__.py', 'license': 'not found', 'description': 'First 100,000,000 bytes of plain text from Wikipedia. Used for testing purposes; see wiki-english-* for proper full Wikipedia datasets.', 'checksum': '68799af40b6bda07dfa47a32612e5364', 'file_name': 'text8.gz', 'read_more': ['http://matmahoney.net/dc/textdata.html'], 'parts': 1}, 'fake-news': {'num_records': 12999, 'record_format': 'dict', 'file_size': 20102776, 'reader_code': 'https://github.com/RaRe-Technologies/gensim-data/releases/download/fake-news/__init__.py', 'license': 'https://creativecommons.org/publicdomain/zero/1.0/', 'fields': {'crawled': 'date the story was archived', 'ord_in_thread': '', 'published': 'date published', 'participants_count': 'number of participants', 'shares': 'number of Facebook shares', 'replies_count': 'number of replies', 'main_img_url': 'image from story', 'spam_score': 'data from webhose.io', 'uuid': 'unique identifier', 'language': 'data from webhose.io', 'title': 'title of story', 'country': 'data from webhose.io', 'domain_rank': 'data from webhose.io', 'author': 'author of story', 'comments': 'number of Facebook comments', 'site_url': 'site URL from BS detector', 'text': 'text of story', 'thread_title': '', 'type': 'type of website (label from BS detector)', 'likes': 'number of Facebook likes'}, 'description': 'News dataset, contains text and metadata from 244 websites and represents 12,999 posts in total from a specific window of 30 days. The data was pulled using the webhose.io API, and because it's coming from their crawler, not all websites identified by their BS Detector are present in this dataset. Data sources that were missing a label were simply assigned a label of 'bs'. There are (ostensibly) no genuine, reliable, or trustworthy news sources represented in this dataset (so far), so don't trust anything you read.', 'checksum': '5e64e942df13219465927f92dcefd5fe', 'file_name': 'fake-news.gz', 'read_more': ['https://www.kaggle.com/mrisdal/fake-news'], 'parts': 1}, '20-newsgroups': {'num_records': 18846, 'record_format': 'dict', 'file_size': 14483581, 'reader_code': 'https://github.com/RaRe-Technologies/gensim-data/releases/download/20-newsgroups/__init__.py', 'license': 'not found', 'fields': {'topic': 'name of topic (20 variant of possible values)', 'set': 'marker of original split (possible values 'train' and 'test')', 'data': '', 'id': 'original id in folder name'}, 'description': 'The notorious collection of approximately 20,000 newsgroup posts, partitioned (nearly) evenly across 20 different newsgroups.', 'checksum': 'c92fd4f6640a86d5ba89eaad818a9891', 'file_name': '20-newsgroups.gz', 'read_more': ['http://qwone.com/~jason/20Newsgroups/'], 'parts': 1}, '__testing_matrix-synopsis': {'description': '[THIS IS ONLY FOR TESTING] Synopsis of the movie matrix.', 'checksum': '1767ac93a089b43899d54944b07d9dc5', 'file_name': '__testing_matrix-synopsis.gz', 'read_more': ['http://www.imdb.com/title/tt0133093/plotsummary?ref=ttpl_pl_syn#synopsis'], 'parts': 1}, '__testing_multipart-matrix-synopsis': {'description': '[THIS IS ONLY FOR TESTING] Synopsis of the movie matrix.', 'checksum-0': 'c8b0c7d8cf562b1b632c262a173ac338', 'checksum-1': '5ff7fc6818e9a5d9bc1cf12c35ed8b96', 'checksum-2': '966db9d274d125beaac7987202076cba', 'file_name': '__testing_multipart-matrix-synopsis.gz', 'read_more': ['http://www.imdb.com/title/tt0133093/plotsummary?ref=ttpl_pl_syn#synopsis'], 'parts': 3}}, 'models': {'fasttext-wiki-news-subwords-300': {'num_records': 999999, 'file_size': 1005007116, 'base_dataset': 'Wikipedia 2017, UMBC webbase corpus and statmt.org news dataset (16B tokens)', 'reader_code': 'https://github.com/RaRe-Technologies/gensim-data/releases/download/fasttext-wiki-ne
```

ws-subwords-300/___init__.py', 'license': 'https://creativecommons.org/licenses/by-sa/3.0/', 'parameters': {'dimension': 300}, 'description': '1 million word vectors trained on Wikipedia 2017, UMBC webbase corpus and statmt.org news dataset (16B tokens).', 'read_more': ['https://fasttext.cc/docs/en/english-vectors.html', 'https://arxiv.org/abs/1712.09405', 'https://arxiv.org/abs/1607.01759'], 'checksum': 'de2bb3a20c46ce65c9c131e1ad9a77af', 'file_name': 'fasttext-wiki-news-subwords-300.gz', 'parts': 1}, 'conceptnet-numberbatch-17-06-300': {'num_records': 1917247, 'file_size': 1225497562, 'base_dataset': 'ConceptNet, word2vec, GloVe, and OpenSubtitles 2016', 'reader_code': 'https://github.com/RaRe-Technologies/gensim-data/releases/download/conceptnet-numberbatch-17-06-300/___init__.py', 'license': 'https://github.com/commonsense/conceptnet-numberbatch/blob/master/LICENSE.txt', 'parameters': {'dimension': 300}, 'description': 'ConceptNet Numberbatch consists of state-of-the-art semantic vectors (also known as word embeddings) that can be used directly as a representation of word meanings or as a starting point for further machine learning. ConceptNet Numberbatch is part of the ConceptNet open data project. ConceptNet provides lots of ways to compute with word meanings, one of which is word embeddings. ConceptNet Numberbatch is a snapshot of just the word embeddings. It is built using an ensemble that combines data from ConceptNet, word2vec, GloVe, and OpenSubtitles 2016, using a variation on retrofitting.', 'read_more': ['http://aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14972', 'https://github.com/commonsense/conceptnet-numberbatch', 'http://conceptnet.io/'], 'checksum': 'fd642d457adcd0ea94da0cd21b150847', 'file_name': 'conceptnet-numberbatch-17-06-300.gz', 'parts': 1}, 'word2vec-ruscorpora-300': {'num_records': 184973, 'file_size': 208427381, 'base_dataset': 'Russian National Corpus (about 250M words)', 'reader_code': 'https://github.com/RaRe-Technologies/gensim-data/releases/download/word2vec-ruscorpora-300/___init__.py', 'license': 'https://creativecommons.org/licenses/by/4.0/deed.en', 'parameters': {'dimension': 300, 'window_size': 10}, 'description': 'Word2vec Continuous Skipgram vectors trained on full Russian National Corpus (about 250M words). The model contains 185K words.', 'preprocessing': 'The corpus was lemmatized and tagged with Universal PoS', 'read_more': ['https://www.academia.edu/24306935/WebVectors_a_Toolkit_for_Building_Web_Interfaces_for_Vector_Semantic_Models', 'http://rusvectors.org/en/', 'https://github.com/RaRe-Technologies/gensim-data/issues/3'], 'checksum': '9bdebdc8ae6d17d20839dd9b5af10bc4', 'file_name': 'word2vec-ruscorpora-300.gz', 'parts': 1}, 'word2vec-google-news-300': {'num_records': 3000000, 'file_size': 1743563840, 'base_dataset': 'Google News (about 100 billion words)', 'reader_code': 'https://github.com/RaRe-Technologies/gensim-data/releases/download/word2vec-google-news-300/___init__.py', 'license': 'not found', 'parameters': {'dimension': 300}, 'description': 'Pre-trained vectors trained on a part of the Google News dataset (about 100 billion words). The model contains 300-dimensional vectors for 3 million words and phrases. The phrases were obtained using a simple data-driven approach described in 'Distributed Representations of Words and Phrases and their Compositionality' (https://code.google.com/archive/p/word2vec/).", 'read_more': ['https://code.google.com/archive/p/word2vec/', 'https://arxiv.org/abs/1301.3781', 'https://arxiv.org/abs/1310.4546', 'https://www.microsoft.com/en-us/research/publication/linguistic-regularities-in-continuous-space-word-representations/?from=http%3A%2F%2Fresearch.microsoft.com%2Fpubs%2F189726%2Frvecs.pdf'], 'checksum': 'a5e5354d40acb95f9ec66d5977d140ef', 'file_name': 'word2vec-google-news-300.gz', 'parts': 1}, 'glove-wiki-gigaword-50': {'num_records': 400000, 'file_size': 69182535, 'base_dataset': 'Wikipedia 2014 + Gigaword 5 (6B tokens, uncased)', 'reader_code': 'https://github.com/RaRe-Technologies/gensim-data/releases/download/glove-wiki-gigaword-50/___init__.py', 'license': 'http://opendatacommons.org/licenses/pddl/', 'parameters': {'dimension': 50}, 'description': 'Pre-trained vectors based on Wikipedia 2014 + Gigaword, 5.6B tokens, 400K vocab, uncased (https://nlp.stanford.edu/projects/glove/).', 'preprocessing': 'Converted to w2v format with `python -m gensim.scripts.glove2word2vec -i <fname> -o glove-wiki-gigaword-50.txt`.', 'read_more': ['https://nlp.stanford.edu/projects/glove/', 'https://nlp.stanford.edu/pubs/glove.pdf'], 'checksum': 'c289bc5d7f2f02c6dc9f2f9b67641813', 'file_name': 'glove-wiki-gigaword-50.gz', 'parts': 1}, 'glove-wiki-giga


```

word-100': {'num_records': 400000, 'file_size': 134300434, 'base_dataset': 'Wikipedi
a 2014 + Gigaword 5 (6B tokens, uncased)', 'reader_code': 'https://github.com/RaRe-T
echnologies/gensim-data/releases/download/glove-wiki-gigaword-100/__init__.py', 'lic
ense': 'http://opendatacommons.org/licenses/pddl/', 'parameters': {'dimension': 10
0}, 'description': 'Pre-trained vectors based on Wikipedia 2014 + Gigaword 5.6B toke
ns, 400K vocab, uncased (https://nlp.stanford.edu/projects/glove/).', 'preprocessin
g': 'Converted to w2v format with `python -m gensim.scripts.glove2word2vec -i <fname
> -o glove-wiki-gigaword-100.txt`. ', 'read_more': ['https://nlp.stanford.edu/project
s/glove/', 'https://nlp.stanford.edu/pubs/glove.pdf'], 'checksum': '40ec481866001177
b8cd4cb0df92924f', 'file_name': 'glove-wiki-gigaword-100.gz', 'parts': 1}, 'glove-wi
ki-gigaword-200': {'num_records': 400000, 'file_size': 264336934, 'base_dataset': 'W
ikipedia 2014 + Gigaword 5 (6B tokens, uncased)', 'reader_code': 'https://github.co
m/RaRe-Technologies/gensim-data/releases/download/glove-wiki-gigaword-200/__init__.p
y', 'license': 'http://opendatacommons.org/licenses/pddl/', 'parameters': {'dimensio
n': 200}, 'description': 'Pre-trained vectors based on Wikipedia 2014 + Gigaword, 5.
6B tokens, 400K vocab, uncased (https://nlp.stanford.edu/projects/glove/).', 'prepro
cessing': 'Converted to w2v format with `python -m gensim.scripts.glove2word2vec -i
<fname> -o glove-wiki-gigaword-200.txt`. ', 'read_more': ['https://nlp.stanford.edu/p
rojects/glove/', 'https://nlp.stanford.edu/pubs/glove.pdf'], 'checksum': '59652db361
b7a87ee73834a6c391dfc1', 'file_name': 'glove-wiki-gigaword-200.gz', 'parts': 1}, 'gl
ove-wiki-gigaword-300': {'num_records': 400000, 'file_size': 394362229, 'base_databa
se': 'Wikipedia 2014 + Gigaword 5 (6B tokens, uncased)', 'reader_code': 'https://gith
ub.com/RaRe-Technologies/gensim-data/releases/download/glove-wiki-gigaword-300/__ini
t__.py', 'license': 'http://opendatacommons.org/licenses/pddl/', 'parameters': {'dim
ension': 300}, 'description': 'Pre-trained vectors based on Wikipedia 2014 + Gigawor
d, 5.6B tokens, 400K vocab, uncased (https://nlp.stanford.edu/projects/glove/).', 'p
reprocessing': 'Converted to w2v format with `python -m gensim.scripts.glove2word2ve
c -i <fname> -o glove-wiki-gigaword-300.txt`. ', 'read_more': ['https://nlp.stanford.
edu/projects/glove/', 'https://nlp.stanford.edu/pubs/glove.pdf'], 'checksum': '29e93
29ac2241937d55b852e8284e89b', 'file_name': 'glove-wiki-gigaword-300.gz', 'parts':
1}, 'glove-twitter-25': {'num_records': 1193514, 'file_size': 109885004, 'base_databa
se': 'Twitter (2B tweets, 27B tokens, 1.2M vocab, uncased)', 'reader_code': 'http
s://github.com/RaRe-Technologies/gensim-data/releases/download/glove-twitter-25/__in
it__.py', 'license': 'http://opendatacommons.org/licenses/pddl/', 'parameters': {'di
mension': 25}, 'description': 'Pre-trained vectors based on 2B tweets, 27B tokens,
1.2M vocab, uncased (https://nlp.stanford.edu/projects/glove/).', 'preprocessing':
'Converted to w2v format with `python -m gensim.scripts.glove2word2vec -i <fname> -o
glove-twitter-25.txt`. ', 'read_more': ['https://nlp.stanford.edu/projects/glove/',
'https://nlp.stanford.edu/pubs/glove.pdf'], 'checksum': '50db0211d7e7a2dcd362c6b7747
62793', 'file_name': 'glove-twitter-25.gz', 'parts': 1}, 'glove-twitter-50': {'num_r
ecords': 1193514, 'file_size': 209216938, 'base_dataset': 'Twitter (2B tweets, 27B t
okens, 1.2M vocab, uncased)', 'reader_code': 'https://github.com/RaRe-Technologies/g
ensim-data/releases/download/glove-twitter-50/__init__.py', 'license': 'http://opend
atacommons.org/licenses/pddl/', 'parameters': {'dimension': 50}, 'description': 'Pre
-trained vectors based on 2B tweets, 27B tokens, 1.2M vocab, uncased (https://nlp.st
anford.edu/projects/glove/).', 'preprocessing': 'Converted to w2v format with `python
-m gensim.scripts.glove2word2vec -i <fname> -o glove-twitter-50.txt`. ', 'read_more':
['https://nlp.stanford.edu/projects/glove/', 'https://nlp.stanford.edu/pubs/glove.pd
f'], 'checksum': 'c168f18641f8c8a00fe30984c4799b2b', 'file_name': 'glove-twitter-50.
gz', 'parts': 1}, 'glove-twitter-100': {'num_records': 1193514, 'file_size': 4059329
91, 'base_dataset': 'Twitter (2B tweets, 27B tokens, 1.2M vocab, uncased)', 'reader_
code': 'https://github.com/RaRe-Technologies/gensim-data/releases/download/glove-twi
tter-100/__init__.py', 'license': 'http://opendatacommons.org/licenses/pddl/', 'para
meters': {'dimension': 100}, 'description': 'Pre-trained vectors based on 2B tweet
s, 27B tokens, 1.2M vocab, uncased (https://nlp.stanford.edu/projects/glove/).', 'pre
processing': 'Converted to w2v format with `python -m gensim.scripts.glove2word2vec

```

```
-i <fname> -o glove-twitter-100.txt`.', 'read_more': ['https://nlp.stanford.edu/projects/glove/', 'https://nlp.stanford.edu/pubs/glove.pdf'], 'checksum': 'b04f7bed38756d64cf55b58ce7e97b15', 'file_name': 'glove-twitter-100.gz', 'parts': 1}, 'glove-twitter-200': {'num_records': 1193514, 'file_size': 795373100, 'base_dataset': 'Twitter (2B tweets, 27B tokens, 1.2M vocab, uncased)', 'reader_code': 'https://github.com/RaRe-Technologies/gensim-data/releases/download/glove-twitter-200/__init__.py', 'license': 'http://opendatacommons.org/licenses/pddl/', 'parameters': {'dimension': 200}, 'description': 'Pre-trained vectors based on 2B tweets, 27B tokens, 1.2M vocab, uncased (https://nlp.stanford.edu/projects/glove/).', 'preprocessing': 'Converted to w2v format with `python -m gensim.scripts.glove2word2vec -i <fname> -o glove-twitter-200.txt`.', 'read_more': ['https://nlp.stanford.edu/projects/glove/', 'https://nlp.stanford.edu/pubs/glove.pdf'], 'checksum': 'e52e8392d1860b95d5308a525817d8f9', 'file_name': 'glove-twitter-200.gz', 'parts': 1}, '__testing_word2vec-matrix-synopsis': {'description': '[THIS IS ONLY FOR TESTING] Word vecrors of the movie matrix.', 'parameters': {'dimensions': 50}, 'preprocessing': 'Converted to w2v using a preprocessed corpus. Converted to w2v format with `python3.5 -m gensim.models.word2vec -train <input_filename> -iter 50 -output <output_filename>`.', 'read_more': [], 'checksum': '534dcb8b56a360977a269b7bfc62d124', 'file_name': '__testing_word2vec-matrix-synopsis.gz', 'parts': 1}}}
```

Using SciSpacy

In [37]: `!pip install scispacy`

Collecting scispacy

Downloading scispacy-0.5.5-py3-none-any.whl.metadata (18 kB)
 Requirement already satisfied: spacy<3.8.0,>=3.7.0 in /usr/local/lib/python3.11/dist-packages (from scispacy) (3.7.5)
 Requirement already satisfied: scipy in /usr/local/lib/python3.11/dist-packages (from scispacy) (1.13.1)
 Requirement already satisfied: requests<3.0.0,>=2.0.0 in /usr/local/lib/python3.11/dist-packages (from scispacy) (2.32.3)
 Collecting conllu (from scispacy)
 Downloading conllu-6.0.0-py3-none-any.whl.metadata (21 kB)
 Requirement already satisfied: numpy in /usr/local/lib/python3.11/dist-packages (from scispacy) (1.26.4)
 Requirement already satisfied: joblib in /usr/local/lib/python3.11/dist-packages (from scispacy) (1.4.2)
 Requirement already satisfied: scikit-learn>=0.20.3 in /usr/local/lib/python3.11/dist-packages (from scispacy) (1.6.1)
 Collecting pysbd (from scispacy)
 Downloading pysbd-0.3.4-py3-none-any.whl.metadata (6.1 kB)
 Collecting nmslib-metabrainz==2.1.3 (from scispacy)
 Downloading nmslib-metabrainz-2.1.3-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (956 bytes)
 Collecting pybind11>=2.2.3 (from nmslib-metabrainz==2.1.3->scispacy)
 Downloading pybind11-2.13.6-py3-none-any.whl.metadata (9.5 kB)
 Requirement already satisfied: psutil in /usr/local/lib/python3.11/dist-packages (from nmslib-metabrainz==2.1.3->scispacy) (5.9.5)
 Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests<3.0.0,>=2.0.0->scispacy) (3.4.1)
 Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-packages (from requests<3.0.0,>=2.0.0->scispacy) (3.10)
 Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-packages (from requests<3.0.0,>=2.0.0->scispacy) (2.3.0)
 Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-packages (from requests<3.0.0,>=2.0.0->scispacy) (2025.1.31)
 Requirement already satisfied: threadpoolctl>=3.1.0 in /usr/local/lib/python3.11/dist-packages (from scikit-learn>=0.20.3->scispacy) (3.5.0)
 Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.11 in /usr/local/lib/python3.11/dist-packages (from spacy<3.8.0,>=3.7.0->scispacy) (3.0.12)
 Requirement already satisfied: spacy-loggers<2.0.0,>=1.0.0 in /usr/local/lib/python3.11/dist-packages (from spacy<3.8.0,>=3.7.0->scispacy) (1.0.5)
 Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in /usr/local/lib/python3.11/dist-packages (from spacy<3.8.0,>=3.7.0->scispacy) (1.0.12)
 Requirement already satisfied: cymem<2.1.0,>=2.0.2 in /usr/local/lib/python3.11/dist-packages (from spacy<3.8.0,>=3.7.0->scispacy) (2.0.11)
 Requirement already satisfied: preshed<3.1.0,>=3.0.2 in /usr/local/lib/python3.11/dist-packages (from spacy<3.8.0,>=3.7.0->scispacy) (3.0.9)
 Requirement already satisfied: thinc<8.3.0,>=8.2.2 in /usr/local/lib/python3.11/dist-packages (from spacy<3.8.0,>=3.7.0->scispacy) (8.2.5)
 Requirement already satisfied: wasabi<1.2.0,>=0.9.1 in /usr/local/lib/python3.11/dist-packages (from spacy<3.8.0,>=3.7.0->scispacy) (1.1.3)
 Requirement already satisfied: srsly<3.0.0,>=2.4.3 in /usr/local/lib/python3.11/dist-packages (from spacy<3.8.0,>=3.7.0->scispacy) (2.5.1)
 Requirement already satisfied: catalogue<2.1.0,>=2.0.6 in /usr/local/lib/python3.11/dist-packages (from spacy<3.8.0,>=3.7.0->scispacy) (2.0.10)
 Requirement already satisfied: weasel<0.5.0,>=0.1.0 in /usr/local/lib/python3.11/dist-packages (from spacy<3.8.0,>=3.7.0->scispacy) (0.4.1)
 Requirement already satisfied: typer<1.0.0,>=0.3.0 in /usr/local/lib/python3.11/dist

```

-packages (from spacy<3.8.0,>=3.7.0->scispacy) (0.15.1)
Requirement already satisfied: tqdm<5.0.0,>=4.38.0 in /usr/local/lib/python3.11/dist-
-packages (from spacy<3.8.0,>=3.7.0->scispacy) (4.67.1)
Requirement already satisfied: pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4 in /usr/local/li
b/python3.11/dist-packages (from spacy<3.8.0,>=3.7.0->scispacy) (2.10.6)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.11/dist-packages (fr
om spacy<3.8.0,>=3.7.0->scispacy) (3.1.5)
Requirement already satisfied: setuptools in /usr/local/lib/python3.11/dist-packages
(from spacy<3.8.0,>=3.7.0->scispacy) (75.1.0)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.11/dist-pac
kages (from spacy<3.8.0,>=3.7.0->scispacy) (24.2)
Requirement already satisfied: langcodes<4.0.0,>=3.2.0 in /usr/local/lib/python3.11/
dist-packages (from spacy<3.8.0,>=3.7.0->scispacy) (3.5.0)
Requirement already satisfied: language-data>=1.2 in /usr/local/lib/python3.11/dist-
packages (from langcodes<4.0.0,>=3.2.0->spacy<3.8.0,>=3.7.0->scispacy) (1.3.0)
Requirement already satisfied: annotated-types>=0.6.0 in /usr/local/lib/python3.11/d
ist-packages (from pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4->spacy<3.8.0,>=3.7.0->scispa
cy) (0.7.0)
Requirement already satisfied: pydantic-core==2.27.2 in /usr/local/lib/python3.11/di
st-packages (from pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4->spacy<3.8.0,>=3.7.0->scispac
y) (2.27.2)
Requirement already satisfied: typing-extensions>=4.12.2 in /usr/local/lib/python3.1
1/dist-packages (from pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4->spacy<3.8.0,>=3.7.0->sci
spacy) (4.12.2)
Requirement already satisfied: blis<0.8.0,>=0.7.8 in /usr/local/lib/python3.11/dist-
packages (from thinc<8.3.0,>=8.2.2->spacy<3.8.0,>=3.7.0->scispacy) (0.7.11)
Requirement already satisfied: confection<1.0.0,>=0.0.1 in /usr/local/lib/python3.1
1/dist-packages (from thinc<8.3.0,>=8.2.2->spacy<3.8.0,>=3.7.0->scispacy) (0.1.5)
Requirement already satisfied: click>=8.0.0 in /usr/local/lib/python3.11/dist-packag
es (from typer<1.0.0,>=0.3.0->spacy<3.8.0,>=3.7.0->scispacy) (8.1.8)
Requirement already satisfied: shellingham>=1.3.0 in /usr/local/lib/python3.11/dist-
packages (from typer<1.0.0,>=0.3.0->spacy<3.8.0,>=3.7.0->scispacy) (1.5.4)
Requirement already satisfied: rich>=10.11.0 in /usr/local/lib/python3.11/dist-packa
ges (from typer<1.0.0,>=0.3.0->spacy<3.8.0,>=3.7.0->scispacy) (13.9.4)
Requirement already satisfied: cloudpathlib<1.0.0,>=0.7.0 in /usr/local/lib/python3.
11/dist-packages (from weasel<0.5.0,>=0.1.0->spacy<3.8.0,>=3.7.0->scispacy) (0.20.0)
Requirement already satisfied: smart-open<8.0.0,>=5.2.1 in /usr/local/lib/python3.1
1/dist-packages (from weasel<0.5.0,>=0.1.0->spacy<3.8.0,>=3.7.0->scispacy) (7.1.0)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.11/dist-pac
kages (from jinja2->spacy<3.8.0,>=3.7.0->scispacy) (3.0.2)
Requirement already satisfied: marisa-trie>=1.1.0 in /usr/local/lib/python3.11/dist-
packages (from language-data>=1.2->langcodes<4.0.0,>=3.2.0->spacy<3.8.0,>=3.7.0->sci
spacy) (1.2.1)
Requirement already satisfied: markdown-it-py>=2.2.0 in /usr/local/lib/python3.11/di
st-packages (from rich>=10.11.0->typer<1.0.0,>=0.3.0->spacy<3.8.0,>=3.7.0->scispacy)
(3.0.0)
Requirement already satisfied: pygments<3.0.0,>=2.13.0 in /usr/local/lib/python3.11/
dist-packages (from rich>=10.11.0->typer<1.0.0,>=0.3.0->spacy<3.8.0,>=3.7.0->scispac
y) (2.18.0)
Requirement already satisfied: wrapt in /usr/local/lib/python3.11/dist-packages (fro
m smart-open<8.0.0,>=5.2.1->weasel<0.5.0,>=0.1.0->spacy<3.8.0,>=3.7.0->scispacy) (1.
17.2)
Requirement already satisfied: mdurl~=0.1 in /usr/local/lib/python3.11/dist-packages
(from markdown-it-py>=2.2.0->rich>=10.11.0->typer<1.0.0,>=0.3.0->spacy<3.8.0,>=3.7.0
->scispacy) (0.1.2)
Downloading scispacy-0.5.5-py3-none-any.whl (46 kB)

```

```

 46.2/46.2 kB 4.7 MB/s eta 0:00:00
Downloading nmslib_metabrainz-2.1.3-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_
x86_64.whl (14.1 MB)
 14.1/14.1 MB 62.6 MB/s eta 0:00:00
Downloading conllu-6.0.0-py3-none-any.whl (16 kB)
Downloading pysbd-0.3.4-py3-none-any.whl (71 kB)
 71.1/71.1 kB 7.3 MB/s eta 0:00:00
Downloading pybind11-2.13.6-py3-none-any.whl (243 kB)
 243.3/243.3 kB 26.6 MB/s eta 0:00:00
Installing collected packages: pysbd, pybind11, conllu, nmslib-metabrainz, scispacy
Successfully installed conllu-6.0.0 nmslib-metabrainz-2.1.3 pybind11-2.13.6 pysbd-0.
3.4 scispacy-0.5.5
```

In [39]: `!pip install https://s3-us-west-2.amazonaws.com/ai2-s2-scispacy/releases/v0.5.3/en_`

Collecting https://s3-us-west-2.amazonaws.com/ai2-s2-scispace/releases/v0.5.3/en_core_sci_md-0.5.3.tar.gz

Downloading https://s3-us-west-2.amazonaws.com/ai2-s2-scispace/releases/v0.5.3/en_core_sci_md-0.5.3.tar.gz (119.1 MB)

119.1/119.1 MB 14.0 MB/s eta 0:00:00

Preparing metadata (setup.py) ... done

Collecting spacy<3.7.0,>=3.6.1 (from en_core_sci_md==0.5.3)

Downloading spacy-3.6.1-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (25 kB)

Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.11 in /usr/local/lib/python3.11/dist-packages (from spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (3.0.12)

Requirement already satisfied: spacy-loggers<2.0.0,>=1.0.0 in /usr/local/lib/python3.11/dist-packages (from spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (1.0.5)

Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in /usr/local/lib/python3.11/dist-packages (from spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (1.0.12)

Requirement already satisfied: cymem<2.1.0,>=2.0.2 in /usr/local/lib/python3.11/dist-packages (from spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (2.0.11)

Requirement already satisfied: preshed<3.1.0,>=3.0.2 in /usr/local/lib/python3.11/dist-packages (from spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (3.0.9)

Collecting thinc<8.2.0,>=8.1.8 (from spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3)

Downloading thinc-8.1.12-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (15 kB)

Requirement already satisfied: wasabi<1.2.0,>=0.9.1 in /usr/local/lib/python3.11/dist-packages (from spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (1.1.3)

Requirement already satisfied: srsly<3.0.0,>=2.4.3 in /usr/local/lib/python3.11/dist-packages (from spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (2.5.1)

Requirement already satisfied: catalogue<2.1.0,>=2.0.6 in /usr/local/lib/python3.11/dist-packages (from spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (2.0.10)

Collecting typer<0.10.0,>=0.3.0 (from spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3)

Downloading typer-0.9.4-py3-none-any.whl.metadata (14 kB)

Collecting pathy>=0.10.0 (from spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3)

Downloading pathy-0.11.0-py3-none-any.whl.metadata (16 kB)

Collecting smart-open<7.0.0,>=5.2.1 (from spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3)

Downloading smart_open-6.4.0-py3-none-any.whl.metadata (21 kB)

Requirement already satisfied: tqdm<5.0.0,>=4.38.0 in /usr/local/lib/python3.11/dist-packages (from spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (4.67.1)

Requirement already satisfied: numpy>=1.15.0 in /usr/local/lib/python3.11/dist-packages (from spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (1.26.4)

Requirement already satisfied: requests<3.0.0,>=2.13.0 in /usr/local/lib/python3.11/dist-packages (from spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (2.32.3)

Requirement already satisfied: pydantic!=1.8,!<1.8.1,<3.0.0,>=1.7.4 in /usr/local/lib/python3.11/dist-packages (from spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (2.10.6)

Requirement already satisfied: jinja2 in /usr/local/lib/python3.11/dist-packages (from spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (3.1.5)

Requirement already satisfied: setuptools in /usr/local/lib/python3.11/dist-packages (from spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (75.1.0)

Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.11/dist-packages (from spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (24.2)

Requirement already satisfied: langcodes<4.0.0,>=3.2.0 in /usr/local/lib/python3.11/dist-packages (from spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (3.5.0)

Requirement already satisfied: language-data>=1.2 in /usr/local/lib/python3.11/dist-packages (from langcodes<4.0.0,>=3.2.0->spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (1.3.0)

Collecting pathlib-abc==0.1.1 (from pathy>=0.10.0->spacy<3.7.0,>=3.6.1->en_core_sci_

md==0.5.3)

Downloading pathlib_abc-0.1.1-py3-none-any.whl.metadata (18 kB)

Requirement already satisfied: annotated-types>=0.6.0 in /usr/local/lib/python3.11/dist-packages (from pydantic!=1.8,!<1.8.1,<3.0.0,>=1.7.4->spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (0.7.0)

Requirement already satisfied: pydantic-core==2.27.2 in /usr/local/lib/python3.11/dist-packages (from pydantic!=1.8,!<1.8.1,<3.0.0,>=1.7.4->spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (2.27.2)

Requirement already satisfied: typing-extensions>=4.12.2 in /usr/local/lib/python3.11/dist-packages (from pydantic!=1.8,!<1.8.1,<3.0.0,>=1.7.4->spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (4.12.2)

Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests<3.0.0,>=2.13.0->spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (3.4.1)

Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-packages (from requests<3.0.0,>=2.13.0->spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (3.10)

Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-packages (from requests<3.0.0,>=2.13.0->spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (2.3.0)

Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-packages (from requests<3.0.0,>=2.13.0->spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (2025.1.31)

Requirement already satisfied: blis<0.8.0,>=0.7.8 in /usr/local/lib/python3.11/dist-packages (from thinc<8.2.0,>=8.1.8->spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (0.7.11)

Requirement already satisfied: confection<1.0.0,>=0.0.1 in /usr/local/lib/python3.11/dist-packages (from thinc<8.2.0,>=8.1.8->spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (0.1.5)

Requirement already satisfied: click<9.0.0,>=7.1.1 in /usr/local/lib/python3.11/dist-packages (from typer<0.10.0,>=0.3.0->spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (8.1.8)

Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.11/dist-packages (from jinja2->spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (3.0.2)

Requirement already satisfied: marisa-trie>=1.1.0 in /usr/local/lib/python3.11/dist-packages (from language-data>=1.2->langcodes<4.0.0,>=3.2.0->spacy<3.7.0,>=3.6.1->en_core_sci_md==0.5.3) (1.2.1)

Downloading spacy-3.6.1-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (6.6 MB)

6.6/6.6 MB 82.1 MB/s eta 0:00:00

Downloading pathy-0.11.0-py3-none-any.whl (47 kB)

47.3/47.3 kB 5.0 MB/s eta 0:00:00

Downloading pathlib_abc-0.1.1-py3-none-any.whl (23 kB)

Downloading smart_open-6.4.0-py3-none-any.whl (57 kB)

57.0/57.0 kB 6.7 MB/s eta 0:00:00

Downloading thinc-8.1.12-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (917 kB)

917.4/917.4 kB 62.2 MB/s eta 0:00:00

Downloading typer-0.9.4-py3-none-any.whl (45 kB)

46.0/46.0 kB 4.7 MB/s eta 0:00:00

Building wheels for collected packages: en_core_sci_md

Building wheel for en_core_sci_md (setup.py) ... done

Created wheel for en_core_sci_md: filename=en_core_sci_md-0.5.3-py3-none-any.whl size=119160705 sha256=8a0488472f0facef4ac6849abfe839ad1d9b8d94f2dcdaafbcd9b8e29773dab0

Stored in directory: /root/.cache/pip/wheels/3f/80/c9/fe3472be531b02d08ade3509af1a6f62e5ed22db84fefd

```

Successfully built en_core_sci_md
Installing collected packages: typer, smart-open, pathlib-abc, pathy, thinc, spacy,
en_core_sci_md
  Attempting uninstall: typer
    Found existing installation: typer 0.15.1
    Uninstalling typer-0.15.1:
      Successfully uninstalled typer-0.15.1
  Attempting uninstall: smart-open
    Found existing installation: smart-open 7.1.0
    Uninstalling smart-open-7.1.0:
      Successfully uninstalled smart-open-7.1.0
  Attempting uninstall: thinc
    Found existing installation: thinc 8.2.5
    Uninstalling thinc-8.2.5:
      Successfully uninstalled thinc-8.2.5
  Attempting uninstall: spacy
    Found existing installation: spacy 3.7.5
    Uninstalling spacy-3.7.5:
      Successfully uninstalled spacy-3.7.5
ERROR: pip's dependency resolver does not currently take into account all the packages
that are installed. This behaviour is the source of the following dependency conflicts.
scispacy 0.5.5 requires spacy<3.8.0,>=3.7.0, but you have spacy 3.6.1 which is incompatible.
en-core-web-sm 3.7.1 requires spacy<3.8.0,>=3.7.2, but you have spacy 3.6.1 which is incompatible.
Successfully installed en_core_sci_md-0.5.3 pathlib-abc-0.1.1 pathy-0.11.0 smart-open-6.4.0
spacy-3.6.1 thinc-8.1.12 typer-0.9.4

```

```

In [98]: import pandas as pd
import scispacy
import spacy

#Load Dischar
patients_df_SciSpaCy = pd.read_csv("/content/drive/MyDrive/Colab Notebooks/AIH/Pati

nlp_SciSpaCy = spacy.load('en_core_sci_md') # Load the specified NLP model

# Apply token extraction
patients_df_SciSpaCy["Processed_Text"] = patients_df_SciSpaCy["TEXT"].apply(lambda

#Copy Processed data to google drive
patients_df_SciSpaCy.to_csv(r'Patient_Summary_SciSpacy.csv', index = False)
!cp 'Patient_Summary_SciSpacy.csv' '/content/drive/MyDrive/Colab Notebooks/AIH/Pati

```

```

In [100... import pandas as pd
#Load Processed SciSpacy
patients_df_SciSpaCy = pd.read_csv("/content/drive/MyDrive/Colab Notebooks/AIH/Pati

```

```

In [101... from spacy import displacy
displacy.render(nlp_SciSpaCy(patients_df_SciSpaCy["Processed_Text"][0]), style="en

```


Admission **ENTITY** Date 2140 1 19 Discharge Date **ENTITY** 2140 1 21 Date of Birth 2117
 8 7 Sex F Service MEDICINE Allergies **ENTITY** Penicillins **ENTITY** Attending:[**First
ENTITY Name3 **ENTITY** LF 2297 Chief Complaint headache **ENTITY** Major Surgical
ENTITY or Invasive Procedure Hemodialysis **ENTITY** History of Present Illness Ms. Known
 lastname is a 22 year **ENTITY** old female **ENTITY** with SLE lupus nephritis **ENTITY**
 ESRD **ENTITY** on HD **ENTITY** malignant **ENTITY** HTN h/o TTP **ENTITY** and
 HOCM **ENTITY** who presents with HA **ENTITY** and hypertensive **ENTITY** urgency
ENTITY Awoke **ENTITY** this a.m. with 8/10 left sided frontal HA **ENTITY** was n't sure if
 it was d/t flare **ENTITY** of uveitis **ENTITY** that had started on Monday **ENTITY** or
 d/t HTN **ENTITY** Decided **ENTITY** to skip HD **ENTITY** and come to ED **ENTITY**
 for evaluation **ENTITY** No vision changes numbness weakness **ENTITY** change in gait
 chest pain **ENTITY** SOB **ENTITY** + Diarrhea **ENTITY** x 1 day **ENTITY** In ED
ENTITY patient **ENTITY** was 217/140 but elevated **ENTITY** to 254/152 > received
 labetalol **ENTITY** IV 30 mg x 1 and MSO4 **ENTITY** 4 mg and pressures **ENTITY**
 dropped to SBPs **ENTITY** 208 and HA **ENTITY** improved Repeat labetalol **ENTITY**
 with 50 mg x 1 and repeated dose **ENTITY** of morphine **ENTITY** dropped pressures
ENTITY to 193/134 > labetalol **ENTITY** gtt **ENTITY** started asa given and HA
ENTITY resolved Head CT **ENTITY** negative **ENTITY** for intracranial bleed **ENTITY**
 and CXR **ENTITY** unremarkable ROS cold **ENTITY** for past week **ENTITY** no fevers
ENTITY chills CP SOB N/V + **ENTITY** diarrhea **ENTITY** Upon arrival **ENTITY** to the
 floor patient 's BP **ENTITY** was 191/126 labetalol **ENTITY** gtt **ENTITY** was not
 started No sxs no HA **ENTITY** She states **ENTITY** that she is compliant **ENTITY** with
 all her meds **ENTITY** and her mother **ENTITY** cooks **ENTITY** with no salt **ENTITY**
 and she has been adherent **ENTITY** to diet **ENTITY** Past Medical History 1 Lupus 2134
 Diagnosed **ENTITY** after she began to have swollen fingers **ENTITY** a rash **ENTITY**
 and painful joints 2 **ENTITY** ESRD **ENTITY** secodary **ENTITY** to SLE **ENTITY** 2135
 Was initially on cytoxan 1 dose **ENTITY** every 3 months **ENTITY** for 2 years **ENTITY**

until began dialysis 3 times ENTITY a week ENTITY in 2137 T Th Sat Awaiting ENTITY
 living donor transplant from mother 3 HTN ENTITY 2137 Normal ENTITY BPs ENTITY
 run 180's/120 's Has had 1 hypertensive ENTITY crisis ENTITY that precipitated
 ENTITY seizures ENTITY in the past 4 Uveitis ENTITY secondary to SLE ENTITY 4
 15 5 HOCM ENTITY per Echo ENTITY in 2137 6 Vaginal bleeding ENTITY 2139 9 20
 7 Multiple episodes ENTITY of dialysis reactions ENTITY 8 Anemia ENTITY 9 Coag
 neg Staph bacteremia ENTITY and HD line infection ENTITY 6 15 10 H/O UE ENTITY
 clot ENTITY was on coumadin ENTITY but no longer Social History Lives ENTITY in
 Location ENTITY 669 with mother ENTITY and 16 year ENTITY old brother
 ENTITY Graduated Name2 NI School ENTITY and then got sick ENTITY so currently is
 not working ENTITY or attending ENTITY school Denies ENTITY any T/E/D. Family
 History ENTITY -No history ENTITY of SLE ENTITY -Grandfather has HTN ENTITY
 -Distant history ENTITY of DM ENTITY -No history of clotting disorders ENTITY -
 No other history ENTITY of other autoimmune diseases ENTITY Physical Exam
 ENTITY Vitals 98.0 173/51 86 15 100 RA ENTITY HEENT ENTITY L eye injected
 w/periorbital edema R eye reactive w/ EOMI anicteric sclera MMM ENTITY OP ENTITY
 clear Neck supple ENTITY no LAD no thyromegaly Cardiac RRR ENTITY NL S1 ENTITY
 and S2 + S4 III/VI systolic ejection murmur ENTITY LUSB ENTITY radiating to apex
 ENTITY and axilla ENTITY intensifies ENTITY w/ Valsalva no rub ENTITY Lungs
 CTAB ENTITY no wheezes rhonchi crackles ENTITY Abd soft NTND NABS ENTITY no
 HSM ENTITY no rebound ENTITY or guarding ENTITY GU ENTITY no CVAT
 ENTITY Ext warm 2 + DP ENTITY pulses ENTITY no C/C/E ENTITY L femoral
 dialysis catheter ENTITY Neuro AOx3 CN II-XII ENTITY intact strength/sensation
 ENTITY grossly intact ENTITY Pertinent Results ENTITY UA ENTITY mod bld 100
 protein ENTITY present on prior UAs ENTITY Radiology ENTITY CXR ENTITY
 No acute ENTITY CP ENTITY abnormality EKG ENTITY NSR ENTITY nml
 ENTITY axis nml intervals borderline LAE ENTITY LVH ENTITY J point elevation in

V2,V3 TWI I ENTITY aVL V5 V6 No change ENTITY when compared ENTITY to prior
 on 2139 11 26 CT ENTITY HEAD No intracranial hemorrhage ENTITY Brief Hospital
 Course ENTITY A/P ENTITY Patient ENTITY is a 22 year ENTITY old female
 ENTITY with SLE lupus nephritis ENTITY ESRD ENTITY on HD ENTITY who
 presents with hypertensive ENTITY urgency Hypertensive urgency Unclear precipitant
 ENTITY Possibly secondary to pain ENTITY from worsening ENTITY uveitis ENTITY
 Compliant with ENTITY meds Denies ENTITY illicit ENTITY and tox screen
 ENTITY negative ENTITY Patient ENTITY was started on labetalol ENTITY drip
 ENTITY in ED ENTITY with good BP ENTITY response ENTITY and was
 subsequently transitioned to PO ENTITY anti-hypertensives ENTITY in ICU ENTITY
 with maintenance ENTITY of stable ENTITY SBPs ENTITY in 150s-170s baseline
 ENTITY 170s-190s Per nephrologist 's ENTITY recommendations ENTITY home
 lisinopril ENTITY was increased ENTITY to 40 mg po bid ENTITY from 40 mg po qd
 ENTITY for better baseline ENTITY BP ENTITY control No clinical evidence ENTITY
 of end organ damage ENTITY UA ENTITY difficult to interpret in setting of CRF CE 's
 ENTITY x 1 negative ENTITY Headache ENTITY No evidence ENTITY by CT
 ENTITY for intracranial bleed ENTITY Headaches ENTITY were well controlled with
 morphine sulfate ENTITY and had resolved by time of discharge Uveitis ENTITY
 Followed by outpatient ENTITY optho specialist Optho not consulted per patient 's
 ENTITY request ESRD ENTITY Secondary ENTITY to lupus nephritis ENTITY On
 transplant list ENTITY Patient ENTITY received hemodialysis ENTITY in house
 ENTITY with 500 ml ultrafiltrate ENTITY without complications ENTITY At dry
 weight ENTITY of 45 kg per patient ENTITY Began Sevelamer ENTITY 800 TID
 ENTITY with meals ENTITY Given difficulty ENTITY in interpreting ENTITY renin
 ENTITY and aldosterone ENTITY levels ENTITY in acutely ill ENTITY patients
 ENTITY these were not drawn and will need to be drawn at outpatient ENTITY follow
 up ENTITY Medications ENTITY on Admission Lisinopril ENTITY 40 mg PO ENTITY

QD ENTITY Labetalol ENTITY 600 PO ENTITY TID ENTITY Valsartan ENTITY
 320 mg PO ENTITY QD ENTITY Clonidine ENTITY 0.3 mg transdermal QW
 ENTITY Prednisone ENTITY 40 mg PO ENTITY QD ENTITY Atropine 1 Hospital1
 ENTITY Prednisolone ENTITY Acetate 1 ENTITY Q1H ENTITY Moxifloxacin
 ENTITY eye drops ENTITY qid ENTITY Lorazepam 1 ENTITY mg PO ENTITY
 Q4 6H PRN Discharge Medications 1 Labetalol ENTITY 200 mg Tablet ENTITY Sig Three
 3 Tablet ENTITY PO ENTITY TID ENTITY 3 times a day ENTITY Tablet(s ENTITY
 2 Clonidine ENTITY 0.3 mg/24 hr Patch Weekly Sig ENTITY One 1 Patch Weekly
 ENTITY Transdermal QTHUR ENTITY every Thursday 3 Atropine 1 Drops ENTITY Sig
 One 1 Drop Ophthalmic Hospital1 2 times a day ENTITY 4 Lorazepam 1 ENTITY mg
 Tablet ENTITY Sig One 1 Tablet ENTITY PO Q4 6H ENTITY every 4 to 6 hours
 ENTITY as needed 5 Valsartan ENTITY 160 mg Tablet ENTITY Sig Two 2 Tablet
 ENTITY PO ENTITY DAILY Daily ENTITY 6 Prednisolone ENTITY Acetate 1 Drops
 Suspension Sig ENTITY One 1 Drop Ophthalmic ENTITY Q1H ENTITY every hour 7
 Lisinopril ENTITY 40 mg Tablet ENTITY Sig One 1 Tablet ENTITY PO ENTITY
 twice a day ENTITY Disp:*60 Tablet(s ENTITY Refills:*2 ENTITY 8 Sevelamer
 ENTITY 800 mg Tablet ENTITY Sig One 1 Tablet ENTITY PO ENTITY TID ENTITY
 3 times a day ENTITY Disp:*90 ENTITY Tablet(s ENTITY Refills:*2 ENTITY 9
 Prednisone ENTITY 20 mg Tablet ENTITY Sig Two 2 Tablet ENTITY PO ENTITY
 once a day ENTITY 10 Blood Pressure Kit Kit Sig One 1 Kit ENTITY Miscellaneous once a
 day ENTITY Disp:*1 Kit Refills:*0 Discharge ENTITY Disposition Home Discharge
 ENTITY Diagnosis ENTITY Hypertensive urgency Discharge Condition Good Discharge
 Instructions ENTITY Please take all of your blood pressure medications ENTITY as
 prescribed ENTITY You should adhere to a low-salt diet ENTITY as increased ENTITY
 levels ENTITY of sodium ENTITY can drive your blood pressure ENTITY up You are
 being discharged with ENTITY a prescription ENTITY for a home blood pressure
 monitor ENTITY which you can use to take daily measurements ENTITY You should call

your primary care physician ENTITY for Name Initial PRE ENTITY systolic ENTITY blood pressures ENTITY greater than 180 or if you experience headaches ENTITY nausea vomiting chest ENTITY pain shortness of breath ENTITY or any other concerning symptoms ENTITY Followup Instructions Please ENTITY resume hemodialysis ENTITY according to your regular schedule ENTITY You are scheduled to see Dr. First Name8 NamePattern2 Last Name NamePattern1 4883 in the Division of Nephrology ENTITY on Wednesday 2 3 at 9:30 AM Please call Telephone/Fax 1 435 if you need to reschedule ENTITY You are scheduled to follow-up ENTITY with your primary care physician ENTITY Last Name NamePattern4 First Name4 NamePattern1 ENTITY Last Name NamePattern1 ENTITY 2423 on Tuesday 1 26 at 3:30 PM ENTITY Please call Telephone/Fax 1 250 if you need to reschedule ENTITY You have been referred to see Dr. First Name4 NamePattern1 Last Name NamePattern1 ENTITY 2539 in the Division of Hematology ENTITY for further evaluation ENTITY of your anemia ENTITY This appointment ENTITY is scheduled for 2 9 at 3 p.m. His office is located on the Location ENTITY un of the Hospital Ward ENTITY Name 23 Building on the Hospital1 18 Hospital Ward ENTITY Name 516 Please call Dr.[**Name NI ENTITY 44536 administrative assistant Doctor First ENTITY Name 8982 at Telephone/Fax 1 32192 ENTITY if you need to confirm ENTITY or reschedule ENTITY

Word2Vec and t-SNE Visualization Using SciSpaCy-Processed Data

```
In [141... from gensim.models import Word2Vec
```

```
corpus = build_corpus(patients_df_SciSpaCy, nlp_SciSpaCy)
```

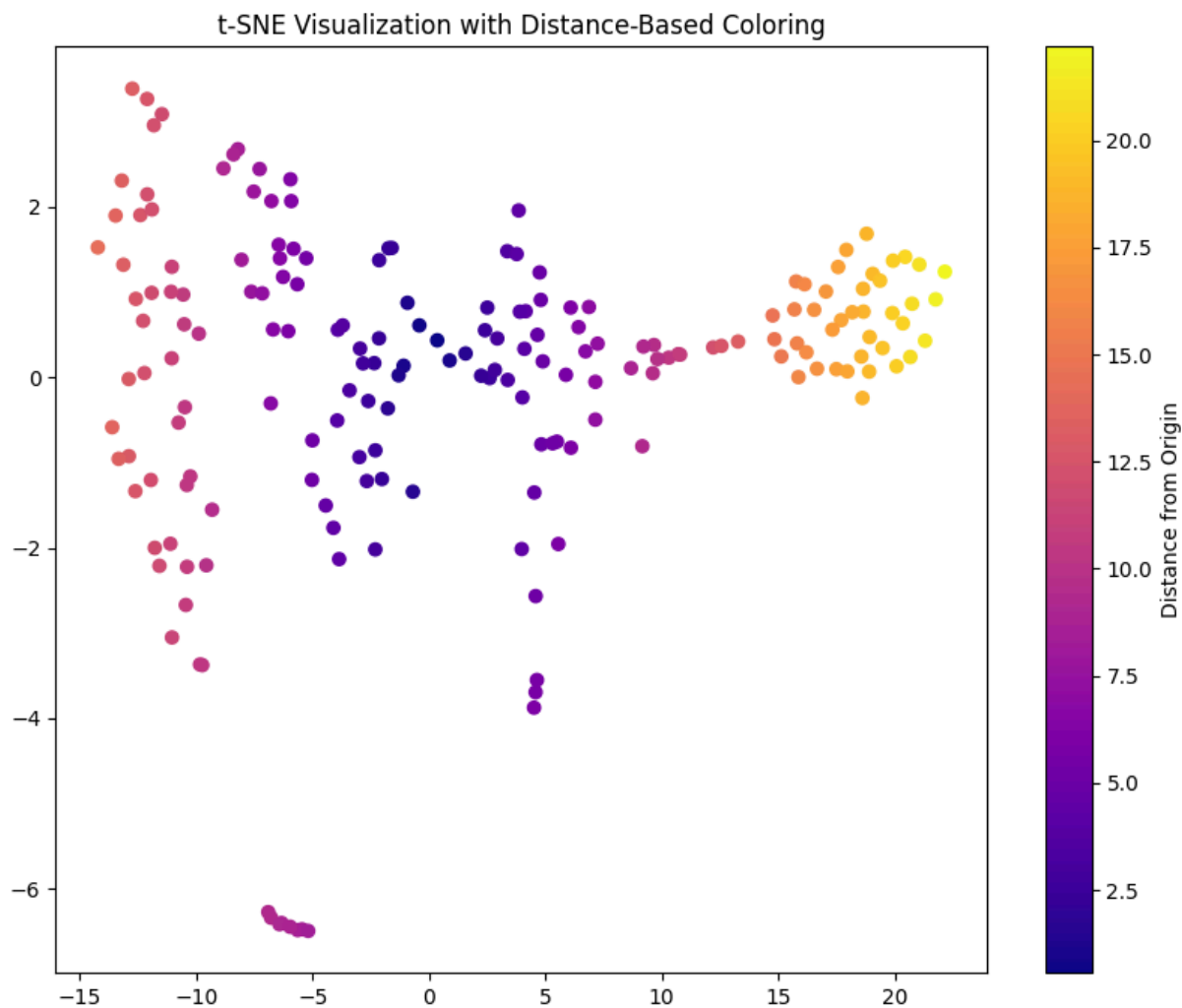
Minimum word count: 11, Maximum word count: 1152, Median word count: 474

```
In [153... model_word2vec = Word2Vec(corpus, min_count=50, window=3, vector_size=100)
```

```
In [154... model_word2vec.wv.similar_by_key("Clonidine"), model_word2vec.wv.similar_by_key("BP
```

```
In [155... vocabs = model_word2vec.wv.key_to_index.keys()
new_v = np.array(list(vocabs))
tsne plot(model_word2vec,new_v)
```





en_ner_bc5cdr_md Visualization Using SciSpaCy-Processed Data.

- It is a SciSpaCy model, not a standard SpaCy model. It is specifically trained for biomedical named entity recognition (NER), focusing on diseases and chemicals.

```
In [11]: !pip install https://s3-us-west-2.amazonaws.com/ai2-s2-scispace/releases/v0.4.0/en_
```

Collecting https://s3-us-west-2.amazonaws.com/ai2-s2-scispace/releases/v0.4.0/en_ner_bc5cdr_md-0.4.0.tar.gz

Downloading https://s3-us-west-2.amazonaws.com/ai2-s2-scispace/releases/v0.4.0/en_ner_bc5cdr_md-0.4.0.tar.gz (125.1 MB)

125.1/125.1 MB 12.9 MB/s eta 0:00:00

Preparing metadata (setup.py) ... done

Collecting spacy<3.1.0,>=3.0.1 (from en_ner_bc5cdr_md==0.4.0)

Downloading spacy-3.0.9.tar.gz (989 kB)

989.2/989.2 kB 19.0 MB/s eta 0:00:00

Installing build dependencies ... done

Getting requirements to build wheel ... done

Preparing metadata (pyproject.toml) ... done

Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.5 in /usr/local/lib/python3.11/dist-packages (from spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0) (3.0.12)

Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in /usr/local/lib/python3.11/dist-packages (from spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0) (1.0.12)

Requirement already satisfied: cymem<2.1.0,>=2.0.2 in /usr/local/lib/python3.11/dist-packages (from spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0) (2.0.11)

Requirement already satisfied: preshed<3.1.0,>=3.0.2 in /usr/local/lib/python3.11/dist-packages (from spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0) (3.0.9)

Collecting thinc<8.1.0,>=8.0.3 (from spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0)

Using cached thinc-8.0.17-cp311-cp311-linux_x86_64.whl

Requirement already satisfied: blis<0.8.0,>=0.4.0 in /usr/local/lib/python3.11/dist-packages (from spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0) (0.7.11)

Collecting wasabi<1.1.0,>=0.8.1 (from spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0)

Using cached wasabi-0.10.1-py3-none-any.whl.metadata (28 kB)

Requirement already satisfied: srsly<3.0.0,>=2.4.1 in /usr/local/lib/python3.11/dist-packages (from spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0) (2.5.1)

Requirement already satisfied: catalogue<2.1.0,>=2.0.4 in /usr/local/lib/python3.11/dist-packages (from spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0) (2.0.10)

Collecting typer<0.4.0,>=0.3.0 (from spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0)

Downloading typer-0.3.2-py3-none-any.whl.metadata (12 kB)

Requirement already satisfied: pathy>=0.3.5 in /usr/local/lib/python3.11/dist-packages (from spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0) (0.11.0)

Requirement already satisfied: smart-open<7.0.0,>=5.2.1 in /usr/local/lib/python3.11/dist-packages (from spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0) (6.4.0)

Requirement already satisfied: tqdm<5.0.0,>=4.38.0 in /usr/local/lib/python3.11/dist-packages (from spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0) (4.67.1)

Requirement already satisfied: numpy>=1.15.0 in /usr/local/lib/python3.11/dist-packages (from spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0) (1.26.4)

Requirement already satisfied: requests<3.0.0,>=2.13.0 in /usr/local/lib/python3.11/dist-packages (from spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0) (2.32.3)

Collecting pydantic!=1.8,!1.8.1,<1.9.0,>=1.7.4 (from spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0)

Using cached pydantic-1.8.2-py3-none-any.whl.metadata (103 kB)

Requirement already satisfied: jinja2 in /usr/local/lib/python3.11/dist-packages (from spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0) (3.1.5)

Requirement already satisfied: setuptools in /usr/local/lib/python3.11/dist-packages (from spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0) (75.1.0)

Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.11/dist-packages (from spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0) (24.2)

Requirement already satisfied: pathlib-abc==0.1.1 in /usr/local/lib/python3.11/dist-packages (from pathy>=0.3.5->spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0) (0.1.1)

Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.11/dist-packages (from pydantic!=1.8,!1.8.1,<1.9.0,>=1.7.4->spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0) (4.12.2)


```

Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests<3.0.0,>=2.13.0->spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-packages (from requests<3.0.0,>=2.13.0->spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-packages (from requests<3.0.0,>=2.13.0->spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0) (2.3.0)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-packages (from requests<3.0.0,>=2.13.0->spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0) (2025.1.31)
Collecting click<7.2.0,>=7.1.1 (from typer<0.4.0,>=0.3.0->spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0)
  Downloading click-7.1.2-py2.py3-none-any.whl.metadata (2.9 kB)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.11/dist-packages (from jinja2->spacy<3.1.0,>=3.0.1->en_ner_bc5cdr_md==0.4.0) (3.0.2)
Using cached pydantic-1.8.2-py3-none-any.whl (126 kB)
Downloading typer-0.3.2-py3-none-any.whl (21 kB)
Using cached wasabi-0.10.1-py3-none-any.whl (26 kB)
Downloading click-7.1.2-py2.py3-none-any.whl (82 kB)
      82.8/82.8 kB 10.1 MB/s eta 0:00:00
Building wheels for collected packages: en_ner_bc5cdr_md, spacy
  Building wheel for en_ner_bc5cdr_md (setup.py) ... done
  Created wheel for en_ner_bc5cdr_md: filename=en_ner_bc5cdr_md-0.4.0-py3-none-any.whl size=125666861 sha256=c9eff27bf4446a8e5874ebd29e7c3e2f680501cf9db90c3dc73ad6ee541d2f3d
  Stored in directory: /root/.cache/pip/wheels/96/37/fc/c6eee4239ccf271541084f54f737f15ed00f62f88406d22471
  Building wheel for spacy (pyproject.toml) ... done
  Created wheel for spacy: filename=spacy-3.0.9-cp311-cp311-linux_x86_64.whl size=27129738 sha256=3446819c9ec226a7fd4787bb64f0d5b89eae5a120e976d5018ab3bf4edcb6
  Stored in directory: /root/.cache/pip/wheels/15/7a/5d/414d7808f5065725156f1dbfc946cda0ba24cfbd3c4f2e3cb
Successfully built en_ner_bc5cdr_md spacy
Installing collected packages: wasabi, pydantic, click, typer, thinc, spacy, en_ner_bc5cdr_md
  Attempting uninstall: wasabi
    Found existing installation: wasabi 1.1.3
    Uninstalling wasabi-1.1.3:
      Successfully uninstalled wasabi-1.1.3
  Attempting uninstall: pydantic
    Found existing installation: pydantic 2.10.6
    Uninstalling pydantic-2.10.6:
      Successfully uninstalled pydantic-2.10.6
  Attempting uninstall: click
    Found existing installation: click 8.1.8
    Uninstalling click-8.1.8:
      Successfully uninstalled click-8.1.8
  Attempting uninstall: typer
    Found existing installation: typer 0.9.4
    Uninstalling typer-0.9.4:
      Successfully uninstalled typer-0.9.4
  Attempting uninstall: thinc
    Found existing installation: thinc 8.1.12
    Uninstalling thinc-8.1.12:

```

```
Successfully uninstalled thinc-8.1.12
Attempting uninstall: spacy
Found existing installation: spacy 3.6.1
Uninstalling spacy-3.6.1:
Successfully uninstalled spacy-3.6.1
```

ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behaviour is the source of the following dependency conflicts.

en-core-sci-md 0.5.3 requires spacy<3.7.0,>=3.6.1, but you have spacy 3.0.9 which is incompatible.

scispacy 0.5.5 requires spacy<3.8.0,>=3.7.0, but you have spacy 3.0.9 which is incompatible.

flask 3.1.0 requires click>=8.1.3, but you have click 7.1.2 which is incompatible.

google-genai 0.8.0 requires pydantic<3.0.0dev,>=2.0.0, but you have pydantic 1.8.2 which is incompatible.

langchain 0.3.17 requires pydantic<3.0.0,>=2.7.4, but you have pydantic 1.8.2 which is incompatible.

albumations 2.0.3 requires pydantic>=2.9.2, but you have pydantic 1.8.2 which is incompatible.

openai 1.61.1 requires pydantic<3,>=1.9.0, but you have pydantic 1.8.2 which is incompatible.

langchain-core 0.3.33 requires pydantic<3.0.0,>=2.5.2; python_full_version < "3.12.4", but you have pydantic 1.8.2 which is incompatible.

dask 2024.10.0 requires click>=8.1, but you have click 7.1.2 which is incompatible.

wandb 0.19.6 requires pydantic<3,>=2.6, but you have pydantic 1.8.2 which is incompatible.

en-core-web-sm 3.7.1 requires spacy<3.8.0,>=3.7.2, but you have spacy 3.0.9 which is incompatible.

Successfully installed click-7.1.2 en_ner_bc5cdr_md-0.4.0 pydantic-1.8.2 spacy-3.0.9 thinc-8.0.17 typer-0.3.2 wasabi-0.10.1

```
In [2]: !pip install -U spacy
        !pip install -U pydantic
        !pip install scispacy
```

Requirement already satisfied: spacy in /usr/local/lib/python3.11/dist-packages (3.0.9)
Collecting spacy
 Downloading spacy-3.8.4-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (27 kB)
Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.11 in /usr/local/lib/python3.11/dist-packages (from spacy) (3.0.12)
Requirement already satisfied: spacy-loggers<2.0.0,>=1.0.0 in /usr/local/lib/python3.11/dist-packages (from spacy) (1.0.5)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in /usr/local/lib/python3.11/dist-packages (from spacy) (1.0.12)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in /usr/local/lib/python3.11/dist-packages (from spacy) (2.0.11)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in /usr/local/lib/python3.11/dist-packages (from spacy) (3.0.9)
Collecting thinc<8.4.0,>=8.3.4 (from spacy)
 Downloading thinc-8.3.4-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (15 kB)
Requirement already satisfied: wasabi<1.2.0,>=0.9.1 in /usr/local/lib/python3.11/dist-packages (from spacy) (0.10.1)
Requirement already satisfied: srsly<3.0.0,>=2.4.3 in /usr/local/lib/python3.11/dist-packages (from spacy) (2.5.1)
Requirement already satisfied: catalogue<2.1.0,>=2.0.6 in /usr/local/lib/python3.11/dist-packages (from spacy) (2.0.10)
Requirement already satisfied: weasel<0.5.0,>=0.1.0 in /usr/local/lib/python3.11/dist-packages (from spacy) (0.4.1)
Requirement already satisfied: typer<1.0.0,>=0.3.0 in /usr/local/lib/python3.11/dist-packages (from spacy) (0.3.2)
Requirement already satisfied: tqdm<5.0.0,>=4.38.0 in /usr/local/lib/python3.11/dist-packages (from spacy) (4.67.1)
Requirement already satisfied: numpy>=1.19.0 in /usr/local/lib/python3.11/dist-packages (from spacy) (1.26.4)
Requirement already satisfied: requests<3.0.0,>=2.13.0 in /usr/local/lib/python3.11/dist-packages (from spacy) (2.32.3)
Requirement already satisfied: pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4 in /usr/local/lib/python3.11/dist-packages (from spacy) (1.8.2)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.11/dist-packages (from spacy) (3.1.5)
Requirement already satisfied: setuptools in /usr/local/lib/python3.11/dist-packages (from spacy) (75.1.0)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.11/dist-packages (from spacy) (24.2)
Requirement already satisfied: langcodes<4.0.0,>=3.2.0 in /usr/local/lib/python3.11/dist-packages (from spacy) (3.5.0)
Requirement already satisfied: language-data>=1.2 in /usr/local/lib/python3.11/dist-packages (from langcodes<4.0.0,>=3.2.0->spacy) (1.3.0)
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.11/dist-packages (from pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4->spacy) (4.12.2)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests<3.0.0,>=2.13.0->spacy) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-packages (from requests<3.0.0,>=2.13.0->spacy) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-packages (from requests<3.0.0,>=2.13.0->spacy) (2.3.0)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-packages (from requests<3.0.0,>=2.13.0->spacy) (2025.1.31)

```

Collecting blis<1.3.0,>=1.2.0 (from thinc<8.4.0,>=8.3.4->spacy)
  Downloading blis-1.2.0-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.
  metadata (7.7 kB)
Requirement already satisfied: confection<1.0.0,>=0.0.1 in /usr/local/lib/python3.11/dist-packages (from thinc<8.4.0,>=8.3.4->spacy) (0.1.5)
Requirement already satisfied: click<7.2.0,>=7.1.1 in /usr/local/lib/python3.11/dist-packages (from typer<1.0.0,>=0.3.0->spacy) (7.1.2)
Requirement already satisfied: cloudpathlib<1.0.0,>=0.7.0 in /usr/local/lib/python3.11/dist-packages (from weasel<0.5.0,>=0.1.0->spacy) (0.20.0)
Requirement already satisfied: smart-open<8.0.0,>=5.2.1 in /usr/local/lib/python3.11/dist-packages (from weasel<0.5.0,>=0.1.0->spacy) (6.4.0)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.11/dist-packages (from jinja2->spacy) (3.0.2)
Requirement already satisfied: marisa-trie>=1.1.0 in /usr/local/lib/python3.11/dist-packages (from language-data>=1.2->langcodes<4.0.0,>=3.2.0->spacy) (1.2.1)
Downloading spacy-3.8.4-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (30.6 MB)
_____ 30.6/30.6 MB 77.9 MB/s eta 0:00:00
Downloading thinc-8.3.4-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (3.9 MB)
_____ 3.9/3.9 MB 104.3 MB/s eta 0:00:00
Downloading blis-1.2.0-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (1.7 MB)
_____ 11.7/11.7 MB 96.8 MB/s eta 0:00:00
Installing collected packages: blis, thinc, spacy
  Attempting uninstall: blis
    Found existing installation: blis 0.7.11
    Uninstalling blis-0.7.11:
      Successfully uninstalled blis-0.7.11
  Attempting uninstall: thinc
    Found existing installation: thinc 8.0.17
    Uninstalling thinc-8.0.17:
      Successfully uninstalled thinc-8.0.17
  Attempting uninstall: spacy
    Found existing installation: spacy 3.0.9
    Uninstalling spacy-3.0.9:
      Successfully uninstalled spacy-3.0.9
ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behaviour is the source of the following dependency conflicts.
en-ner-bc5cdr-md 0.4.0 requires spacy<3.1.0,>=3.0.1, but you have spacy 3.8.4 which is incompatible.
en-core-sci-md 0.5.3 requires spacy<3.7.0,>=3.6.1, but you have spacy 3.8.4 which is incompatible.
scispacy 0.5.5 requires spacy<3.8.0,>=3.7.0, but you have spacy 3.8.4 which is incompatible.
en-core-web-sm 3.7.1 requires spacy<3.8.0,>=3.7.2, but you have spacy 3.8.4 which is incompatible.
Successfully installed blis-1.2.0 spacy-3.8.4 thinc-8.3.4

```

```

Requirement already satisfied: pydantic in /usr/local/lib/python3.11/dist-packages
(1.8.2)
Collecting pydantic
  Downloading pydantic-2.10.6-py3-none-any.whl.metadata (30 kB)
Requirement already satisfied: annotated-types>=0.6.0 in /usr/local/lib/python3.11/d
ist-packages (from pydantic) (0.7.0)
Requirement already satisfied: pydantic-core==2.27.2 in /usr/local/lib/python3.11/di
st-packages (from pydantic) (2.27.2)
Requirement already satisfied: typing-extensions>=4.12.2 in /usr/local/lib/python3.1
1/dist-packages (from pydantic) (4.12.2)
Downloading pydantic-2.10.6-py3-none-any.whl (431 kB)
----- 431.7/431.7 kB 14.6 MB/s eta 0:00:00
Installing collected packages: pydantic
  Attempting uninstall: pydantic
    Found existing installation: pydantic 1.8.2
    Uninstalling pydantic-1.8.2:
      Successfully uninstalled pydantic-1.8.2
ERROR: pip's dependency resolver does not currently take into account all the packag
es that are installed. This behaviour is the source of the following dependency conf
licts.
scispacy 0.5.5 requires spacy<3.8.0,>=3.7.0, but you have spacy 3.8.4 which is incom
patible.
en-core-web-sm 3.7.1 requires spacy<3.8.0,>=3.7.2, but you have spacy 3.8.4 which is
incompatible.
Successfully installed pydantic-2.10.6
Requirement already satisfied: scispacy in /usr/local/lib/python3.11/dist-packages
(0.5.5)
Collecting spacy<3.8.0,>=3.7.0 (from scispacy)
  Downloading spacy-3.7.5-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.wh
l.metadata (27 kB)
Requirement already satisfied: scipy in /usr/local/lib/python3.11/dist-packages (fro
m scispacy) (1.13.1)
Requirement already satisfied: requests<3.0.0,>=2.0.0 in /usr/local/lib/python3.11/d
ist-packages (from scispacy) (2.32.3)
Requirement already satisfied: conllu in /usr/local/lib/python3.11/dist-packages (fr
om scispacy) (6.0.0)
Requirement already satisfied: numpy in /usr/local/lib/python3.11/dist-packages (fro
m scispacy) (1.26.4)
Requirement already satisfied: joblib in /usr/local/lib/python3.11/dist-packages (fr
om scispacy) (1.4.2)
Requirement already satisfied: scikit-learn>=0.20.3 in /usr/local/lib/python3.11/dis
t-packages (from scispacy) (1.6.1)
Requirement already satisfied: pysbd in /usr/local/lib/python3.11/dist-packages (fro
m scispacy) (0.3.4)
Requirement already satisfied: nmslib-metabrainz==2.1.3 in /usr/local/lib/python3.1
1/dist-packages (from scispacy) (2.1.3)
Requirement already satisfied: pybind11>=2.2.3 in /usr/local/lib/python3.11/dist-pac
kages (from nmslib-metabrainz==2.1.3->scispacy) (2.13.6)
Requirement already satisfied: psutil in /usr/local/lib/python3.11/dist-packages (fr
om nmslib-metabrainz==2.1.3->scispacy) (5.9.5)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.1
1/dist-packages (from requests<3.0.0,>=2.0.0->scispacy) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-packag
es (from requests<3.0.0,>=2.0.0->scispacy) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-
packages (from requests<3.0.0,>=2.0.0->scispacy) (2.3.0)

```

```

Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-
packages (from requests<3.0.0,>=2.0.0->scispacy) (2025.1.31)
Requirement already satisfied: threadpoolctl>=3.1.0 in /usr/local/lib/python3.11/dis
t-packages (from scikit-learn>=0.20.3->scispacy) (3.5.0)
Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.11 in /usr/local/lib/python
3.11/dist-packages (from spacy<3.8.0,>=3.7.0->scispacy) (3.0.12)
Requirement already satisfied: spacy-loggers<2.0.0,>=1.0.0 in /usr/local/lib/python
3.11/dist-packages (from spacy<3.8.0,>=3.7.0->scispacy) (1.0.5)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in /usr/local/lib/python3.1
1/dist-packages (from spacy<3.8.0,>=3.7.0->scispacy) (1.0.12)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in /usr/local/lib/python3.11/dist
-packages (from spacy<3.8.0,>=3.7.0->scispacy) (2.0.11)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in /usr/local/lib/python3.11/di
st-packages (from spacy<3.8.0,>=3.7.0->scispacy) (3.0.9)
Traceback (most recent call last):
  File "/usr/local/lib/python3.11/dist-packages/pip/_internal/cli/base_command.py",
line 179, in exc_logging_wrapper
    status = run_func(*args)
             ^^^^^^^^^^^^^^^^^
  File "/usr/local/lib/python3.11/dist-packages/pip/_internal/cli/req_command.py", 1
ine 67, in wrapper
    return func(self, options, args)
           ^^^^^^^^^^^^^^^^^^^^^^^^^
  File "/usr/local/lib/python3.11/dist-packages/pip/_internal/commands/install.py",
line 377, in run
    requirement_set = resolver.resolve(
                     ^^^^^^^^^^^^^^^^^
  File "/usr/local/lib/python3.11/dist-packages/pip/_internal/resolution/resolvelib/
resolver.py", line 95, in resolve
    result = self._result = resolver.resolve(
             ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
  File "/usr/local/lib/python3.11/dist-packages/pip/_vendor/resolvelib/resolvers.p
y", line 546, in resolve
    state = resolution.resolve(requirements, max_rounds=max_rounds)
           ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
  File "/usr/local/lib/python3.11/dist-packages/pip/_vendor/resolvelib/resolvers.p
y", line 427, in resolve
    failure_causes = self._attempt_to_pin_criterion(name)
                    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
  File "/usr/local/lib/python3.11/dist-packages/pip/_vendor/resolvelib/resolvers.p
y", line 239, in _attempt_to_pin_criterion
    criteria = self._get_updated_criteria(candidate)
              ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
  File "/usr/local/lib/python3.11/dist-packages/pip/_vendor/resolvelib/resolvers.p
y", line 230, in _get_updated_criteria
    self._add_to_criteria(criteria, requirement, parent=candidate)
  File "/usr/local/lib/python3.11/dist-packages/pip/_vendor/resolvelib/resolvers.p
y", line 173, in _add_to_criteria
    if not criterion.candidates:
  File "/usr/local/lib/python3.11/dist-packages/pip/_vendor/resolvelib/structs.py",
line 156, in __bool__
    return bool(self._sequence)
           ^^^^^^^^^^^^^^^^^^^^^
  File "/usr/local/lib/python3.11/dist-packages/pip/_internal/resolution/resolvelib/
found_candidates.py", line 174, in __bool__
    return any(self)

```

```

File "/usr/local/lib/python3.11/dist-packages/pip/_internal/resolution/resolvelib/
found_candidates.py", line 162, in <genexpr>
    return (c for c in iterator if id(c) not in self._incompatible_ids)
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "/usr/local/lib/python3.11/dist-packages/pip/_internal/resolution/resolvelib/
found_candidates.py", line 49, in _iter_built
    for version, func in infos:
File "/usr/local/lib/python3.11/dist-packages/pip/_internal/resolution/resolvelib/
factory.py", line 300, in iter_index_candidate_infos
    result = self._finder.find_best_candidate(
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "/usr/local/lib/python3.11/dist-packages/pip/_internal/index/package_finder.p
y", line 884, in find_best_candidate
    candidates = self.find_all_candidates(project_name)
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "/usr/local/lib/python3.11/dist-packages/pip/_internal/index/package_finder.p
y", line 825, in find_all_candidates
    page_candidates = list(page_candidates_it)
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "/usr/local/lib/python3.11/dist-packages/pip/_internal/index/sources.py", lin
e 194, in page_candidates
    yield from self._candidates_from_page(self._link)
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "/usr/local/lib/python3.11/dist-packages/pip/_internal/index/package_finder.p
y", line 789, in process_project_url
    page_links = list(parse_links(index_response))
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "/usr/local/lib/python3.11/dist-packages/pip/_internal/index/collector.py", l
ine 218, in wrapper_wrapper
    return list(fn(page))
    ^^^^^^^^^^^^^^^^^
File "/usr/local/lib/python3.11/dist-packages/pip/_internal/index/collector.py", l
ine 233, in parse_links
    link = Link.from_json(file, page.url)
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "/usr/local/lib/python3.11/dist-packages/pip/_internal/models/link.py", line
273, in from_json
    url = _ensure_quoted_url(urllib.parse.urljoin(page_url, file_url))
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "/usr/local/lib/python3.11/dist-packages/pip/_internal/models/link.py", line
178, in _ensure_quoted_url
    return urllib.parse.urlunparse(result._replace(path=path))
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "/usr/lib/python3.11/urllib/parse.py", line 518, in urlunparse
    return _coerce_result(urlunsplit((scheme, netloc, url, query, fragment)))
    ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
File "/usr/lib/python3.11/urllib/parse.py", line 537, in urlunsplit
    return _coerce_result(url)
    ^^^^^^^^^^^^^^^^^^^^^
File "/usr/lib/python3.11/urllib/parse.py", line 108, in _noop
    def _noop(obj):

```

KeyboardInterrupt

During handling of the above exception, another exception occurred:

Traceback (most recent call last):

```

File "/usr/local/bin/pip3", line 10, in <module>
    sys.exit(main())
          ^^^^^^
File "/usr/local/lib/python3.11/dist-packages/pip/_internal/cli/main.py", line 80,
in main
    return command.main(cmd_args)
           ^^^^^^^^^^^^^^^^^^^^^^
File "/usr/local/lib/python3.11/dist-packages/pip/_internal/cli/base_command.py",
line 100, in main
    return self._main(args)
           ^^^^^^^^^^^^^^^^^^
File "/usr/local/lib/python3.11/dist-packages/pip/_internal/cli/base_command.py",
line 232, in _main
    return run(options, args)
           ^^^^^^^^^^^^^^^^^^
File "/usr/local/lib/python3.11/dist-packages/pip/_internal/cli/base_command.py",
line 215, in exc_logging_wrapper
    logger.critical("Operation cancelled by user")
File "/usr/lib/python3.11/logging/__init__.py", line 1526, in critical
    def critical(self, msg, *args, **kwargs):

```

KeyboardInterrupt

^C

In [156...

```

import en_ner_bc5cdr_md
import spacy
from spacy import displacy
nlp_bc5cdr = en_ner_bc5cdr_md.load()
displacy.render(nlp_bc5cdr(patients_df_SciSpaCy['Processed_Text'][0]), style="ent"

```


Admission Date 2140 1 19 Discharge Date 2140 1 21 Date of Birth 2117 8 **DISEASE** 7 Sex F
 Service MEDICINE Allergies **DISEASE** Penicillins **CHEMICAL** Attending:[**First Name3
CHEMICAL LF 2297 Chief Complaint headache **DISEASE** Major Surgical or Invasive
 Procedure Hemodialysis History of Present Illness Ms. Known lastname is a 22 year old female
 with SLE lupus nephritis ESRD **DISEASE** on HD malignant HTN **DISEASE** h/o TTP
DISEASE and HOCM **DISEASE** who presents with HA and hypertensive **DISEASE**
 urgency Awoke this a.m. with 8/10 left sided frontal HA was n't sure if it was d/t flare of
 uveitis **DISEASE** that had started on Monday or d/t HTN Decided to skip HD **DISEASE**
 and come to ED for evaluation No vision changes numbness weakness **DISEASE** change in
 gait chest pain **DISEASE** SOB + Diarrhea **DISEASE** x 1 day In ED patient was 217/140 but
 elevated to 254/152 > received labetalol **CHEMICAL** IV 30 mg x 1 and MSO4 **CHEMICAL**
 4 mg and pressures dropped to SBPs 208 and HA improved Repeat labetalol **CHEMICAL**
 with 50 mg x 1 and repeated dose of morphine **CHEMICAL** dropped pressures to 193/134 >
 labetalol **CHEMICAL** gtt started asa given and HA resolved Head CT negative for
 intracranial bleed **DISEASE** and CXR unremarkable ROS cold for past week no fevers chills
 CP SOB N/V **CHEMICAL** + diarrhea **DISEASE** Upon arrival to the floor patient 's BP was
 191/126 labetalol **CHEMICAL** gtt was not started No sxs no HA She states that she is
 compliant with all her meds and her mother cooks with no salt **CHEMICAL** and she has been
 adherent to diet Past Medical History 1 Lupus 2134 Diagnosed after she began to have swollen
 fingers a rash **DISEASE** and painful joints **DISEASE** 2 ESRD **DISEASE** secondary to SLE
 2135 Was initially on cytoxan **CHEMICAL** 1 dose every 3 months for 2 years until began
 dialysis 3 times a week in 2137 T Th Sat Awaiting living donor transplant from mother 3 HTN
 2137 Normal BPs run 180's/120 's Has had 1 hypertensive **DISEASE** crisis that precipitated
 seizures **DISEASE** in the past 4 Uveitis **DISEASE** secondary to SLE 4 15 5 HOCM
DISEASE per Echo in 2137 6 Vaginal bleeding **DISEASE** 2139 9 20 7 Multiple episodes of
 dialysis reactions 8 Anemia **DISEASE** 9 Coag neg Staph bacteremia **DISEASE** and HD
DISEASE line infection **DISEASE** 6 15 10 H/O UE clot was on coumadin **CHEMICAL** but

no longer Social History Lives in Location 669 with mother and 16 year old brother Graduated Name2 NI School and then got sick so currently is not working or attending school Denies any T/E/D. Family History -No history of SLE -Grandfather has HTN -Distant history of DM

CHEMICAL -No history of clotting disorders **DISEASE** -No other history of other autoimmune diseases **DISEASE** Physical Exam Vitals 98.0 173/51 86 15 100 RA HEENT L eye injected w/periorbital edema R eye reactive w/ **DISEASE** EOMI anicteric sclera MMM OP clear Neck **CHEMICAL** supple no LAD no thyromegaly **DISEASE** Cardiac RRR NL S1 and S2 + S4 III/VI systolic ejection murmur LUSB radiating to apex and axilla intensifies w/ Valsalva no rub Lungs CTAB **CHEMICAL** no wheezes rhonchi **DISEASE** crackles Abd soft NTND NABS no HSM no rebound or guarding GU **CHEMICAL** no CVAT Ext warm 2 + DP pulses no C/C/E L femoral dialysis catheter Neuro AOx3 CN II-XII intact strength/sensation grossly intact

Pertinent Results UA **CHEMICAL** mod bld 100 protein present on prior UAs **CHEMICAL** Radiology CXR No acute CP abnormality **CHEMICAL** EKG NSR **DISEASE** nml axis nml intervals borderline LAE LVH J point elevation in V2,V3 TWI I aVL V5 V6 No change when compared to prior on 2139 11 26 CT HEAD No intracranial hemorrhage **DISEASE** Brief Hospital Course A/P Patient is a 22 year old female with SLE lupus nephritis ESRD **DISEASE** on HD **DISEASE** who presents with hypertensive **DISEASE** urgency Hypertensive **DISEASE** urgency Unclear precipitant Possibly secondary to pain **DISEASE** from worsening uveitis **DISEASE** Compliant with meds Denies illicit and tox screen negative Patient was started on labetalol **CHEMICAL** drip in ED with good BP response and was subsequently transitioned to PO anti-hypertensives in ICU with maintenance of stable SBPs in 150s-170s baseline 170s-190s Per nephrologist 's recommendations home lisinopril **CHEMICAL** was increased to 40 mg po bid from 40 mg po qd for better baseline BP control No clinical evidence of end organ damage UA **DISEASE** difficult to interpret in setting of CRF **DISEASE** CE 's x 1 negative Headache **DISEASE** No evidence by CT for intracranial bleed **DISEASE** Headaches **DISEASE** were well controlled with morphine **CHEMICAL** sulfate and had resolved by time of discharge Uveitis **DISEASE** Followed by outpatient ophtho specialist

Ophtho **CHEMICAL** not consulted per patient 's request ESRD **DISEASE** Secondary to
 lupus nephritis **DISEASE** On transplant list Patient received hemodialysis in house with 500
 ml ultrafiltrate without complications At dry weight of 45 kg per patient Began Sevelamer 800
 TID with meals Given difficulty in interpreting renin and aldosterone **CHEMICAL** levels in
 acutely ill patients these were not drawn and will need to be drawn at outpatient follow up
 Medications on Admission Lisinopril **CHEMICAL** 40 mg PO QD Labetalol **CHEMICAL** 600
 PO TID Valsartan **CHEMICAL** 320 mg PO QD Clonidine **CHEMICAL** 0.3 mg transdermal
 QW Prednisone **CHEMICAL** 40 mg PO QD Atropine **CHEMICAL** 1 Hospital1
 Prednisolone Acetate **CHEMICAL** 1 Q1H Moxifloxacin **CHEMICAL** eye drops qid
 Lorazepam **CHEMICAL** 1 mg PO Q4 6H PRN Discharge Medications 1 Labetalol **CHEMICAL**
 200 mg Tablet Sig Three 3 Tablet PO TID 3 times a day Tablet(s **CHEMICAL** 2 Clonidine
CHEMICAL 0.3 mg/24 hr Patch Weekly Sig One 1 Patch Weekly Transdermal QTHUR every
 Thursday 3 Atropine **CHEMICAL** 1 Drops **CHEMICAL** Sig One 1 Drop Ophthalmic
 Hospital1 2 times a day 4 Lorazepam **CHEMICAL** 1 mg Tablet Sig One 1 Tablet PO Q4 6H
 every 4 to 6 hours as needed 5 Valsartan **CHEMICAL** 160 mg Tablet Sig Two 2 Tablet PO
 DAILY Daily 6 Prednisolone Acetate **CHEMICAL** 1 Drops **CHEMICAL** Suspension Sig One
 1 Drop Ophthalmic Q1H every hour 7 Lisinopril **CHEMICAL** 40 mg Tablet Sig One 1 Tablet
 PO twice a day Disp:*60 Tablet(s Refills:*2 **CHEMICAL** 8 Sevelamer **CHEMICAL** 800 mg
 Tablet Sig One 1 Tablet PO TID 3 times a day Disp:*90 Tablet(s Refills:*2 **CHEMICAL** 9
 Prednisone **CHEMICAL** 20 mg Tablet Sig Two 2 Tablet PO once a day 10 Blood Pressure Kit
 Kit Sig One 1 Kit Miscellaneous once a day Disp:*1 **CHEMICAL** Kit Refills:*0 Discharge
 Disposition Home Discharge Diagnosis Hypertensive **DISEASE** urgency Discharge Condition
 Good Discharge Instructions Please take all of your blood pressure medications as prescribed
 You should adhere to a low-salt diet as increased levels of sodium **CHEMICAL** can drive your
 blood pressure up You are being discharged with a prescription for a home blood pressure
 monitor which you can use to take daily measurements You should call your primary care
 physician for Name Initial PRE systolic blood pressures greater than 180 or if you experience

headaches nausea vomiting chest pain shortness of breath **DISEASE** or any other concerning symptoms Followup Instructions Please resume hemodialysis according to your regular schedule You are scheduled to see Dr. First Name8 NamePattern2 Last Name NamePattern1 **CHEMICAL** 4883 in the Division of Nephrology on Wednesday 2 3 at 9:30 AM Please call Telephone/Fax 1 435 if you need to reschedule You are scheduled to follow-up with your primary care physician Last Name NamePattern4 First Name4 NamePattern1 **CHEMICAL** Last Name NamePattern1 **CHEMICAL** 2423 on Tuesday 1 26 at 3:30 PM Please call Telephone/Fax 1 250 if you need to reschedule You have been referred to see Dr. First Name4 NamePattern1 **CHEMICAL** Last Name NamePattern1 **CHEMICAL** 2539 in the Division of Hematology for further evaluation of your anemia **DISEASE** This appointment is scheduled for 2 9 at 3 p.m. His office is located on the Location un of the Hospital Ward Name 23 Building on the Hospital1 18 Hospital Ward Name 516 Please call Dr.[**Name NI 44536 administrative assistant Doctor First Name 8982 at Telephone/Fax 1 32192 if you need to confirm or reschedule

MedSpacy Visualization Using SciSpaCy-Processed Data.

```
In [8]: !pip install medspacy
```

Collecting medspacy

Downloading medspacy-1.3.1.tar.gz (244 kB)

```

_____ 0.0/244.6 kB ? eta -:-:-
_____ 244.6/244.6 kB 8.4 MB/s eta 0:00:00

```

Installing build dependencies ... done

Getting requirements to build wheel ... done

Preparing metadata (pyproject.toml) ... done

Collecting spacy<3.8 (from medspacy)

Using cached spacy-3.7.5-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (27 kB)

Collecting PyRuSH>=1.0.8 (from medspacy)

Downloading PyRuSH-1.0.9-cp311-cp311-manylinux_2_5_x86_64.manylinux1_x86_64.manylinux2_17_x86_64.manylinux2014_x86_64.whl.metadata (5.2 kB)

Requirement already satisfied: pysbd==0.3.4 in /usr/local/lib/python3.11/dist-packages (from medspacy) (0.3.4)

Requirement already satisfied: jsonschema in /usr/local/lib/python3.11/dist-packages (from medspacy) (4.23.0)

Collecting medspacy_quickumls==3.2 (from medspacy)

Downloading medspacy_quickumls-3.2.tar.gz (69 kB)

```

_____ 69.9/69.9 kB 7.2 MB/s eta 0:00:00

```

Preparing metadata (setup.py) ... done

Requirement already satisfied: numpy>=1.8.2 in /usr/local/lib/python3.11/dist-packages (from medspacy_quickumls==3.2->medspacy) (1.26.4)

Collecting unicode>=0.4.19 (from medspacy_quickumls==3.2->medspacy)

Downloading Unicode-1.3.8-py3-none-any.whl.metadata (13 kB)

Requirement already satisfied: nltk>=3.3 in /usr/local/lib/python3.11/dist-packages (from medspacy_quickumls==3.2->medspacy) (3.9.1)

Collecting pysimstring (from medspacy_quickumls==3.2->medspacy)

Downloading pysimstring-1.2.1.tar.gz (68 kB)

```

_____ 68.4/68.4 kB 7.8 MB/s eta 0:00:00

```

Preparing metadata (setup.py) ... done

Collecting medspacy_unqlite>=0.8.1 (from medspacy_quickumls==3.2->medspacy)

Downloading medspacy_unqlite-0.9.6-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (1.4 kB)

Requirement already satisfied: pytest>=6 in /usr/local/lib/python3.11/dist-packages (from medspacy_quickumls==3.2->medspacy) (8.3.4)

Requirement already satisfied: six in /usr/local/lib/python3.11/dist-packages (from medspacy_quickumls==3.2->medspacy) (1.17.0)

Requirement already satisfied: Cython in /usr/local/lib/python3.11/dist-packages (from PyRuSH>=1.0.8->medspacy) (3.0.11)

Requirement already satisfied: setuptools in /usr/local/lib/python3.11/dist-packages (from PyRuSH>=1.0.8->medspacy) (75.1.0)

Collecting PyFastNER>=1.0.8 (from PyRuSH>=1.0.8->medspacy)

Downloading PyFastNER-1.0.10-py3-none-any.whl.metadata (15 kB)

Collecting quicksectx>=0.3.5 (from PyRuSH>=1.0.8->medspacy)

Downloading quicksectx-0.4.0-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (4.6 kB)

Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.11 in /usr/local/lib/python3.11/dist-packages (from spacy<3.8->medspacy) (3.0.12)

Requirement already satisfied: spacy-loggers<2.0.0,>=1.0.0 in /usr/local/lib/python3.11/dist-packages (from spacy<3.8->medspacy) (1.0.5)

Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in /usr/local/lib/python3.11/dist-packages (from spacy<3.8->medspacy) (1.0.12)

Requirement already satisfied: cymem<2.1.0,>=2.0.2 in /usr/local/lib/python3.11/dist-packages (from spacy<3.8->medspacy) (2.0.11)

Requirement already satisfied: preshed<3.1.0,>=3.0.2 in /usr/local/lib/python3.11/di

```

st-packages (from spacy<3.8->medspacy) (3.0.9)
Collecting thinc<8.3.0,>=8.2.2 (from spacy<3.8->medspacy)
  Downloading thinc-8.2.5-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl
1.metadata (15 kB)
Requirement already satisfied: wasabi<1.2.0,>=0.9.1 in /usr/local/lib/python3.11/dist-
t-packages (from spacy<3.8->medspacy) (0.10.1)
Requirement already satisfied: srsly<3.0.0,>=2.4.3 in /usr/local/lib/python3.11/dist
-packages (from spacy<3.8->medspacy) (2.5.1)
Requirement already satisfied: catalogue<2.1.0,>=2.0.6 in /usr/local/lib/python3.11/
dist-packages (from spacy<3.8->medspacy) (2.0.10)
Requirement already satisfied: weasel<0.5.0,>=0.1.0 in /usr/local/lib/python3.11/dis
t-packages (from spacy<3.8->medspacy) (0.4.1)
Requirement already satisfied: typer<1.0.0,>=0.3.0 in /usr/local/lib/python3.11/dist
-packages (from spacy<3.8->medspacy) (0.3.2)
Requirement already satisfied: tqdm<5.0.0,>=4.38.0 in /usr/local/lib/python3.11/dist
-packages (from spacy<3.8->medspacy) (4.67.1)
Requirement already satisfied: requests<3.0.0,>=2.13.0 in /usr/local/lib/python3.11/
dist-packages (from spacy<3.8->medspacy) (2.32.3)
Requirement already satisfied: pydantic!=1.8,!<1.8.1,<3.0.0,>=1.7.4 in /usr/local/li
b/python3.11/dist-packages (from spacy<3.8->medspacy) (2.10.6)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.11/dist-packages (fr
om spacy<3.8->medspacy) (3.1.5)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.11/dist-pac
kages (from spacy<3.8->medspacy) (24.2)
Requirement already satisfied: langcodes<4.0.0,>=3.2.0 in /usr/local/lib/python3.11/
dist-packages (from spacy<3.8->medspacy) (3.5.0)
Requirement already satisfied: attrs>=22.2.0 in /usr/local/lib/python3.11/dist-packa
ges (from jsonschema->medspacy) (25.1.0)
Requirement already satisfied: jsonschema-specifications>=2023.03.6 in /usr/local/li
b/python3.11/dist-packages (from jsonschema->medspacy) (2024.10.1)
Requirement already satisfied: referencing>=0.28.4 in /usr/local/lib/python3.11/dist
-packages (from jsonschema->medspacy) (0.36.2)
Requirement already satisfied: rpds-py>=0.7.1 in /usr/local/lib/python3.11/dist-pack
ages (from jsonschema->medspacy) (0.22.3)
Requirement already satisfied: language-data>=1.2 in /usr/local/lib/python3.11/dist-
packages (from langcodes<4.0.0,>=3.2.0->spacy<3.8->medspacy) (1.3.0)
Requirement already satisfied: click in /usr/local/lib/python3.11/dist-packages (fro
m nltk>=3.3->medspacy_quickumls==3.2->medspacy) (7.1.2)
Requirement already satisfied: joblib in /usr/local/lib/python3.11/dist-packages (fr
om nltk>=3.3->medspacy_quickumls==3.2->medspacy) (1.4.2)
Requirement already satisfied: regex>=2021.8.3 in /usr/local/lib/python3.11/dist-pac
kages (from nltk>=3.3->medspacy_quickumls==3.2->medspacy) (2024.11.6)
Requirement already satisfied: annotated-types>=0.6.0 in /usr/local/lib/python3.11/d
ist-packages (from pydantic!=1.8,!<1.8.1,<3.0.0,>=1.7.4->spacy<3.8->medspacy) (0.7.
0)
Requirement already satisfied: pydantic-core==2.27.2 in /usr/local/lib/python3.11/di
st-packages (from pydantic!=1.8,!<1.8.1,<3.0.0,>=1.7.4->spacy<3.8->medspacy) (2.27.
2)
Requirement already satisfied: typing-extensions>=4.12.2 in /usr/local/lib/python3.1
1/dist-packages (from pydantic!=1.8,!<1.8.1,<3.0.0,>=1.7.4->spacy<3.8->medspacy) (4.
12.2)
Requirement already satisfied: iniconfig in /usr/local/lib/python3.11/dist-packages
(from pytest>=6->medspacy_quickumls==3.2->medspacy) (2.0.0)
Requirement already satisfied: pluggy<2,>=1.5 in /usr/local/lib/python3.11/dist-pack
ages (from pytest>=6->medspacy_quickumls==3.2->medspacy) (1.5.0)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.1

```

```

1/dist-packages (from requests<3.0.0,>=2.13.0->spacy<3.8->medspacy) (3.4.1)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-packag
es (from requests<3.0.0,>=2.13.0->spacy<3.8->medspacy) (3.10)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-
packages (from requests<3.0.0,>=2.13.0->spacy<3.8->medspacy) (2.3.0)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-
packages (from requests<3.0.0,>=2.13.0->spacy<3.8->medspacy) (2025.1.31)
Collecting blis<0.8.0,>=0.7.8 (from thinc<8.3.0,>=8.2.2->spacy<3.8->medspacy)
  Using cached blis-0.7.11-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.wh
l.metadata (7.4 kB)
Requirement already satisfied: confection<1.0.0,>=0.0.1 in /usr/local/lib/python3.1
1/dist-packages (from thinc<8.3.0,>=8.2.2->spacy<3.8->medspacy) (0.1.5)
Requirement already satisfied: cloudpathlib<1.0.0,>=0.7.0 in /usr/local/lib/python3.
11/dist-packages (from weasel<0.5.0,>=0.1.0->spacy<3.8->medspacy) (0.20.0)
Requirement already satisfied: smart-open<8.0.0,>=5.2.1 in /usr/local/lib/python3.1
1/dist-packages (from weasel<0.5.0,>=0.1.0->spacy<3.8->medspacy) (6.4.0)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.11/dist-pac
kages (from jinja2->spacy<3.8->medspacy) (3.0.2)
Requirement already satisfied: marisa-trie>=1.1.0 in /usr/local/lib/python3.11/dist-
packages (from language-data>=1.2->langcodes<4.0.0,>=3.2.0->spacy<3.8->medspacy) (1.
2.1)
Downloading PyRuSH-1.0.9-cp311-cp311-manylinux_2_5_x86_64.manylinux1_x86_64.manylinu
x_2_17_x86_64.manylinux2014_x86_64.whl (211 kB)
_____ 211.8/211.8 kB 19.0 MB/s eta 0:00:00
Downloading spacy-3.7.5-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl
(6.6 MB)
_____ 6.6/6.6 MB 109.9 MB/s eta 0:00:00
Downloading medspacy_unqlite-0.9.6-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x
86_64.whl (419 kB)
_____ 419.7/419.7 kB 37.4 MB/s eta 0:00:00
Downloading PyFastNER-1.0.10-py3-none-any.whl (22 kB)
Downloading quicksectx-0.4.0-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.
whl (1.1 MB)
_____ 1.1/1.1 MB 70.7 MB/s eta 0:00:00
Downloading thinc-8.2.5-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl
(920 kB)
_____ 920.2/920.2 kB 61.0 MB/s eta 0:00:00
Downloading Unidecode-1.3.8-py3-none-any.whl (235 kB)
_____ 235.5/235.5 kB 24.2 MB/s eta 0:00:00
Using cached blis-0.7.11-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl
(10.2 MB)
Building wheels for collected packages: medspacy, medspacy_quickumls, pysimstring
  Building wheel for medspacy (pyproject.toml) ... done
  Created wheel for medspacy: filename=medspacy-1.3.1-py3-none-any.whl size=313656 s
ha256=30719daf9fbfa0f8d3bc1eb07fdd99438a178850ae20edd4c3bc0f82a04ece10
  Stored in directory: /root/.cache/pip/wheels/86/7a/11/d0a022246fbc360e3565129da2fe
ab9721c8cfc77a28071cd4
  Building wheel for medspacy_quickumls (setup.py) ... done
  Created wheel for medspacy_quickumls: filename=medspacy_quickumls-3.2-py3-none-an
y.whl size=98385 sha256=4fdae0b50cf9946fa21c29732ea7fd3362e91fa6feface2a579ea710cd08
25e1
  Stored in directory: /root/.cache/pip/wheels/ec/7d/89/b535af193474457500ee5c185017
c502083c44ae44d0a39d21
  Building wheel for pysimstring (setup.py) ... done
  Created wheel for pysimstring: filename=pysimstring-1.2.1-cp311-cp311-linux_x86_6
4.whl size=1326530 sha256=dee89c552917341b451ea8742c31c3fe8747e4499f37c4ea6148719a50

```

e65ba9

Stored in directory: /root/.cache/pip/wheels/bf/f2/a6/333de60c73d3a47ad895b02532f0a7bfeb723d00474775b558

Successfully built medspacy medspacy_quickumls pysimstring

Installing collected packages: pysimstring, medspacy_unqlite, unicode, quicksectx, blis, PyFastNER, thinc, spacy, PyRuSH, medspacy_quickumls, medspacy

Attempting uninstall: blis

Found existing installation: blis 1.2.0

Uninstalling blis-1.2.0:

Successfully uninstalled blis-1.2.0

Attempting uninstall: thinc

Found existing installation: thinc 8.3.4

Uninstalling thinc-8.3.4:

Successfully uninstalled thinc-8.3.4

Attempting uninstall: spacy

Found existing installation: spacy 3.8.4

Uninstalling spacy-3.8.4:

Successfully uninstalled spacy-3.8.4

ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behaviour is the source of the following dependency conflicts.

en-ner-bc5cdr-md 0.4.0 requires spacy<3.1.0,>=3.0.1, but you have spacy 3.7.5 which is incompatible.

en-core-sci-md 0.5.3 requires spacy<3.7.0,>=3.6.1, but you have spacy 3.7.5 which is incompatible.

Successfully installed PyFastNER-1.0.10 PyRuSH-1.0.9 blis-0.7.11 medspacy-1.3.1 medspacy_quickumls-3.2 medspacy_unqlite-0.9.6 pysimstring-1.2.1 quicksectx-0.4.0 spacy-3.7.5 thinc-8.2.5 unicode-1.3.8

In [161...

```
import medspacy
from spacy import displacy
from medspacy.ner import TargetRule
from medspacy.visualization import visualize_ent

# Load MedspaCy NLP pipeline
nlp_medspacy = medspacy.load()

# Add rules for target concept extraction
target_matcher = nlp_medspacy.get_pipe("medspacy_target_matcher")
# Define custom rules for better entity detection
target_rules = [
    TargetRule("hyperlipidemia", "DISEASE"),
    TargetRule("O2", "CHEMICAL"),
    TargetRule("FiO2", "CHEMICAL"),
    TargetRule("hypertension", "DISEASE"),
    TargetRule("hypertensive urgency", "DISEASE"),
    TargetRule("obesity", "CONDITION"),
    TargetRule("cardiac", "DISEASE"),
    TargetRule("SLE", "DISEASE"),
    TargetRule("lupus nephritis", "DISEASE"),
    TargetRule("ESRD", "DISEASE"),
    TargetRule("dialysis", "TREATMENT"),
    TargetRule("hemodialysis", "TREATMENT"),
    TargetRule("SBP", "MEASUREMENT"),
    TargetRule("HR", "MEASUREMENT"),
    TargetRule("TPN", "TREATMENT"),
```



```
TargetRule("Prednisone", "MEDICATION"),
TargetRule("Lisinopril", "MEDICATION"),
TargetRule("Labetalol", "MEDICATION"),
TargetRule("Clonidine", "MEDICATION"),
TargetRule("Valsartan", "MEDICATION"),
TargetRule("Sevelamer", "MEDICATION"),
TargetRule("Atropine", "MEDICATION"),
TargetRule("Morphine sulfate", "MEDICATION"),
TargetRule("Diarrhea", "SYMPTOM"),
TargetRule("Headache", "SYMPTOM"),
TargetRule("nausea", "SYMPTOM"),
TargetRule("vomiting", "SYMPTOM"),
TargetRule("shortness of breath", "SYMPTOM"),
TargetRule("fever", "SYMPTOM"),
TargetRule("chills", "SYMPTOM")
]

target_matcher.add(target_rules)

# Process the shift note
doc = nlp_medspacy(patients_df_SciSpaCy['Processed_Text'][0])

# visualize
visualize_ent(doc)
```

Admission Date 2140 1 19 Discharge Date 2140 1 21 Date of Birth 2117 8 7 Sex F Service

MEDICINE Allergies Penicillins Attending:[**First Name3 LF 2297 Chief Complaint **headache**

SYMPTOM Major Surgical or Invasive Procedure **Hemodialysis TREATMENT** History of

Present Illness Ms. Known lastname is a 22 year old female with **SLE DISEASE** **lupus**

nephritis DISEASE **ESRD DISEASE** on HD malignant HTN h/o TTP and HOCM who

presents with HA and **hypertensive urgency DISEASE** Awoke this a.m. with 8/10 left sided

frontal HA was n't sure **if HYPOTHETICAL** it was d/t flare of uveitis that had started on

Monday or d/t HTN Decided to skip HD and come to ED for evaluation **No**

NEGATED_EXISTENCE vision changes numbness weakness change in gait chest pain SOB +

Diarrhea SYMPTOM x 1 day In ED patient was 217/140 but elevated to 254/152 > received

labetolol IV 30 mg x 1 and MSO4 4 mg and pressures dropped to SBPs 208 and HA improved

Repeat labetolol with 50 mg x 1 and repeated dose of morphine dropped pressures to 193/134

> labetolol gtt started asa given and HA resolved Head CT negative for intracranial bleed and

CXR unremarkable ROS cold for past week **no NEGATED_EXISTENCE** fevers **chills SYMPTOM**

CP SOB N/V + **diarrhea SYMPTOM** Upon arrival to the floor patient 's BP was 191/126

labetolol gtt was not started No sxs no HA She states that she is compliant with all her meds

and her **mother FAMILY** cooks **with no NEGATED_EXISTENCE** salt and she has been

adherent to diet **Past Medical History HISTORICAL** 1 Lupus 2134 Diagnosed after she began

to have swollen fingers a rash and painful joints 2 **ESRD DISEASE** secondary to **SLE DISEASE**

2135 Was initially on cytoxan 1 dose every 3 months for 2 years until began **dialysis**

TREATMENT 3 times a week in 2137 T Th Sat Awaiting living donor transplant from mother 3

HTN 2137 Normal BPs run 180's/120 's Has had 1 hypertensive crisis that precipitated seizures

in the past 4 Uveitis secondary to **SLE DISEASE** 4 15 5 HOCM per Echo in 2137 6 Vaginal

bleeding 2139 9 20 7 Multiple episodes of **dialysis TREATMENT** reactions 8 Anemia 9 Coag

neg Staph bacteremia and HD line infection 6 15 10 H/O UE clot was on coumadin but no

longer Social History Lives in Location 669 with mother and 16 year old brother Graduated

Name2 NI School and then got sick so currently is not working or attending school **Denies**

NEGATED_EXISTENCE any T/E/D. **Family FAMILY** History -No **history HISTORICAL** of
SLE DISEASE -Grandfather has HTN -Distant history of DM -No history of clotting disorders
 -No other **history HISTORICAL** of other autoimmune diseases Physical Exam Vitals 98.0
 173/51 86 15 100 RA HEENT L eye injected w/periorbital edema R eye reactive w/ EOMI
 anicteric sclera MMM OP clear Neck supple no LAD **no NEGATED_EXISTENCE** thyromegaly
Cardiac DISEASE RRR NL S1 and S2 + S4 III/VI systolic ejection murmur LUSB radiating to
 apex and axilla intensifies w/ Valsalva no rub Lungs CTAB no wheezes rhonchi crackles Abd soft
 NTND NABS no HSM no rebound or guarding GU no CVAT Ext warm 2 + DP pulses **no**
NEGATED_EXISTENCE C/C/E L femoral **dialysis TREATMENT** catheter Neuro AOx3 CN II-XII
 intact strength/sensation grossly intact Pertinent Results UA mod bld 100 protein present on
 prior UAs Radiology CXR No acute CP abnormality EKG NSR nml axis nml intervals borderline
 LAE LVH J point elevation in V2,V3 TWI I aVL V5 V6 No change when compared to prior on
 2139 11 26 CT HEAD **No NEGATED_EXISTENCE** intracranial hemorrhage Brief Hospital Course
 A/P Patient is a 22 year old female with **SLE DISEASE** **lupus nephritis DISEASE** **ESRD**
DISEASE on HD who presents with **hypertensive urgency DISEASE** **Hypertensive urgency**
DISEASE Unclear precipitant Possibly secondary to pain from worsening uveitis Compliant
 with meds **Denies NEGATED_EXISTENCE** illicit and tox screen negative Patient was started on
 labetalol drip in ED with good BP response and was subsequently transitioned to PO anti-
 hypertensives in ICU with maintenance of stable SBPs in 150s-170s baseline 170s-190s Per
 nephrologist 's recommendations home **lisinopril MEDICATION** was increased to 40 mg po
 bid from 40 mg po qd for better baseline BP control **No NEGATED_EXISTENCE** clinical
 evidence of end organ damage UA difficult **ro POSSIBLE_EXISTENCE** interpret in setting of
 CRF CE 's x 1 negative **Headache SYMPTOM** **No evidence NEGATED_EXISTENCE** by CT for
 intracranial bleed Headaches were well controlled with **morphine sulfate MEDICATION** and
 had resolved by time of discharge Uveitis Followed by outpatient optho specialist Optho **not**
NEGATED_EXISTENCE consulted per patient 's request **ESRD DISEASE** Secondary to **lupus**
nephritis DISEASE On transplant list Patient received **hemodialysis TREATMENT** in house

with 500 ml ultrafiltrate without complications At dry weight of 45 kg per patient Began
 Sevelamer 800 TID with meals Given difficulty in interpreting renin and aldosterone levels in
 acutely ill patients these were **not NEGATED_EXISTENCE** drawn and will need to be drawn at
 outpatient follow up Medications on Admission **Lisinopril MEDICATION** 40 mg PO QD
Labetalol MEDICATION 600 PO TID **Valsartan MEDICATION** 320 mg PO QD **Clonidine**
MEDICATION 0.3 mg transdermal QW **Prednisone MEDICATION** 40 mg PO QD **Atropine**
MEDICATION 1 Hospital1 Prednisolone Acetate 1 Q1H Moxifloxacin eye drops qid Lorazepam
 1 mg PO Q4 6H PRN Discharge Medications 1 **Labetalol MEDICATION** 200 mg Tablet Sig
 Three 3 Tablet PO TID 3 times a day Tablet(s 2 **Clonidine MEDICATION** 0.3 mg/24 hr
MEASUREMENT Patch Weekly Sig One 1 Patch Weekly Transdermal QTHUR every Thursday 3
Atropine MEDICATION 1 Drops Sig One 1 Drop Ophthalmic Hospital1 2 times a day 4
 Lorazepam 1 mg Tablet Sig One 1 Tablet PO Q4 6H every 4 to 6 hours as needed
HYPOTHETICAL 5 **Valsartan MEDICATION** 160 mg Tablet Sig Two 2 Tablet PO DAILY Daily 6
 Prednisolone Acetate 1 Drops Suspension Sig One 1 Drop Ophthalmic Q1H every hour 7
Lisinopril MEDICATION 40 mg Tablet Sig One 1 Tablet PO twice a day Disp:*60 Tablet(s
 Refills:*2 8 **Sevelamer MEDICATION** 800 mg Tablet Sig One 1 Tablet PO TID 3 times a day
 Disp:*90 Tablet(s Refills:*2 9 **Prednisone MEDICATION** 20 mg Tablet Sig Two 2 Tablet PO
 once a day 10 Blood Pressure Kit Kit Sig One 1 Kit Miscellaneous once a day Disp:*1 Kit
 Refills:*0 Discharge Disposition Home Discharge Diagnosis **Hypertensive urgency DISEASE**
 Discharge Condition Good Discharge Instructions Please take all of your blood pressure
 medications as prescribed You should adhere to a low-salt diet as increased levels of sodium
 can drive your blood pressure up You are being discharged with a prescription for a home
 blood pressure monitor which you can use to take daily measurements You should call your
 primary care physician for Name Initial PRE systolic blood pressures greater than 180 or if
HYPOTHETICAL you experience headaches **nausea SYMPTOM** **vomiting SYMPTOM**
 chest pain **shortness of breath SYMPTOM** or **any other NEGATED_EXISTENCE** concerning
 symptoms Followup Instructions Please resume **hemodialysis TREATMENT** according to your

regular schedule You are scheduled to see Dr. First Name8 NamePattern2 Last Name
NamePattern1 4883 in the Division of Nephrology on Wednesday 2 3 at 9:30 AM Please call
Telephone/Fax 1 435 if you need to reschedule You are scheduled to follow-up with your
primary care physician Last Name NamePattern4 First Name4 NamePattern1 Last Name
NamePattern1 2423 on Tuesday 1 26 at 3:30 PM Please call Telephone/Fax 1 250 if you need to
reschedule You have been referred to see Dr. First Name4 NamePattern1 Last Name
NamePattern1 2539 in the Division of Hematology for further evaluation of your anemia This
appointment is scheduled for 2 9 at 3 p.m. His office is located on the Location un of the
Hospital Ward Name 23 Building on the Hospital1 18 Hospital Ward Name 516 Please call Dr.
[**Name NI 44536 administrative assistant Doctor First Name 8982 at Telephone/Fax 1 32192 if
you need to confirm or reschedule