

**Title:** A Focus on Media Coverage and Geographical Representation

**Team Members:** Kossi Akplaka, Saloua Daouki

During the semester, our class delved into Natural Language Processing (NLP) and its diverse applications, such as sentiment analysis and spam detection. This project aims to extend our understanding of NLP, specifically focusing on Named Entity Recognition (NER). NER plays a pivotal role in extracting and categorizing entities, which are crucial for gaining insights from unstructured text data.

**Project Objectives:**

The project will primarily investigate the applications and limitations of NER. To achieve this, we will utilize NLP libraries, including spaCy, and transformer models like GPTs and BERT.

**Research Questions:**

1. **Media Coverage Disparities:** Does media coverage differ between the conflicts in Israel and Ukraine? Using Named Entity Recognition (NER), we'll analyze news articles to see if there's a noticeable discrepancy in how these two separate wars are reported.
2. **Geographical Representation in NYT Articles:** The second research question aims to explore how The New York Times (NYT) covers different regions of the world. Using NLP techniques, including NER, we will analyze metadata from NYT articles to identify and categorize named entities associated with various geographical locations. This analysis will provide insights into the geographic focus of NYT articles, contributing to a comprehensive understanding of media representation.

**Data Sources:** New York Times API or Bing News Search API.

**Programming Language:** Python