	Bansilal Ramnath Agarwal Charitable Trust's Vishwakarma Institute of Information Technology **Department of** **Artificial Intelligence and Data Science**

Name: **Siddhesh Dilip Khairnar**		
Class: **TY**	Division: **B**	Roll No: **372028**
Semester: **6th**	Academic Year: **2023-24**	
Subject Name & Code: **Natural Language and Processing & ADUA32203**		
Title of Assignment: **Recognize Named Entities from a given paragraph.**		
Date of Performance: **15-03-2024**	Date of Submission: **22-03-2024**	

ASSIGNMENT NO: - 9

Aim: Recognize Named Entities from a given paragraph.

❖ **Theory:**

1. Brief discussion Name Entities and their types

- Named entities are specific words or phrases that refer to objects, people, places, organizations, or other types of entities that have a proper name.
- These named entities are crucial for understanding the meaning of a piece of text and extracting valuable information from it.
- There are several types of named entities that are commonly used in natural language processing (NLP). Some of the most common types include:

1) **Person:** Refers to the name of an individual, such as "John Smith" or "Marilyn Monroe".
2) **Organization:** Refers to the name of a company, institution, or other type of organization, such as "Apple Inc." or "The United Nations".
3) **Location:** Refers to the name of a place or geographic location, such as "New York City" or "Mount Everest".
4) **Date:** Refers to a specific date or time, such as "June 12th, 2022" or "2:30 PM".
5) **Time:** Refers to a specific time of day, such as "8:00 AM" or "12:30 PM".
6) **Money:** Refers to a specific amount of money, such as "$50" or "€100".
7) **Percentage:** Refers to a specific percentage, such as "10%" or "50.5%".
8) **Product:** Refers to the name of a specific product or brand, such as "iPhone" or "Coca-Cola".
9) **Event:** Refers to the name of a specific event, such as "The Super Bowl" or "The Oscars".

- Named entity recognition (NER) is a process in NLP that aims to identify and classify these named entities in text data, and is a crucial step in many NLP applications, such as information retrieval, question answering, and text summarization.

❖ **CODE**:

```
import nltk

nltk.download('averaged_perceptron_tagger')
nltk.download('maxent_ne_chunker')
nltk.download('words')
nltk.download('brown')

```

```
[nltk_data] Downloading package averaged_perceptron_tagger to
[nltk_data]     /root/nltk_data...
[nltk_data]   Unzipping taggers/averaged_perceptron_tagger.zip.
[nltk_data] Downloading package maxent_ne_chunker to
[nltk_data]     /root/nltk_data...
[nltk_data]   Unzipping chunkers/maxent_ne_chunker.zip.
[nltk_data] Downloading package words to /root/nltk_data...
[nltk_data]   Unzipping corpora/words.zip.
[nltk_data] Downloading package brown to /root/nltk_data...
[nltk_data]   Unzipping corpora/brown.zip.
True
```

```
from nltk import word_tokenize, pos_tag, ne_chunk

sentence = "John Smith is a software engineer at Google."
tokens = word_tokenize(sentence)
tagged = pos_tag(tokens)
ner = ne_chunk(tagged)
print(ner)

```

```
(S
  (PERSON John/NNP)
  (PERSON Smith/NNP)
  is/VBZ
  a/DT
  software/NN
  engineer/NN
  at/IN
  (ORGANIZATION Google/NNP)
  ./.)
```

```
[ ]
import spacy

nlp = spacy.load("en_core_web_sm") # Load the English language model

text = "John Smith is a software engineer at Google."

doc = nlp(text) # Parse the text

for ent in doc.ents: # Iterate over named entities
    print(ent.text, ent.label_) # Print entity text and label

```

```
John Smith PERSON
Google ORG
```

```
[ ]
!pip install textblob
```

```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Requirement already satisfied: textblob in /usr/local/lib/python3.9/dist-packages (0.17.1)
Requirement already satisfied: nltk>=3.1 in /usr/local/lib/python3.9/dist-packages (from textblob) (3.8.1)
Requirement already satisfied: click in /usr/local/lib/python3.9/dist-packages (from nltk>=3.1->textblob) (8.1.3)
Requirement already satisfied: regex>=2021.8.3 in /usr/local/lib/python3.9/dist-packages (from nltk>=3.1->textblob) (2022.10.31)
Requirement already satisfied: tqdm in /usr/local/lib/python3.9/dist-packages (from nltk>=3.1->textblob) (4.65.0)
Requirement already satisfied: joblib in /usr/local/lib/python3.9/dist-packages (from nltk>=3.1->textblob) (1.2.0)
```

```
[ ]
from textblob import TextBlob

# create a textblob object
text = "John Smith works at Apple in California."
blob = TextBlob(text)

# extract named entities
entities = blob.noun_phrases

# print named entities
print(entities)

```

```
['john smith', 'apple', 'california']
```

- **Conclusion: Thus, we have successfully understood how to Recognize Named Entities from a given paragraph.**