

# NLP Basics and Text Processing

## Introduction to NLP

Natural Language Processing (NLP) involves teaching computers to understand, interpret, and generate human language. It combines computer science, linguistics, and artificial intelligence.

Key tasks include:

- Machine Translation
- Sentiment Analysis
- Part-of-Speech (POS) Tagging
- Named Entity Recognition (NER)
- Summarization, Parsing, Dialog Systems

Challenges:

- Ambiguity
- Non-standard text (social media)
- World knowledge requirements

## Basic Text Processing

Regular Expressions:

- Pattern matching language.

Tokenization:

- Splitting text into words.

Normalization:

- Lowercasing, punctuation removal, stemming, and lemmatization.

Sentence Segmentation:

- Identifying sentence boundaries, handling abbreviations and periods.