

# Overview of NLP

---

Assignment 0

CS 4395.001: Human Language Technologies

Hannah Valena, HCV180000

## What is NLP?

Natural Language Processing (NLP) describes how computers interpret and interact with the human language. This includes human language that is spoken and written, and it aims to both understand and generate natural human language using linguistics and a variety of techniques.

## Relationship between AI and NLP

Artificial intelligence (AI) describes how computers simulate human intelligence by performing human-like actions and problem-solving processes. NLP branches from AI by focusing specifically on human language and communication.

## Natural language understanding vs. natural language generation

Natural language understanding takes either speech or text as input and interprets the meaning of the input. On the other hand, natural language generation involves computers forming valid sentences and phrases as output.

## Examples of modern NLP applications

Some examples of modern NLP applications are: - Speech to text, eg "Hey Alexa", "Hey Siri", "OK Google" - Spell check - Email filters and spam email identification - Predictive text - Search results

## The 3 main approaches to NLP

### Approach 1: Rules-based

The rules-based approach is the oldest approach to NLP. It is based on grammar and syntax rules, which often use regular expressions to generate or interpret sentences. Even if a list of exceptions are provided with the rules, this approach is problematic because of the complexity and nuances of human language.

An example of this is Eliza, the computer chatbot that doubles as a "therapist." Eliza uses regular expressions and code to emulate chatting to a therapist.

### Approach 2: Statistical and probabilistic

Statistical and probabilistic approaches to NLP are successors to the rules-based approach. These approaches involve taking a collection of data as input and repeatedly learning from this data. Language models can be created from counting word frequencies and computing the probabilities of word sequences.

An example of the usage of these models is predictive text, which is often used in messaging (text, email, etc.) and search queries.

These approaches also utilize machine learning algorithms, such as decision trees and regression analysis, to learn from large datasets. Although these require a moderate amount of processing power, they show an improvement from the rules-based approach.

### Approach 3: Deep learning

Deep learning, the newest approach to NLP, evolved from statistical and probabilistic approaches after computers' processing power skyrocketed. Deep learning uses neural networks and its variations to process and learn from huge collections of data. Vast improvements to NLP were made possible with deep learning, making computers act more and more like humans.

An example of deep learning is sentiment analysis, which can classify an inputted language expression as negative, positive, or neutral. A popular browser extension, Grammarly, uses sentiment analysis to determine the tone of text.

## **My interest in NLP**

NLP is becoming an integral part of daily life, and I am interested in learning about its current and future applications. I am excited to learn more about NLP through completing related projects, like making my own chatbot and using sentiment analysis on news articles. I want to have a deeper understanding of all three NLP approaches, from using regex and rules to using neural networks and machine learning.