- According to wikipedia, "**Natural language processing (NLP)** is an area of computer science and artificial intelligence concerned with the interactions between computers and human (natural) languages, in particular how to program computers to process and analyze large amounts of natural language data.

- Often when performing analysis, lots of data is numerical, such as sales numbers, physical measurements, quantifiable categories.
- Computers are very good at handling direct numerical information.
- But what do we do about **text data?**

- As humans we can tell there is a plethora of information inside of text documents.
- But a computer needs specialized processing techniques in order to “understand” raw text data.
- Text data is highly unstructured and can be in multiple languages!

- Natural Language Processing attempts to use a variety of techniques in order to create structure out of text data.
- In this section we will be discussing some of these basic techniques, which are built into libraries such as Spacy and NLTK.

- Example Use Cases:
  - Classifying Emails as Spam vs Legitimate
  - Sentiment Analysis of Text Movie Reviews
  - Analyzing Trends from written customer feedback forms.
  - Understanding text commands, “Hey Google, play this song”.

- Natural Language Processing is constantly evolving and great strides are made every month!
- We will focus on fundamental ideas that all state of the art techniques are based off.
- Let’s begin by learning about the basics of using the Spacy library.