**NLP and LLMs in Social Sciences**

**8am – 5pm**

Suggestion: 3-hour bootcamp on “Python for Data Science” a day before for those who have little experience with Python

**1. Introduction to NLP and LLMs (1 Hour)**

The basic concepts of NLP and LLMs, their history, and their significance.

* Presentation on the evolution of NLP.
* Overview of LLMs like GPT-3, BERT, Mistral, Llama and their mechanisms.
* LM Studio, Hugging Face
* Discussion on the relevance of NLP in social sciences.

**2. Basics of AI and Deep Learning (2 Hours)**

* Introduction to Artificial Intelligence: Definitions, scope, and history.
* Basics of neural networks: Neurons, layers, activation functions.
* Overview of deep learning architectures and their applications.
* Selected architectures: NN, RNN, LSTM
* Simple demonstrations using Python to illustrate how neural networks work (e.g., using TensorFlow).

**3. Key Concepts in NLP (1.5 Hours)**

* Text preprocessing: Tokenization, stemming, lemmatization.
* Vectorization techniques: Bag of Words, TF-IDF.
* Introduction to Word Embeddings and Models like word2vec, GloVe.
* Hands-on Python session using libraries like NLTK and Gensim to demonstrate text preprocessing and generating word embeddings.

Lunch Break (1 Hour) 12:30-1:30

**4. Introduction to Transformers and LLMs (1.5 Hour)**

- The architecture and functioning of transformers and their application in creating LLMs.

* Explanation of the Transformer model: Attention mechanisms, position encoding.
* Overview of how LLMs are trained and fine-tuned.
* Interactive demonstration using a pre-trained transformer model from the Hugging Face library to analyze text.

**5. Practical Session: Building an NLP Project (2 Hours)**

Four different real data sets will be provided.