

**Superior University Gold Campus**

**PAI Lab Task # 9**

Name: Syed Ejiz Ul Hassan Kazmi

Roll. No: SU92-BSDSM-F23-025

Section: BSDS-4A

Instructor: Sir Rasikh Ali

**Named Entity Relationship with spacy and NLTK**

This task focuses on extracting useful information from text using Natural Language Processing (NLP) techniques.  
It covers:

* Named Entity Recognition (NER) using SpaCy
* Relationship Extraction
* NER using NLTK and SpaCy together

## **1.** **Named Entity Recognition (NER) using SpaCy**

**NER** is the process of locating and classifying named entities (persons, organizations, locations, dates, etc.) in text.

### Steps:

* The SpaCy library is imported.
* The small English model en\_core\_web\_sm is loaded.
* An example text "Donald Trump is president of America" is analyzed.
* Detected entities and their types are printed.

### Output:

### 

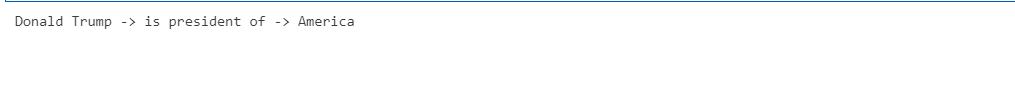
## **2.** **Relationship Extraction**

**Relationship Extraction** means identifying connections between entities, such as "X is president of Y."

### Steps:

* The tokens from the parsed text are looped over.
* If a word is an attribute and its head is an auxiliary verb, the subject and object are identified.
* If a subject and a location entity (GPE) are found, the relationship is printed.

**Output:**



## **3.** **Named Entity Recognition (NER) using NLTK and SpaCy**

This part combines web scraping, NLTK (Natural Language Toolkit), and SpaCy to perform NER on a real-world example.

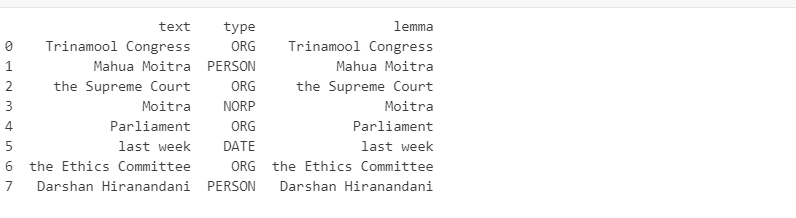
### Steps:

* Install and import beautifulsoup4, pandas, nltk, and requests.
* Define a sample text discussing a real political event.
* Process the text with SpaCy.
* Extract and display named entities along with:
  + The entity text
  + Start and end character positions
  + The type of entity.

### 4. ****Organizing Entities into a DataFrame****

* The extracted entities are organized into a neat table using pandas DataFrame, listing:
  + Text
  + Type
  + Lemma

**Output:**



\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_