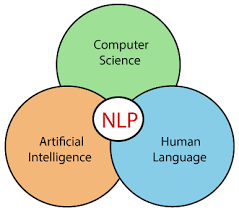
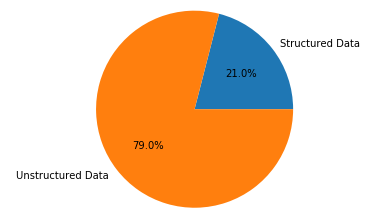
* **What is NLP?**

NLP stands for Natural Language Processing or Neuro Linguistic Programming concerned with the interactions between computers and human languages, in particular how to program computers to process and analyze large amount natural data which gives machines the ability to read, understand and derive meaning from human languages.



NLP is the way to make computers to understand unstructured data as there are total of 6500 languages that exists so coming to the 21st century according to the industry estimates only 21% of the available data is present in the structured form data is being generated as I speak, tweet and send messages on WhatsApp or the various of the crops of Facebook and majority of the state exists in the textual form which is highly unstructured in nature.



* **Applications of NLP**



1. **Sentimental Analysis**

It is one of the most important application be it Twitter or Facebook sentimental analysis as it’s being used heavily. Try out this [online sentiment analyzer](https://monkeylearn.com/sentiment-analysis-online/) to see how NLP sorts your text by emotions.

1. **Text Classification**

Email Filtering is the one of the famous application of text classification, you might have already noticed that whenever a mail arrives, it gets classified into pre-defined set of classes like spam, primary, social and promotions etc. Another example is classification of news articles into various categories.

1. **Chabot’s**

You might have used the customer chat services provided by various companies to smooth out customer experience that solves basic quires of the customer for improvement in their products and also keeping customer satisfied and the process behind all of that is because of the NLP. Recently, google released a chatbot named **“meena chatbot”** and you can also watch a movie **“Her 2013”** which is based on NLP as well.

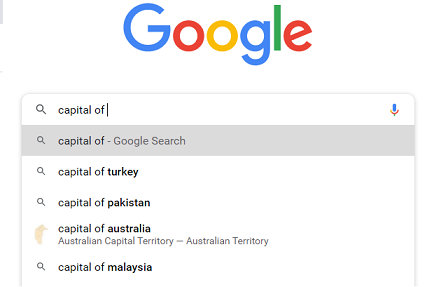
1. **Virtual Assistants** (Speech recognition + Cognition)

Voice assistant like Google Assistant, Apple Siri, Amazon Alexa and Windows Cortana are all amazing applications of NLP.

1. **Text Summarization**

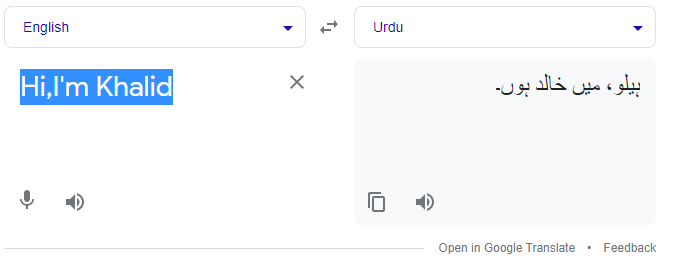
Automatic summarization is pretty self-explanatory. It summarizes text by extracting the most important information by simplifying the process of going through vast amounts of data, such as scientific papers, news content or legal documentation.

1. **Auto-Correct and Auto-Complete**

We often see this feature in Google search bar when we type 2 or 3 letters, it shows the possible search terms and also show relevant results even with typos after auto correction.

1. **Language Translator**

Another amazing application is of Google Translator which supports 108 languages of the world and still expanding it.



1. **Grammar Checker**

[Grammarly](https://www.grammarly.com/?q=brand&utm_source=google&utm_medium=cpc&utm_campaign=brand_f1&utm_content=sa360test&utm_term=gramerly&matchtype=e&placement=&network=g&gclid=CjwKCAiA1aiMBhAUEiwACw25MTPaBnlRnB5ejRqp4djUtjRdSUKvquIg91vg4S33TNUfd-rBJu9OqxoCeeQQAvD_BwE&gclsrc=aw.ds) is one of the widely used application of NLP which correct grammar, spelling, suggests better synonyms and help in delivering better clarity and engagement. This application is based on Transformers which were introduced in 2017.

1. **Hiring and Recruitment**

Mostly companies receive hundreds of resumes of different candidates so automatic process is required to shortlist the best candidate based on their technical skills.

1. **Paraphrasing**

* **NLP Components**

In order to produce significant and actionable insights from natural data it is important to get acquainted with the techniques of text analysis and NLP.

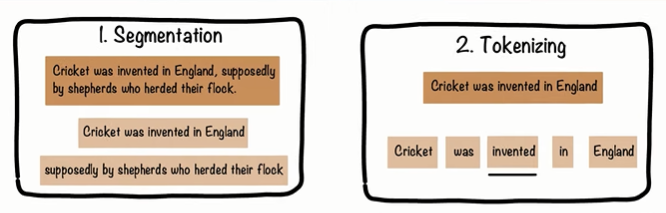
Following are the main parts of a standard NLP pipeline:

1. **Sentence Segmentation**

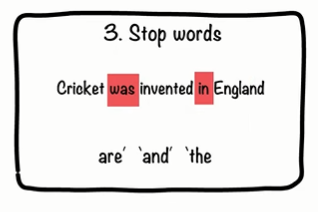
Separation of each sentence in the text.

1. **Word Tokenization**

Separation of each word of a sentence.



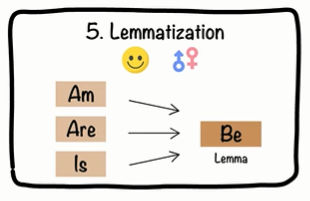
1. **Identify Stop words**

It’s a hard coded list of all stop words which is used to remove stop words from the training data to improve model accuracy

1. **Text Lemmatization**

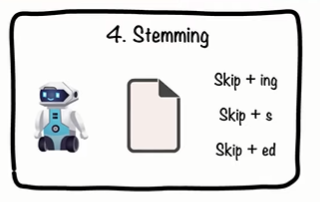
Groups together different inflected forms of a word, called lemma by using look-

up tables of words. Its output is a word. Like Going, Gone and went will be,

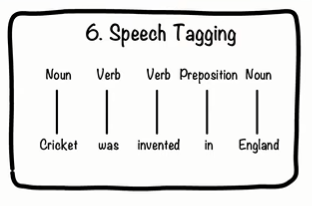
 marked as Go and Is, am, are as be. It is mostly on verbs.

1. **Text Stemming**

Normalize words into its base or root form. Somehow similar to Lemmatization. Skips, skipping, skipped as skip



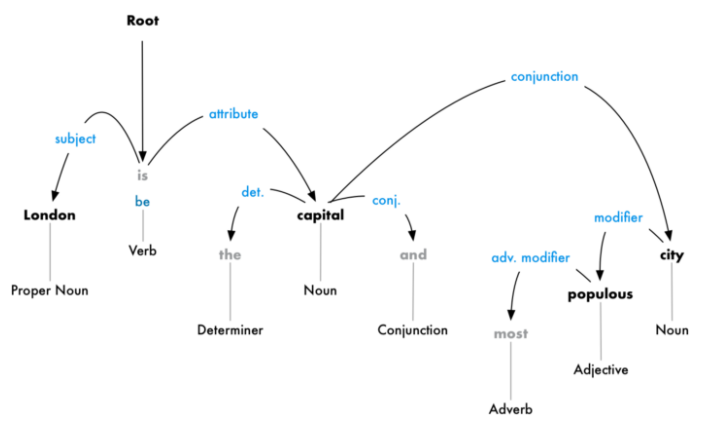
.

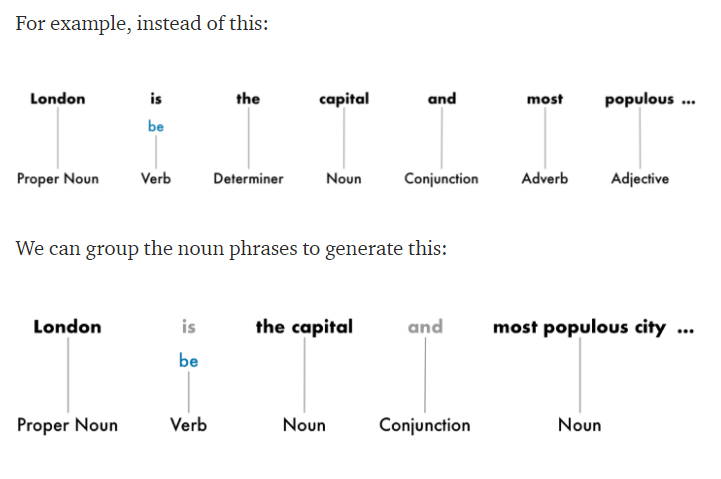
1. **Predict Part of Speech**

Tag each word to its part of speech.

1. **Dependency Parsing & Finding Noun Phrases**(Chunking)

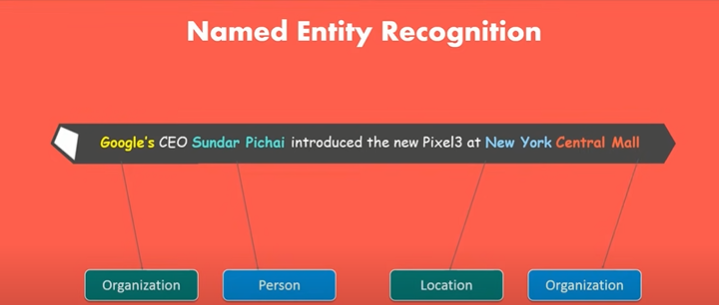
Picking up individual pieces of information and grouping them into bigger pieces based on their nouns.





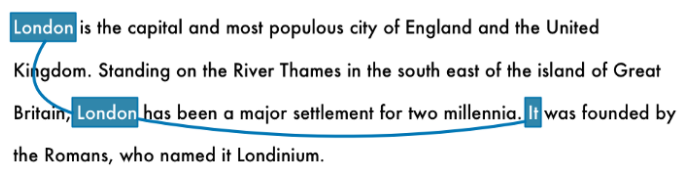
1. **NER Prediction**

Find people’s names, company names, geographical locations, product names, dates and times, amounts of money, names of events.

****

1. **Co-reference Resolution**

It is basically to trace back pronouns then to get all the relevant information about a noun using “Neuralcoref” library like in below example;



* **NLP Major Libraries**

1. SpaCy
2. Textacy
3. Neuralcoref
4. NLTK tool kit.

* **BERT** (Bi-directional Encoder Representations from transformers) is the SOTA (state of the art) NLP model that’s being used behind Google search bar, NER models, Meena chatbot and in many other NLP applications.

References:

1. [Natural Language Processing is Fun! | by Adam Geitgey | Medium](https://medium.com/@ageitgey/natural-language-processing-is-fun-9a0bff37854e)