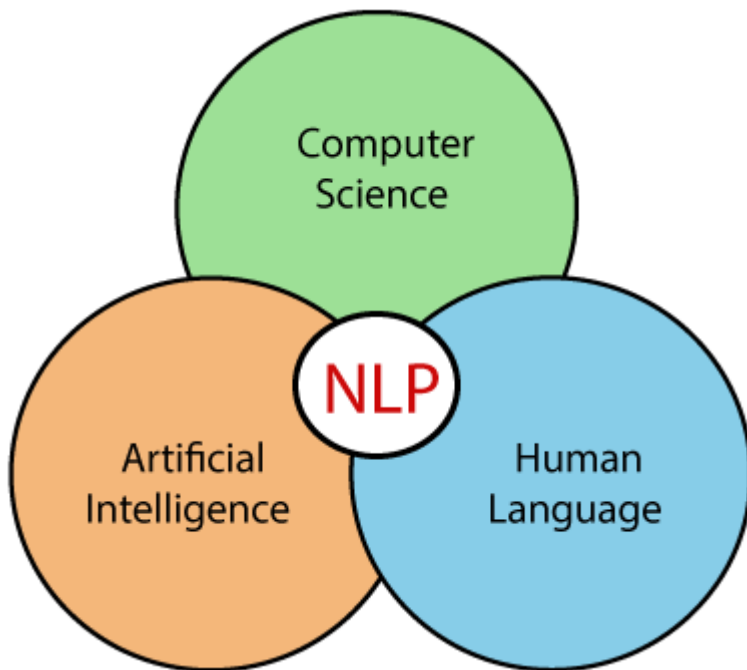


## What is NLP?

NLP stands for **Natural Language Processing**, which is a part of **Computer Science**, **Human language**, and **Artificial Intelligence**. It is the technology that is used by machines to understand, analyse, manipulate, and interpret human's languages. It helps developers to organize knowledge for performing tasks such as **translation**, **automatic summarization**, **Named Entity Recognition (NER)**, **speech recognition**, **relationship extraction**, and **topic segmentation**.



---

## History of NLP

### (1940-1960) - Focused on Machine Translation (MT)

The Natural Languages Processing started in the year 1940s.

**1948** - In the Year 1948, the first recognisable NLP application was introduced in Birkbeck College, London.

**1950s** - In the Year 1950s, there was a conflicting view between linguistics and computer science. Now, Chomsky developed his first book syntactic structures and claimed that language is generative in nature.

In 1957, Chomsky also introduced the idea of Generative Grammar, which is rule based descriptions of syntactic structures.

### (1960-1980) - Flavored with Artificial Intelligence (AI)

In the year 1960 to 1980, the key developments were:

#### **Augmented Transition Networks (ATN)**

Augmented Transition Networks is a finite state machine that is capable of recognizing regular languages.

## Case Grammar

Case Grammar was developed by **Linguist Charles J. Fillmore** in the year 1968. Case Grammar uses languages such as English to express the relationship between nouns and verbs by using the preposition.

In Case Grammar, case roles can be defined to link certain kinds of verbs and objects.

**For example:** "Neha broke the mirror with the hammer". In this example case grammar identify Neha as an agent, mirror as a theme, and hammer as an instrument.

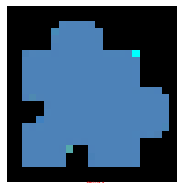
In the year 1960 to 1980, key systems were:

### SHRDLU

SHRDLU is a program written by **Terry Winograd** in 1968-70. It helps users to communicate with the computer and moving objects. It can handle instructions such as "pick up the green boll" and also answer the questions like "What is inside the black box." The main importance of SHRDLU is that it shows those syntax, semantics, and reasoning about the world that can be combined to produce a system that understands a natural language.

### LUNAR

LUNAR is the classic example of a Natural Language database interface system that is used ATNs and Woods' Procedural Semantics. It was capable of translating elaborate natural language expressions into database queries and handle 78% of requests without errors.



## 1980 - Current

Till the year 1980, natural language processing systems were based on complex sets of hand-written rules. After 1980, NLP introduced machine learning algorithms for language processing.

In the beginning of the year 1990s, NLP started growing faster and achieved good process accuracy, especially in English Grammar. In 1990 also, an electronic text introduced, which provided a good resource for training and examining natural language programs. Other factors may include the availability of computers with fast CPUs and more memory. The major factor behind the advancement of natural language processing was the Internet.

Now, modern NLP consists of various applications, like **speech recognition, machine translation, and machine text reading**. When we combine all these applications then it allows the artificial intelligence to gain knowledge of the world. Let's consider the example of AMAZON ALEXA, using this robot you can ask the question to Alexa, and it will reply to you.

---

## Advantages of NLP

- NLP helps users to ask questions about any subject and get a direct response within seconds.
  - NLP offers exact answers to the question means it does not offer unnecessary and unwanted information.
  - NLP helps computers to communicate with humans in their languages.
  - It is very time efficient.
  - Most of the companies use NLP to improve the efficiency of documentation processes, accuracy of documentation, and identify the information from large databases.
- 

## Disadvantages of NLP

A list of disadvantages of NLP is given below:

- NLP may not show context.
  - NLP is unpredictable
  - NLP may require more keystrokes.
  - NLP is unable to adapt to the new domain, and it has a limited function that's why NLP is built for a single and specific task only.
- 

## Components of NLP

There are the following two components of NLP -

### 1. Natural Language Understanding (NLU)

Natural Language Understanding (NLU) helps the machine to understand and analyse human language by extracting the metadata from content such as concepts, entities, keywords, emotion, relations, and semantic roles.

NLU mainly used in Business applications to understand the customer's problem in both spoken and written language.

NLU involves the following tasks -

- It is used to map the given input into useful representation.
- It is used to analyze different aspects of the language.

## 2. Natural Language Generation (NLG)

Natural Language Generation (NLG) acts as a translator that converts the computerized data into natural language representation. It mainly involves Text planning, Sentence planning, and Text Realization.

Note: The NLU is difficult than NLG.

### Difference between NLU and NLG

NLU	NLG
NLU is the process of reading and interpreting language.	NLG is the process of writing or generating language.
It produces non-linguistic outputs from natural language inputs.	It produces constructing natural language outputs from non-linguistic inputs.

### Applications of NLP

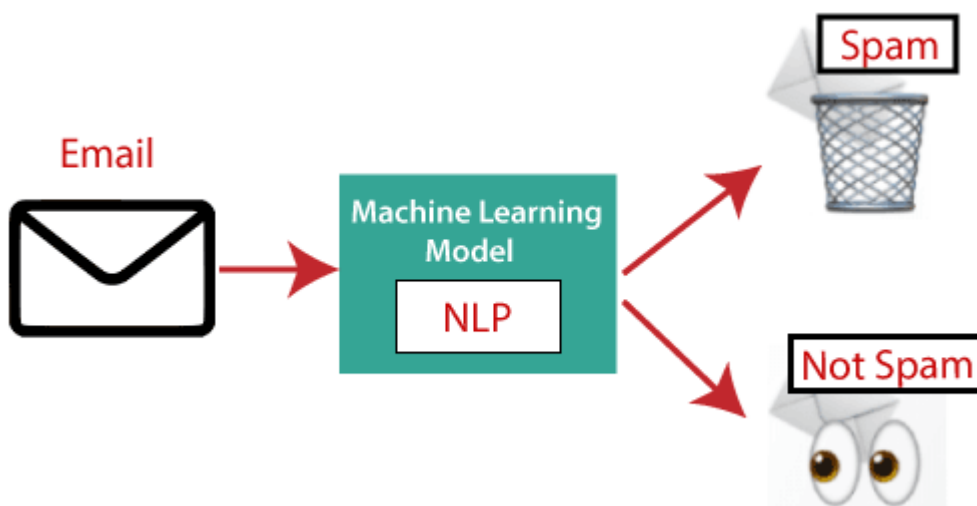
There are the following applications of NLP -

#### 1. Question Answering

Question Answering focuses on building systems that automatically answer the questions asked by humans in a natural language.

#### 2. Spam Detection

Spam detection is used to detect unwanted e-mails getting to a user's inbox.



#### 3. Sentiment Analysis

Sentiment Analysis is also known as **opinion mining**. It is used on the web to analyse the attitude, behaviour, and emotional state of the sender. This application is implemented through a

combination of NLP (Natural Language Processing) and statistics by assigning the values to the text (positive, negative, or natural), identify the mood of the context (happy, sad, angry, etc.)



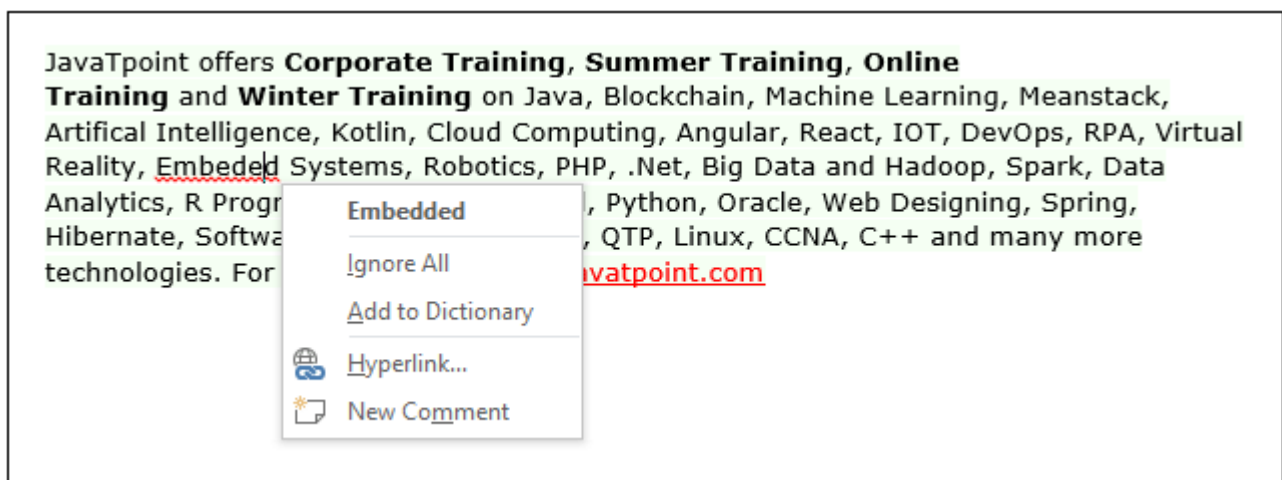
#### 4. Machine Translation

Machine translation is used to translate text or speech from one natural language to another natural language.

**Example:** Google Translator

#### 5. Spelling correction

Microsoft Corporation provides word processor software like MS-word, PowerPoint for the spelling correction.

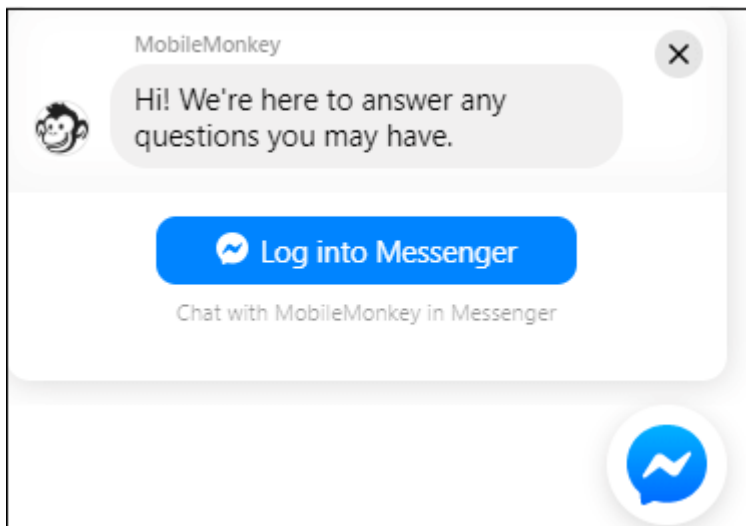


#### 6. Speech Recognition

Speech recognition is used for converting spoken words into text. It is used in applications, such as mobile, home automation, video recovery, dictating to Microsoft Word, voice biometrics, voice user interface, and so on.

#### 7. Chatbot

Implementing the Chatbot is one of the important applications of NLP. It is used by many companies to provide the customer's chat services.



## 8. Information extraction

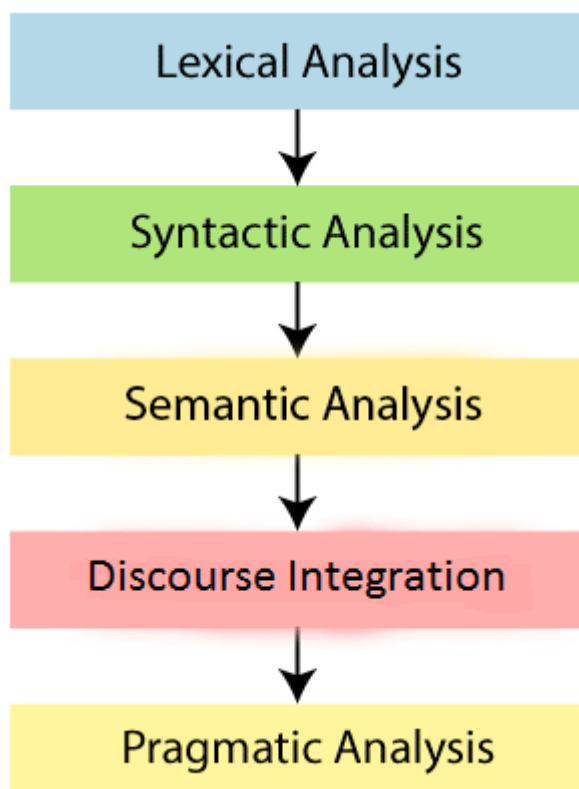
Information extraction is one of the most important applications of NLP. It is used for extracting structured information from unstructured or semi-structured machine-readable documents.

## 9. Natural Language Understanding (NLU)

It converts a large set of text into more formal representations such as first-order logic structures that are easier for the computer programs to manipulate notations of the natural language processing.

### Phases of NLP

There are the following five phases of NLP:



## **1. Lexical Analysis and Morphological**

The first phase of NLP is the Lexical Analysis. This phase scans the source code as a stream of characters and converts it into meaningful lexemes. It divides the whole text into paragraphs, sentences, and words.

## **2. Syntactic Analysis (Parsing)**

Syntactic Analysis is used to check grammar, word arrangements, and shows the relationship among the words.

**Example:** kAgra goes to the Poonam

In the real world, Agra goes to the Poonam, does not make any sense, so this sentence is rejected by the Syntactic analyzer.

## **3. Semantic Analysis**

Semantic analysis is concerned with the meaning representation. It mainly focuses on the literal meaning of words, phrases, and sentences.

## **4. Discourse Integration**

Discourse Integration depends upon the sentences that proceeds it and also invokes the meaning of the sentences that follow it.

## **5. Pragmatic Analysis**

Pragmatic is the fifth and last phase of NLP. It helps you to discover the intended effect by applying a set of rules that characterize cooperative dialogues.

**For Example:** "Open the door" is interpreted as a request instead of an order.