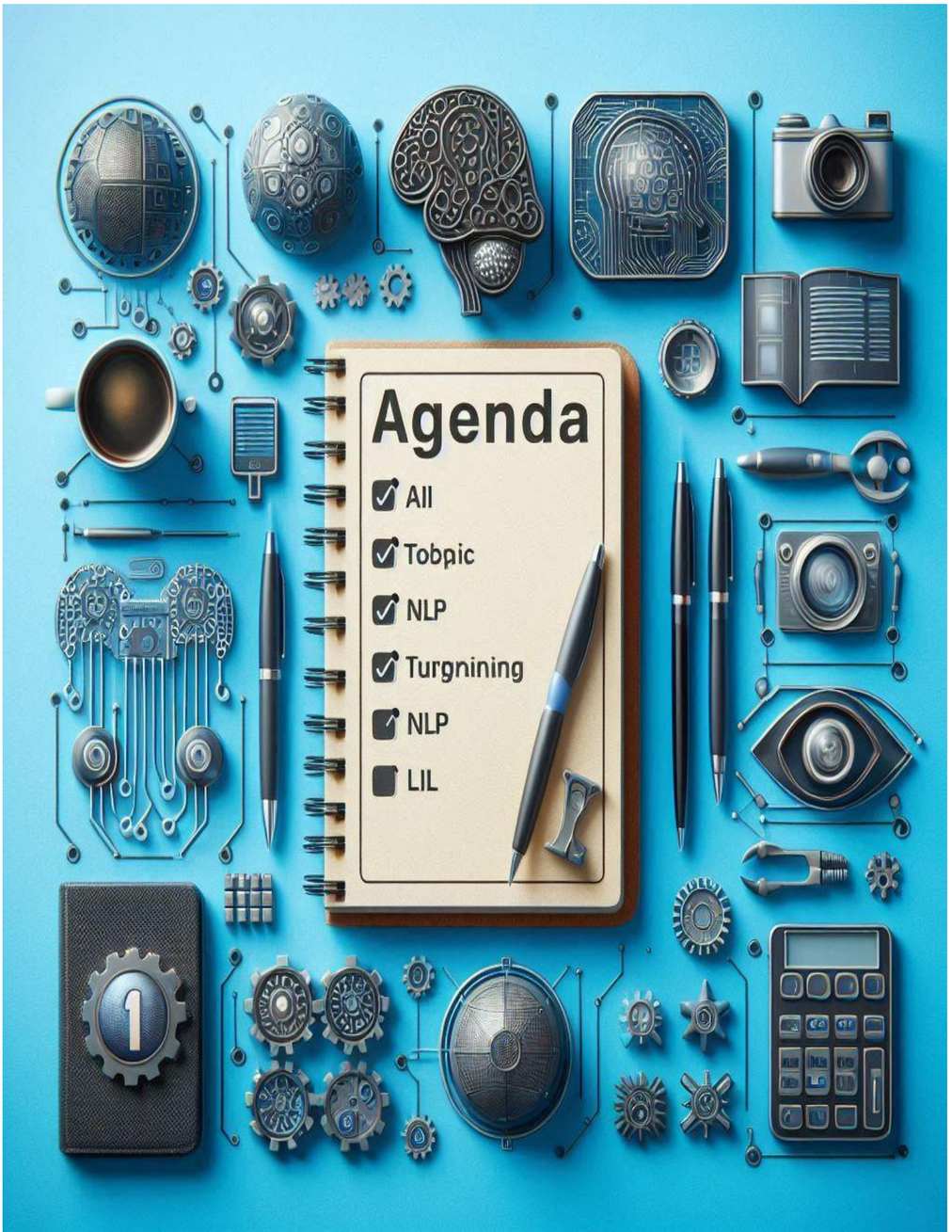
National Telecommunication Institute

*What are your expectations for the course?*

# Course Outline

1. Introduction to NLP course and Basic Concepts
   - Tokenization
   - Sentence
   - Segmentation
   - POS Tagging
2. NLP Basic Concepts
   - Stemming
   - Lemmatization
   - Named Entity Recognition
   - Stop Words
3. NLP Basic Concepts
   - Matchers
   - Text Visualization
   - Syntax Structure
4. Simple Processing
   - Bag of Words
   - Text Vectors
   - TF-IDF
5. Simple Processing
   - Word Embedding
   - Word2Vec
   - Text Similarity
   - Distance Similarity
6. Advanced Processing
   - Text Classification
   - Text Clustering
7. Modeling & Text Generation
   - LDA
   - N-Grams
   - Text Generation
8. Modern NLP Architectures
   - Attention Mechanism
   - Transformer
9. Large Language Models
   - LLMs (BERT, GPT)
   - Fine-tuning LLMs



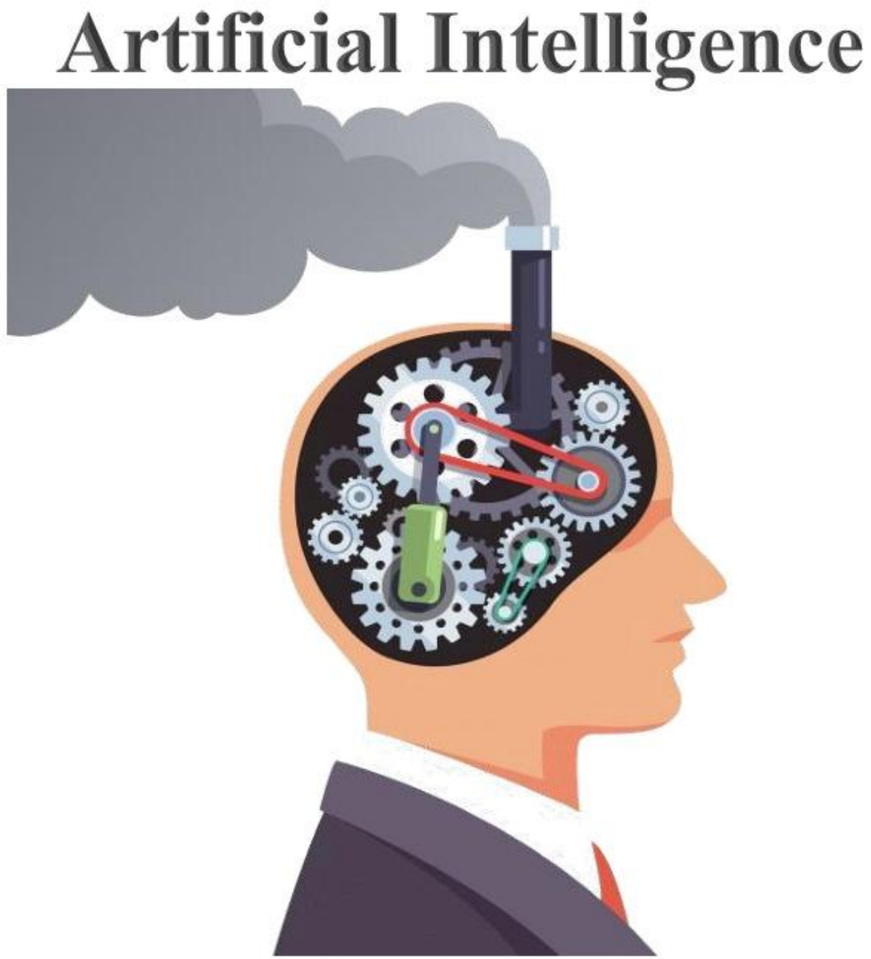

# Agenda

- Introduction to Natural Language Processing
- What is Natural Language Processing (NLP)?
- Natural Language Understanding(NLU) and Natural Language Generation(NLG)
- Applications of Natural Language Processing(NLP)
- Applications of Natural Language Understanding(NLU)
- Applications of Natural Language Generation(NLG)
- Challenges of natural language processing
- How does natural language processing work?
- Evolution of NLP
- NLP Pipeline

National Telecommunication Institute

# What is AI?

# What is AI?

# What is Intelligence?

According to the theory of multiple intelligences proposed by Professor Howard Gardner, multiple intelligences are manifested by eight capabilities:

- Linguistic-verbal intelligence
- Logical-mathematical intelligence
- Visual-spatial intelligence
- Bodily-kinesthetic intelligence
- Musical-rhythmic and harmonic intelligence
- Interpersonal intelligence
- Intrapersonal intelligence
- Naturalistic intelligence

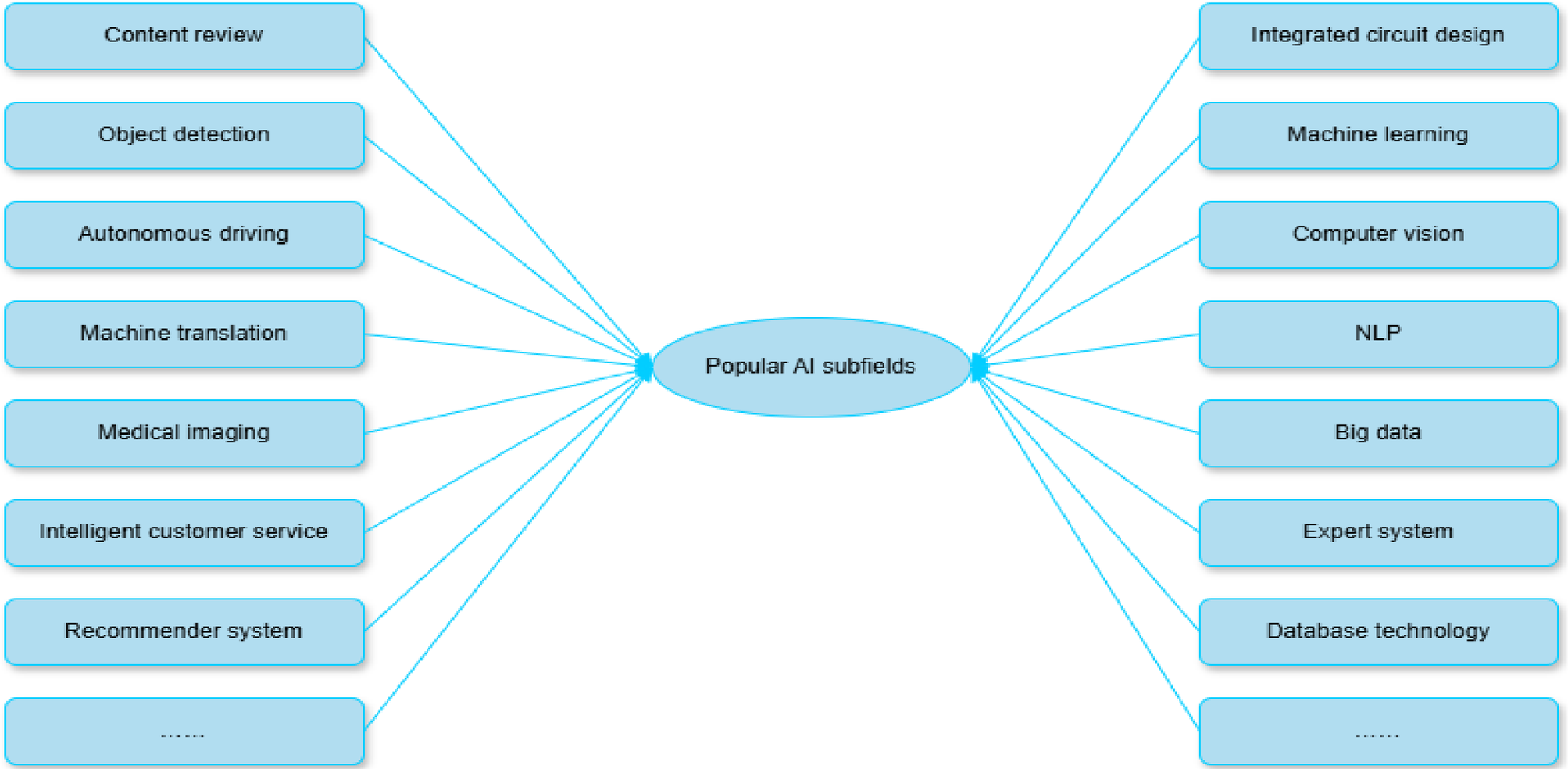
# What is AI?

**Artificial Intelligence (AI)** is the field of computer science focused on creating systems or machines that can perform tasks that typically require **human intelligence**.

These tasks include:

- **Learning** (e.g., from data or experience)
- **Understanding language** (Natural Language Processing)
- **Recognizing patterns and images** (Computer Vision)
- **Making decisions** (based on rules or data)
- **Controlling robots or autonomous systems**

# Popular AI Subfields



# What is Natural Language Processing?

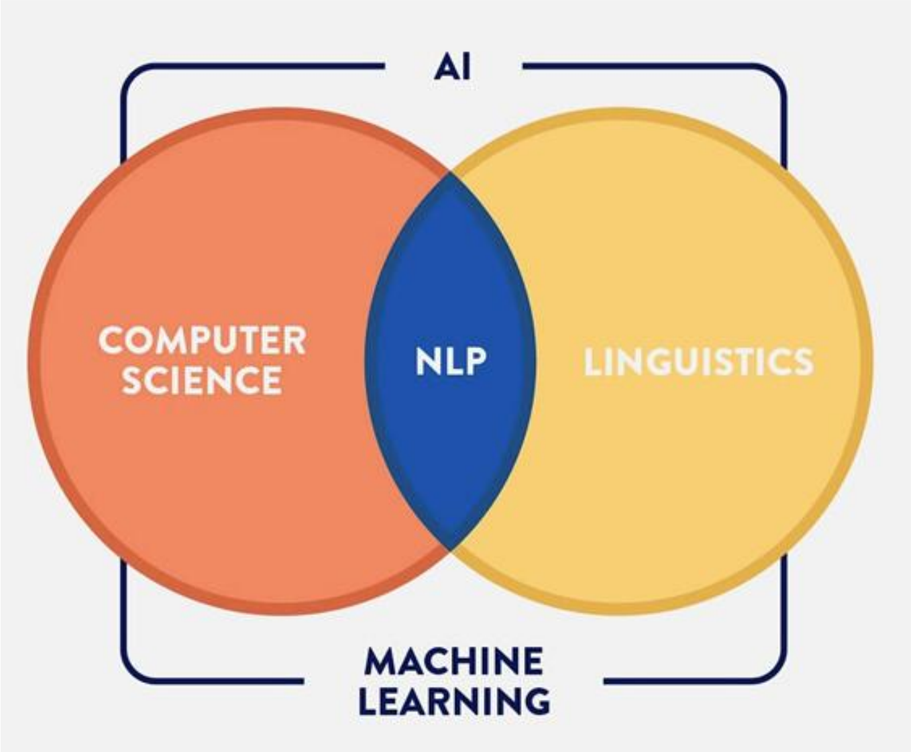
- Natural language processing (NLP) is a subfield of linguistics, computer science, and artificial intelligence concerned with the interactions between computers and human language, how to program computers to process and analyze large amounts of natural language data.
- By “natural language” we mean a language that is used for everyday communication by humans, such as Arabic, English, Spanish….etc.
- NLP is not to be confused with the abbreviation that stands for Neuro-Linguistic Programming (برمجة لغوية عصبية) which is a psychological approach that involves analyzing strategies used by successful individuals and applying them to reach a personal goal.

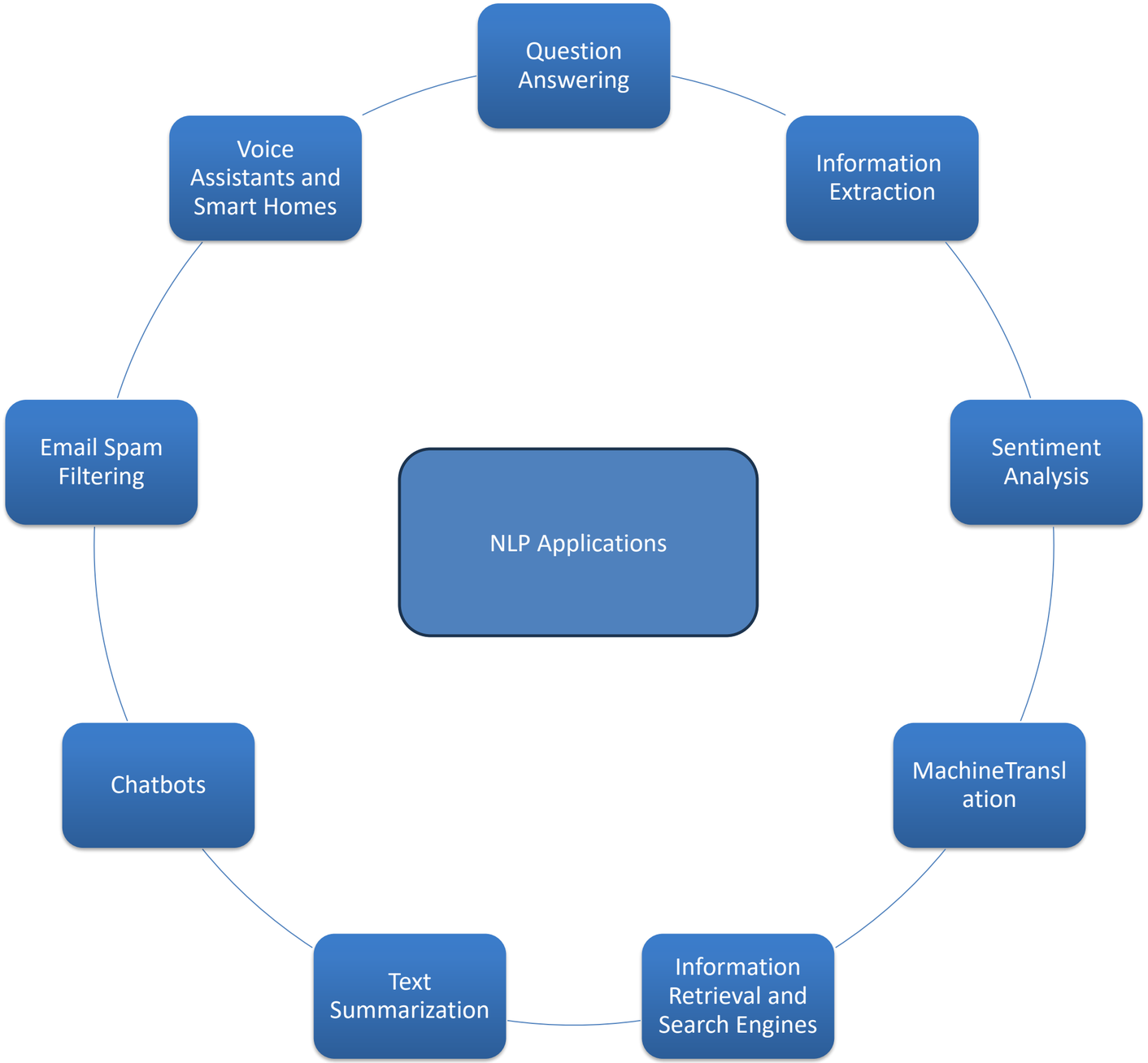
# What is NLP? (cont.)

- NLP is an Intersection of several fields
  - Computer Science
  - Artificial Intelligence
  - Linguistics

- NLP is a difficult task because it involves a lot of unstructured data.

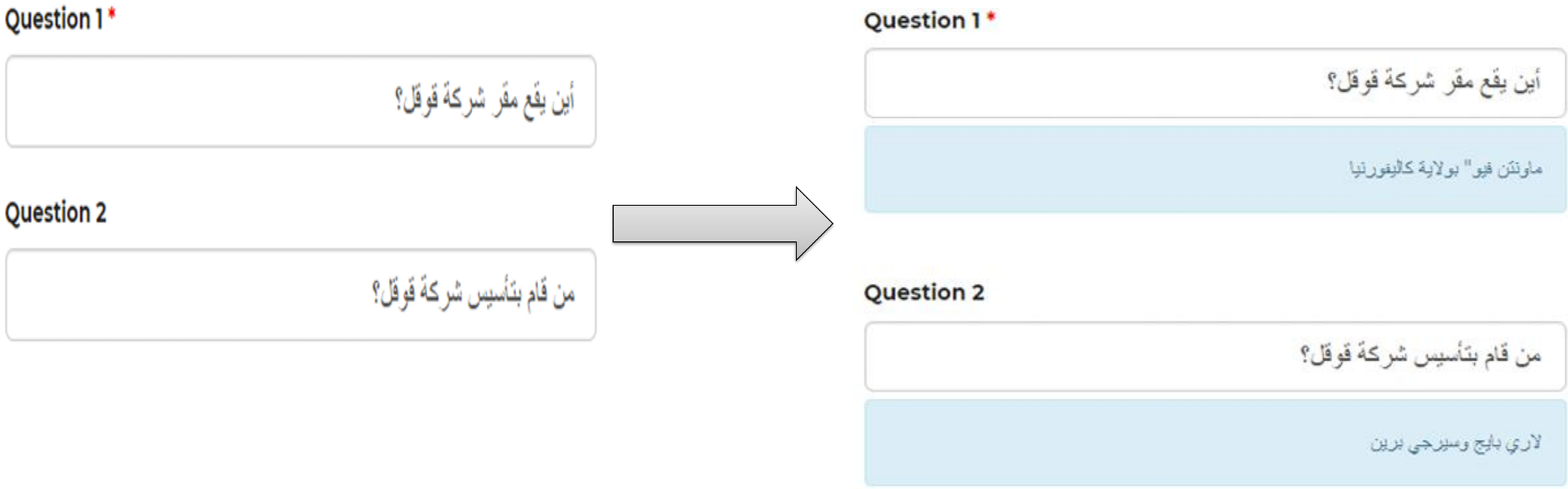
- Understanding context is an issue in NLP– that requires semantic analysis and machine learning to get a handle on it.




Question Answering
Information Extraction
Sentiment Analysis
MachineTranslation
Information Retrieval and Search Engines
Text Summarization
Chatbots
Email Spam Filtering
Voice Assistants and Smart Homes
NLP Applications

# NLP Applications
# 1.Question Answering

Question & Answer Demo Using BERT NLP for Arabic Language



# NLP Applications
# 2.InformationExtraction

Subject: curriculum meeting

Date: January 15,2012

To: Dan Jurafsky

Hi Dan,

We've now scheduled the curriculum meeting. It will be in Gates 159 tomorrow from10:00--11:30.

Event: Curriculums
Date: Jan-16-2012
Start:10:00am
End: 11:30am
Where: Gates 159

# NLP Applications
# 3. Sentiment Analysis

✓ • nice and compact to carry!

✓ • since the camera is small and light, I won't need to carry 4 around those heavy, bulky professional cameras either!

✗ • the camera feels flimsy, is plastic and very light in weight you have to be very delicate in the handling of this camera

# NLP Applications: 4.MachineTranslation

Fully automatic

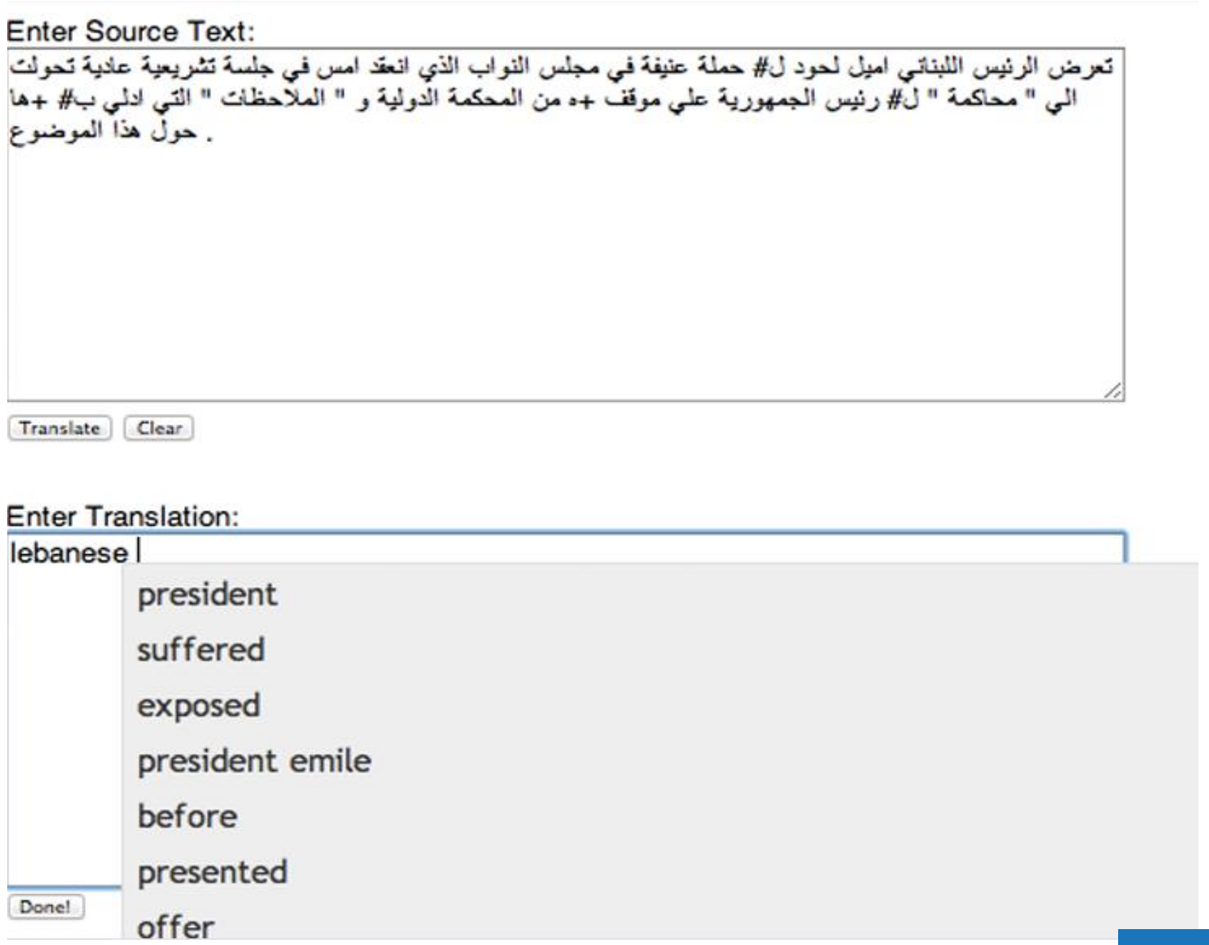
Enter Source Text:
这不过是一个时间的问题

Translation from Stanford's Phrasal

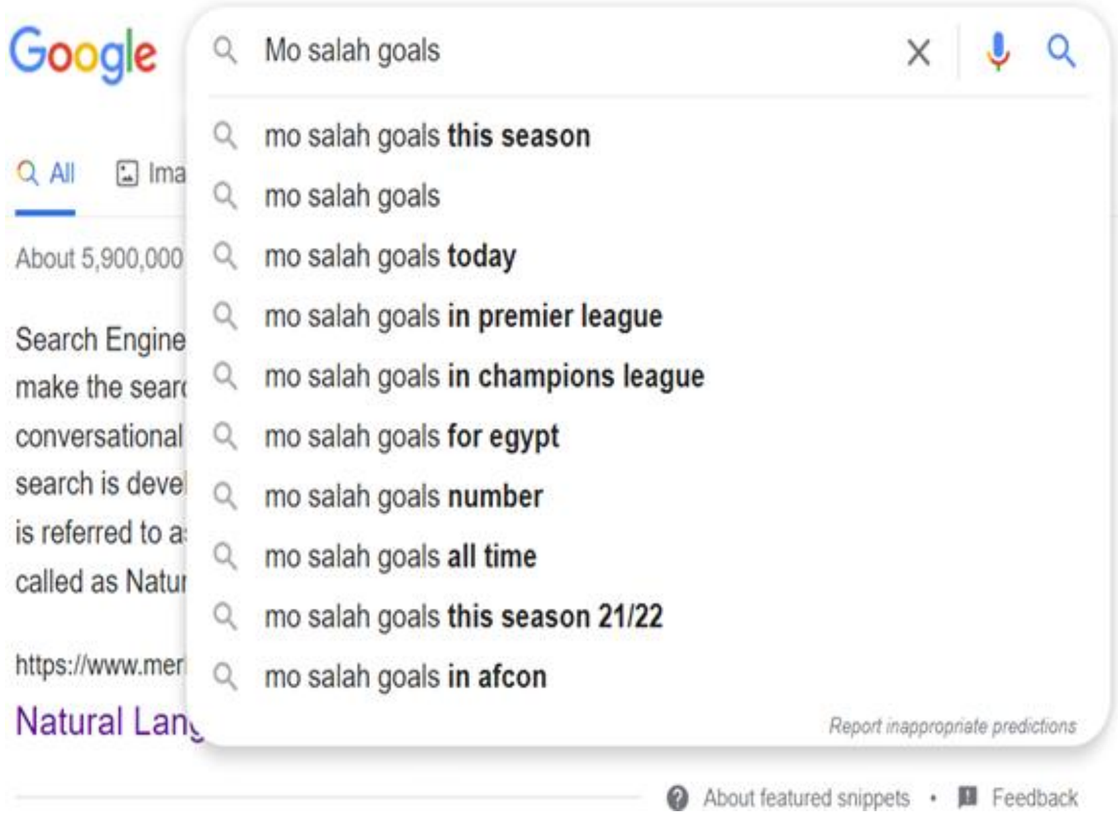
This is only a matter of time.

Google Voice Translator:
https://youtu.be/Pk6a6mvOoJA

Helping human translators by suggesting next words for example



# NLP Applications:
# 5. Information Retrieval and Search Engines

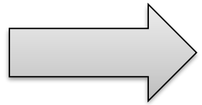

# NLP Applications:
# 6. Text Summarization

Extractive Vs. Abstractive

- Extractive: Identify and extract the important sections merely of the input text.
- Abstractive: Produce new text in a generalized form and different structure.

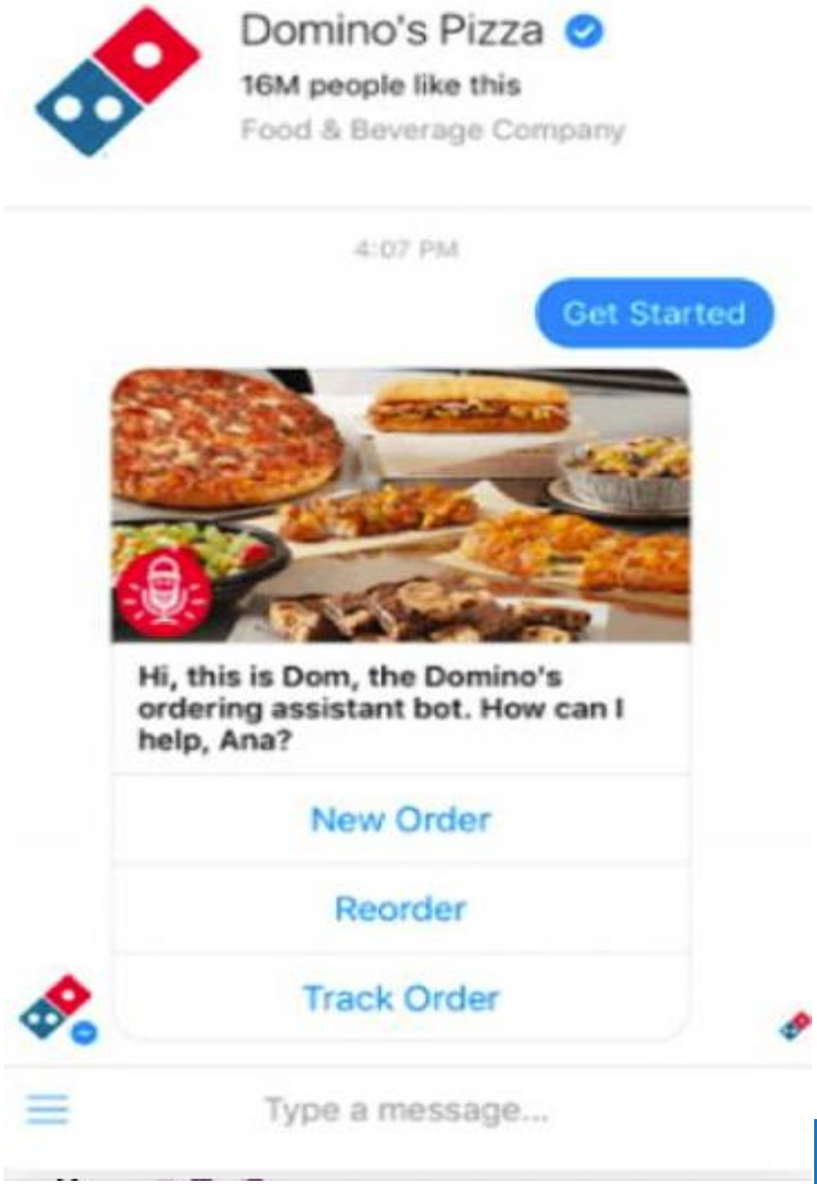
# NLP Applications:
# 6. Text Summarization

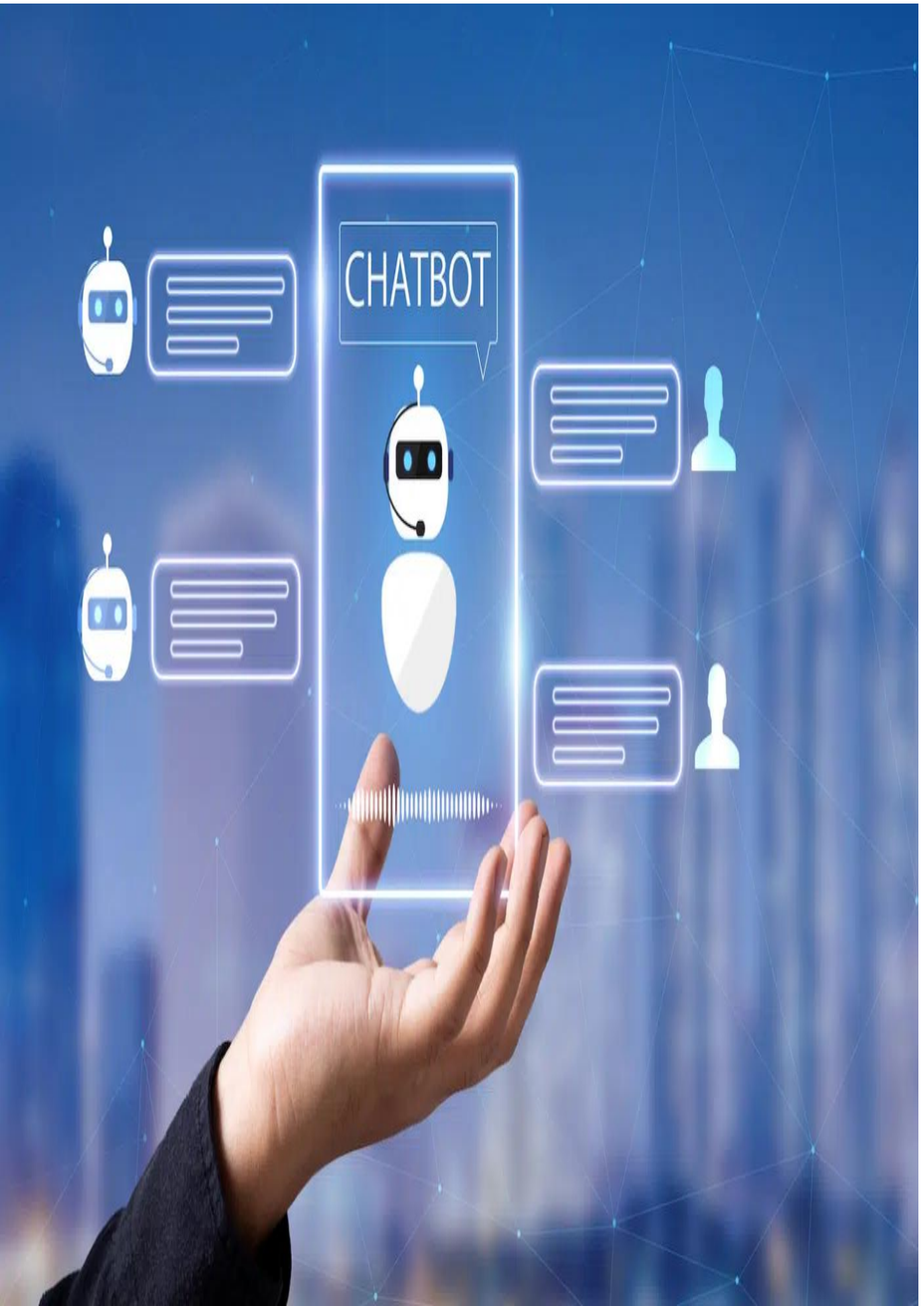
Natural Language Processing (NLP) is a field of Artificial Intelligence that enables computers to understand, interpret, and generate human language. It combines computational linguistics with machine learning and deep learning models. NLP powers applications such as chatbots, translation services, sentiment analysis, and voice assistants like Siri and Alexa. It is widely used across industries such as healthcare, finance, and customer service to automate and enhance communication processes.



NLP is a branch of AI that helps computers process and understand human language. It powers tools like chatbots and voice assistants and is used in various industries to improve communication.

# NLP Applications:
# 7. Chatbots



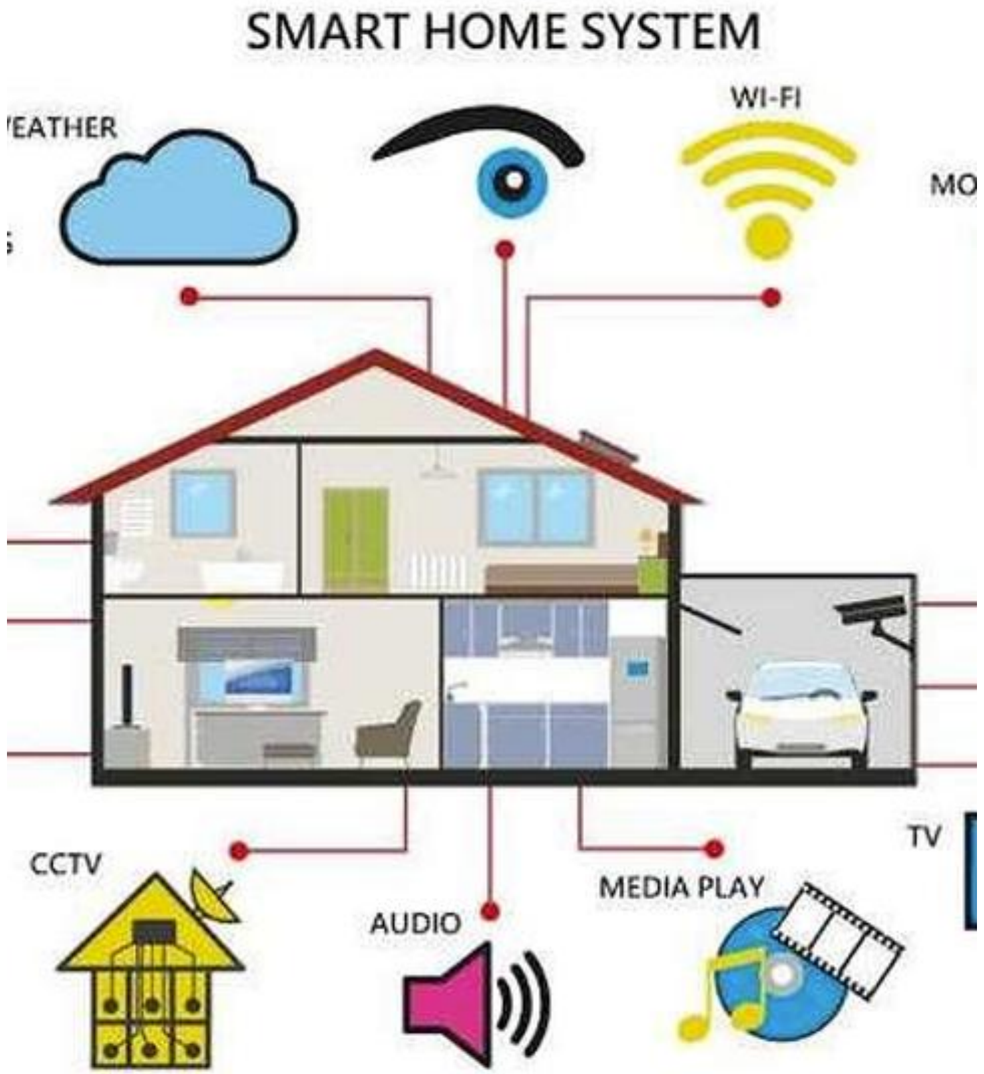

# Chatbots Advanced

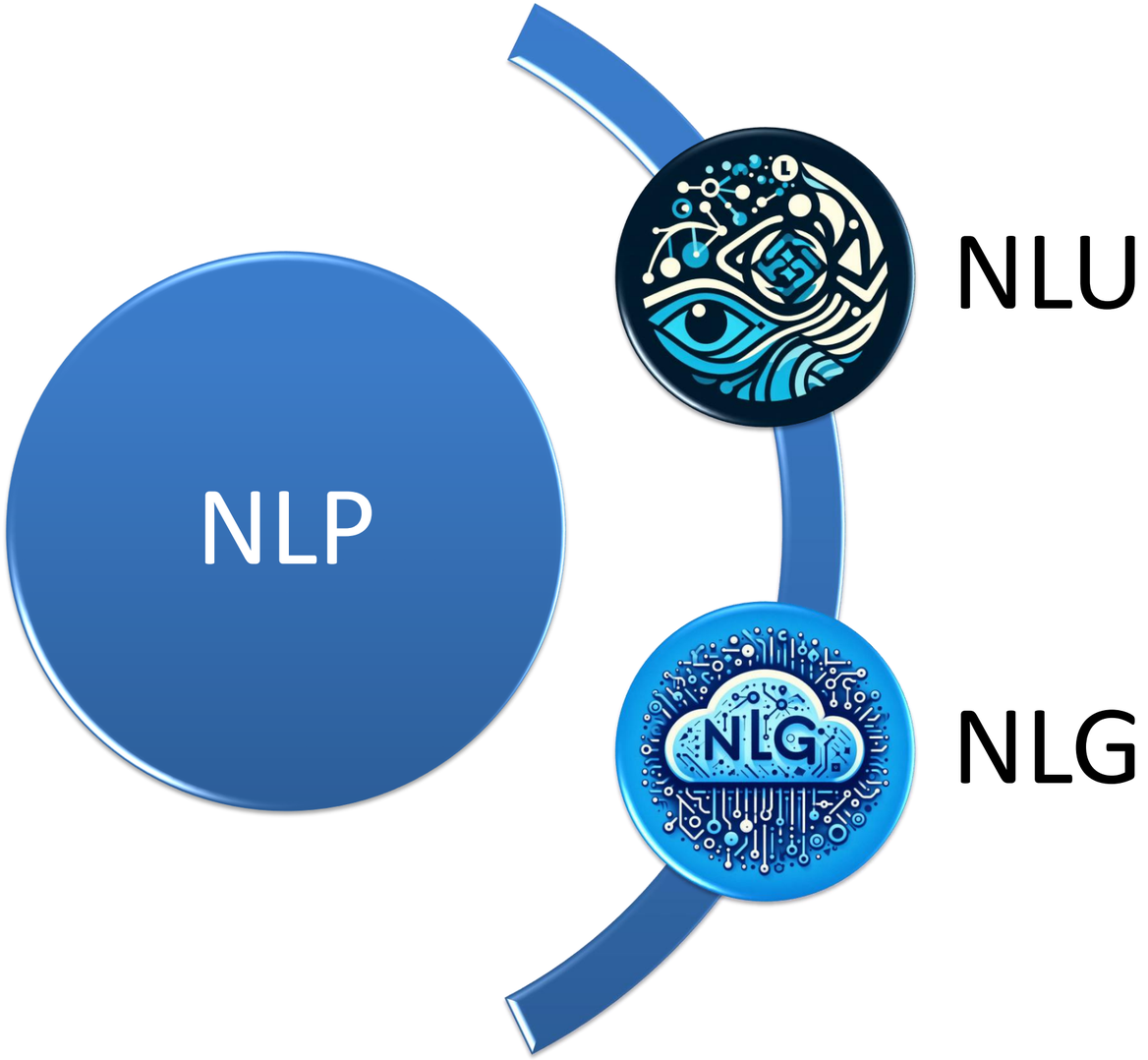
1. Facilitate Seamless Live Communication
2. Make Customer Service Available 24/7
3. Save Time and Money
4. Reduce People-to-People Interactions with Customers
5. Eliminate Tedious Time-Consuming Tasks
6. Offer a Smoother Customer Journey
7. Reduce Stress for Consumers
8. Eliminate Interactive Voice Response (IVR) Systems
9. Humanize Your Brand
10. Make Marketing More Targeted
11. Help Grow Your Business
12. Get Constant Improvement Over Time With Machine Learning

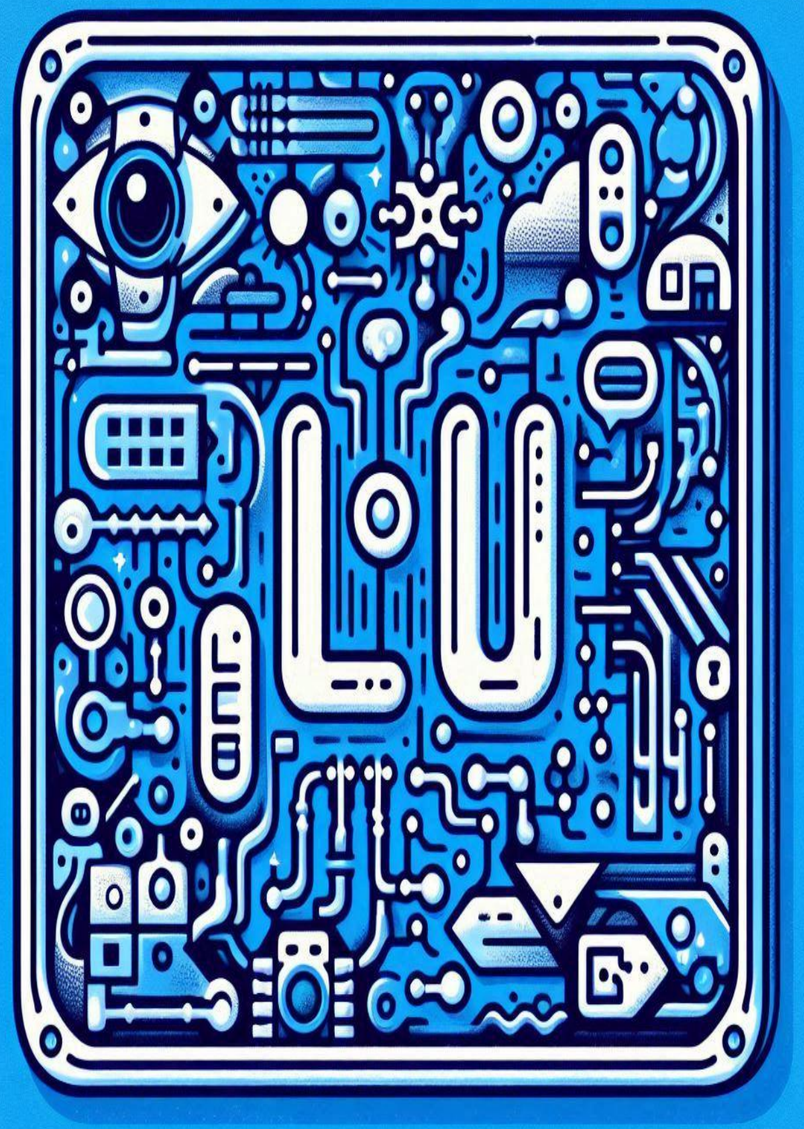
# NLP Applications:
# 8. Email Spam Filtering

# NLP Applications:
# 9. Voice Assistants and Smart Homes

- Amazon Alexa.
- •Google Assistant: “Okay Google!”
https://youtu.be/hIHsgqID 9Xc
https://youtu.be/ 81uyYKPpWww
- •Apple Siri.
- •Microsoft Cortana.
- •Samsung Bixby: “Hey Bixby!”
https://youtu.be/pu 9QNm5ITWg

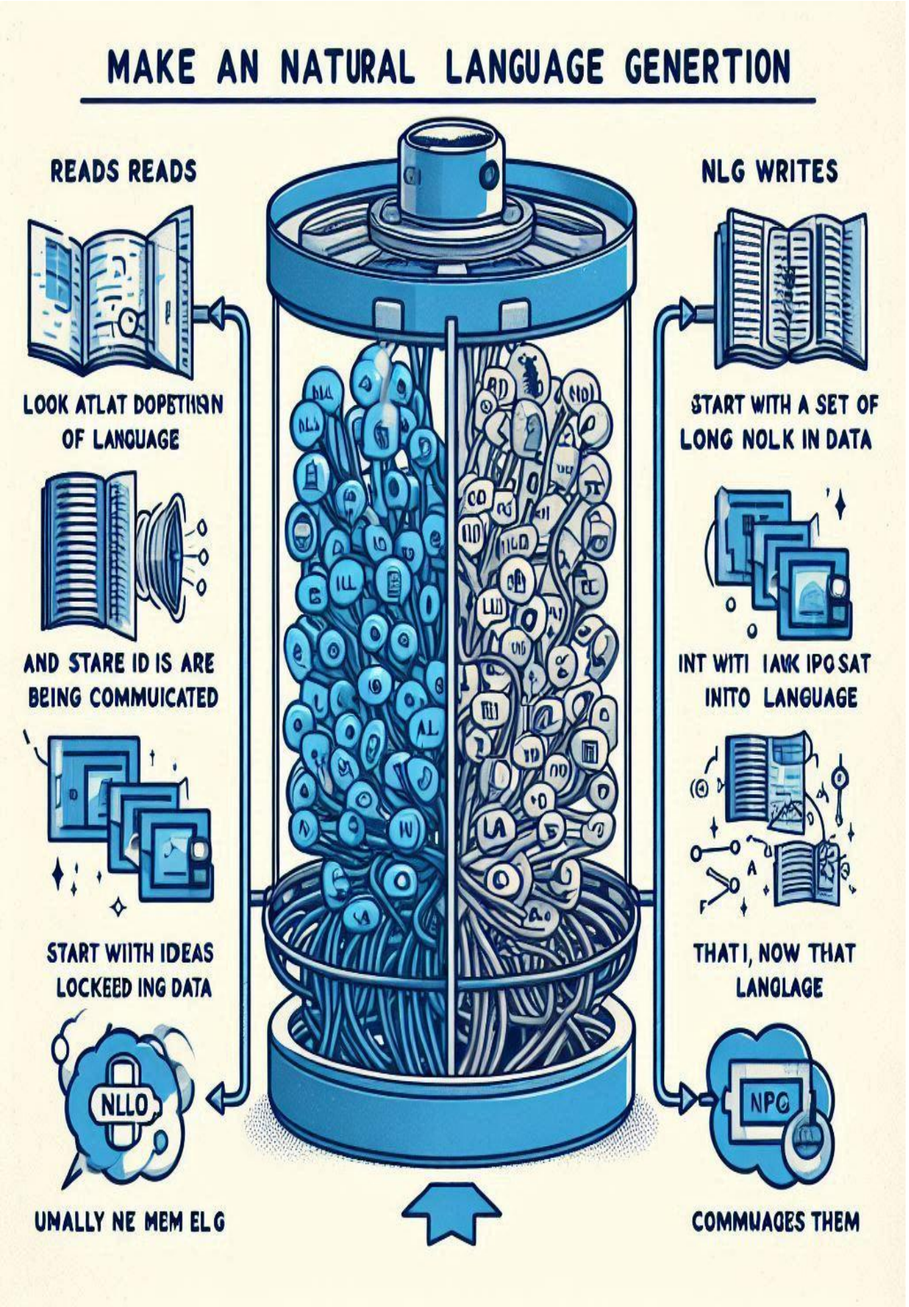



NLP
NLU
NLG



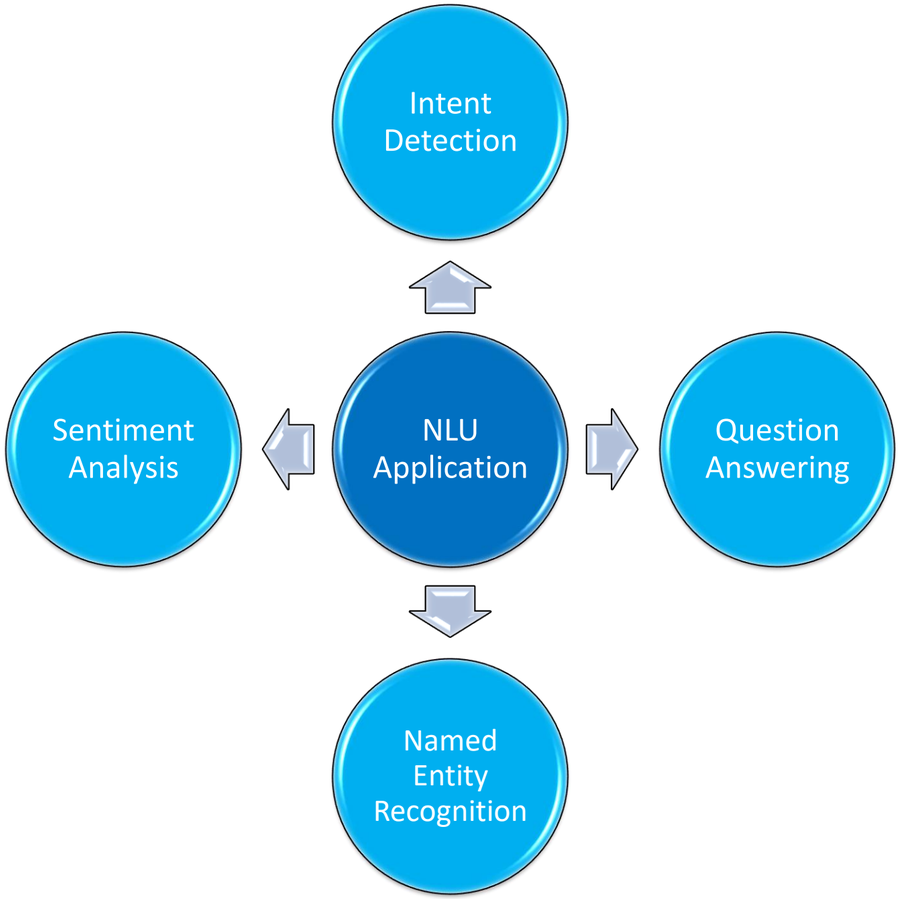
# Natural Language Understanding (NLU)

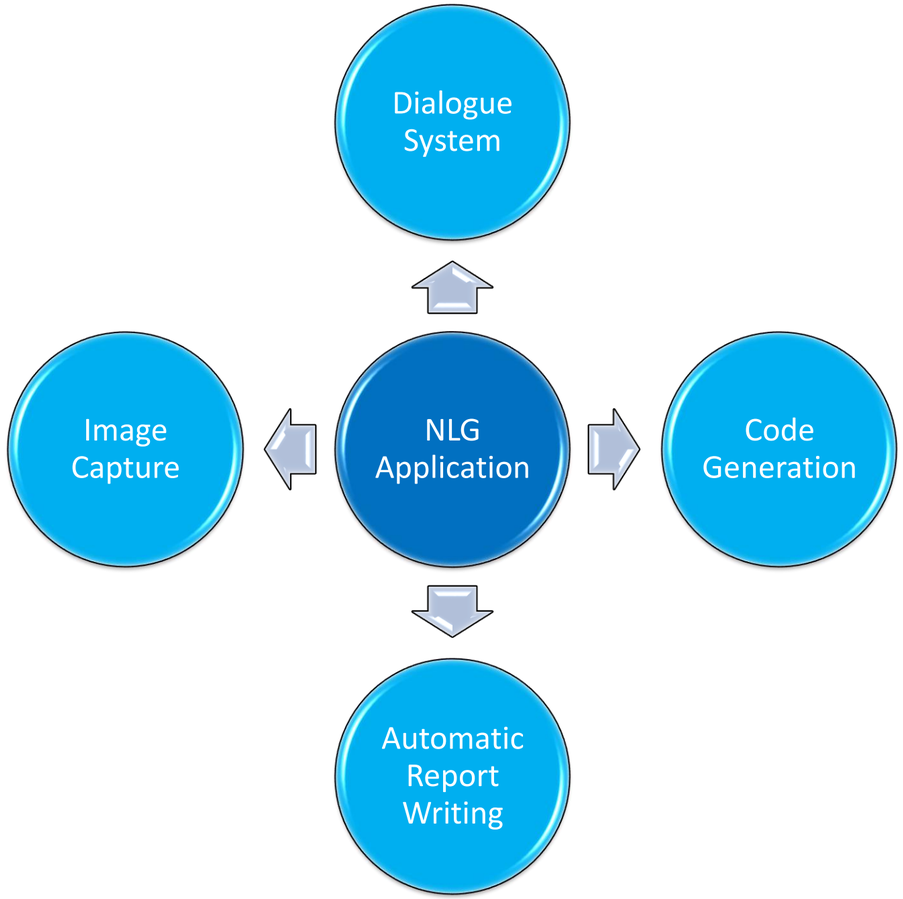
Natural language understanding (NLU) is a sub-branch of NLP and deals with these nuances via machine reading comprehension rather than simply understanding literal meanings. The aim of NLP and NLU is to help computers understand human language well enough that they can use language in a natural way.




# Natural Language Generation (NLG)

- NLG is the opposite direction of NLU.
- NLU reads while NLG writes.
- NLU systems look at language and figure out what ideas are being communicated.
- NLG systems start with a set of ideas locked in data and turn them into language that, in turn, communicates them
- Usually need NLU to perform NLG!


Intent Detection
Sentiment Analysis
NLU Application
Question Answering
Named Entity Recognition


Dialogue System
Image Capture
NLG Application
Code Generation
Automatic Report Writing

National Telecommunication Institute

# Challenges and Limitations of NLP

- Contextual
- Synonyms
- Irony and sarcasm
- Errors in text or speech
- Ambiguity
- Low-resource language

# Contextual Words and Phrases

Meaning depends on context

**Examples:**

- *He sat on the **bank** of the river.* vs *She went to the **bank** to open an account.*
- *The light is too **bright**.* vs *She has a **bright** future.*
- I **ran** to the store because we **ran** out of milk.
- May you **run** the program now?

# Synonyms & Paraphrasing

use many different words to express the same idea

**Examples:**

- *I'm very tired = I'm exhausted*
- *He bought a new vehicle = He purchased a new car*
- *I don't like it = I'm not a fan of it*
- *أنا مكتئب = أشعر بالحزن الشديد*

# Irony and Sarcasm

Words say one thing, mean the opposite.

**Examples:**

- Oh fantastic, the internet is down again!
- Great job! (after a clear failure)
- I just love getting stuck in traffic.
- فالح يا اخي

# Domain-specific language:

Different businesses and industries often use very different language

**Examples:**

- The patient has tachycardia.
- Deploy the container to Kubernetes.
- Use a sigmoid activation function
- There is problem in cloud

# Low-resource languages

Languages with limited data and models

**Examples:**

- Amazigh (limited NLP tools)
- Nubian dialects in Egypt and Sudan
- Arabic dialects (Sa'idi, Algerian) vs Modern Standard Arabic

# Ambiguity

Sentences or words with multiple meanings

**Examples:**

- The old man and woman left.
- I saw her duck
- Flying planes can be dangerous.

# Ambiguity

Sentences or words with multiple meanings

**Examples:**

- The old man and woman left.
- Flying planes can be dangerous.
- I saw the boy on the beach with my binoculars.
- قص الرجل الرواية

# Errors in text & speech

- U=you
- انشاء الله = ان شاء الله
- I havv too many erors

# Colloquialisms and idioms

- العين الحمراء
- Get cold feet
- dark horse

# Colloquialisms and Slang

- This movie is fire!
- وشك منور
- She ghosted me

# Lack of Research in Fields

- Sarcasm detection in Arabic tweets
- Legal document simplification
- Multilingual summarization for Arabic-English

# How does NLP really work?

# How does NLP work?

NLP commonly used terms:

- Phonemes: he smallest unit of sound to make a meaningful difference to a word; for example, the word cat contains three phonemes /k/-/a/-/t/;

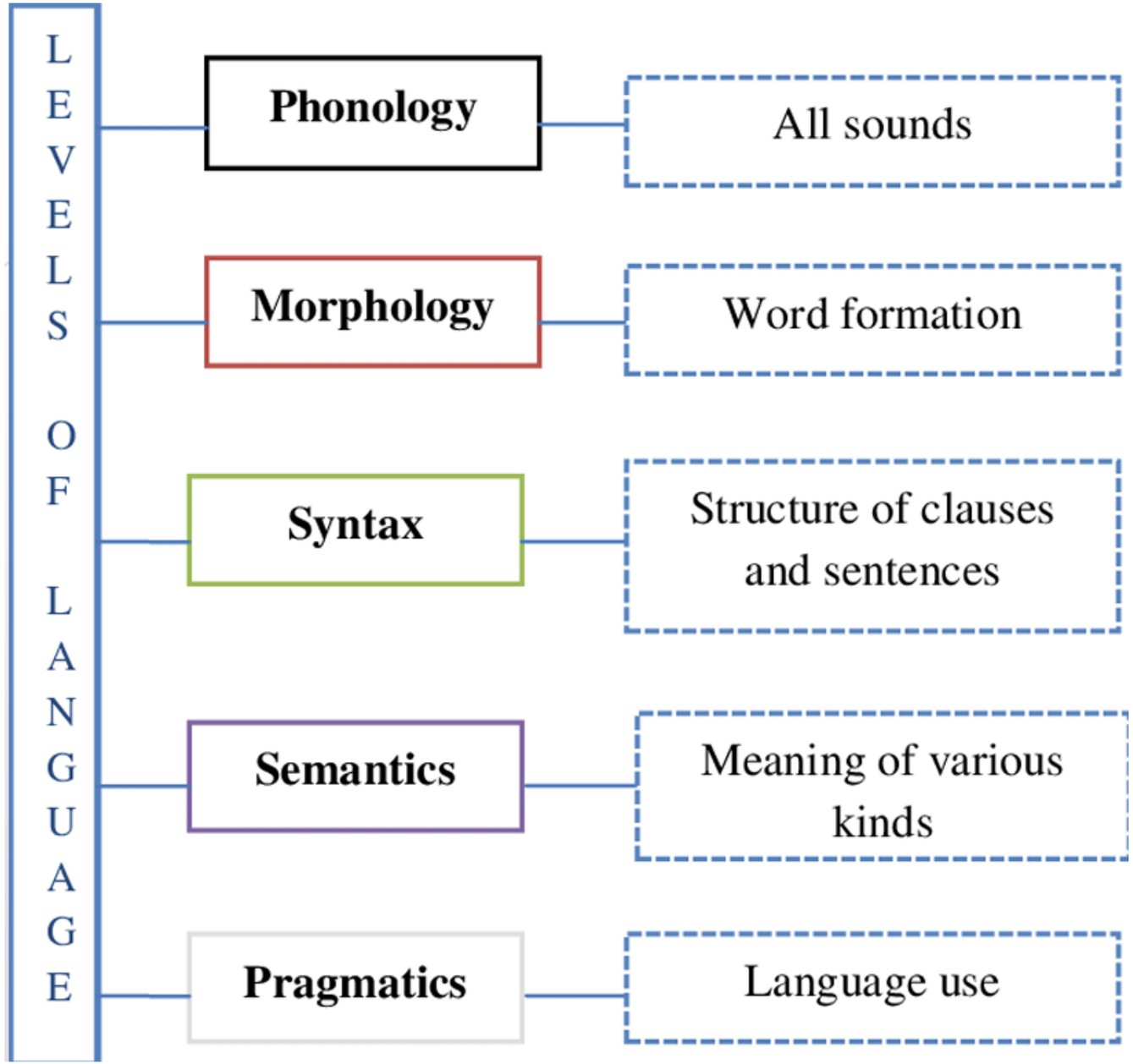
- Morphemes: morphemes are the basic units of meaning within words; for example, a free morpheme like cat is a word in its own right but bound morphemes like affixes (e.g.-er,-ing, un-) occur only in combination with a base (e.g. cooker).

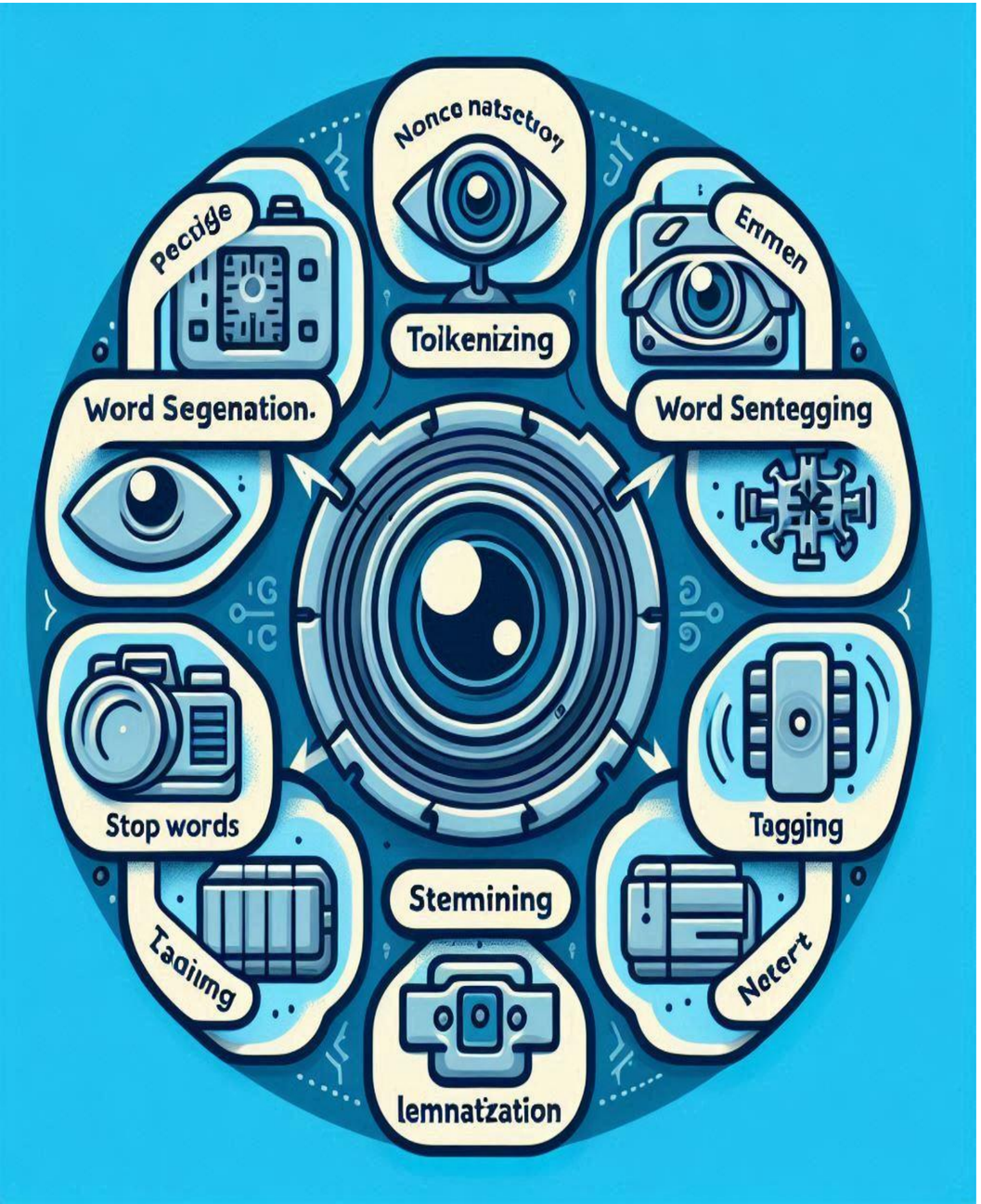
- Syntax: how words and sentences are constructed from these two building blocks. The most basic syntax follows a subject + verb + direct object formula. That is, "Jillian hit the ball."

- Semantics: the meaning of those words and sentences.

- Discourse: semantics in context. Conversation, persuasive writing, etc.

# Levels of NLP






# How does NLP work?

- The basic tasks in NLP are word level analysis including:
  - Tokenizing (also known as word segmentation)
  - Part-of-speech tagging (POS)
  - Named Entity Recognition (NER)
  - Stop Words Removal
  - Stemming
  - Lemmatization

- Syntax analysis or parsing is the process that follows to draw out exact meaning based on the structure of the sentence using the rules of formal grammar

# How does NLP work?

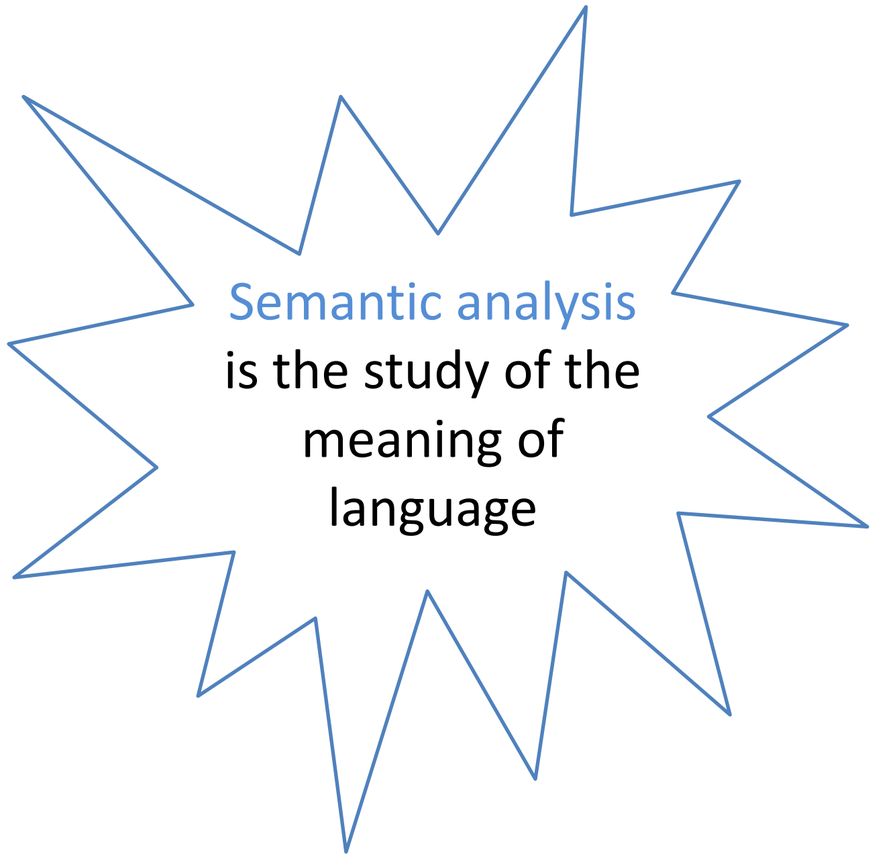
- Semantic analysis would help the computer learn about less literal meanings that go beyond the standard lexicon. This is often linked to sentiment analysis.

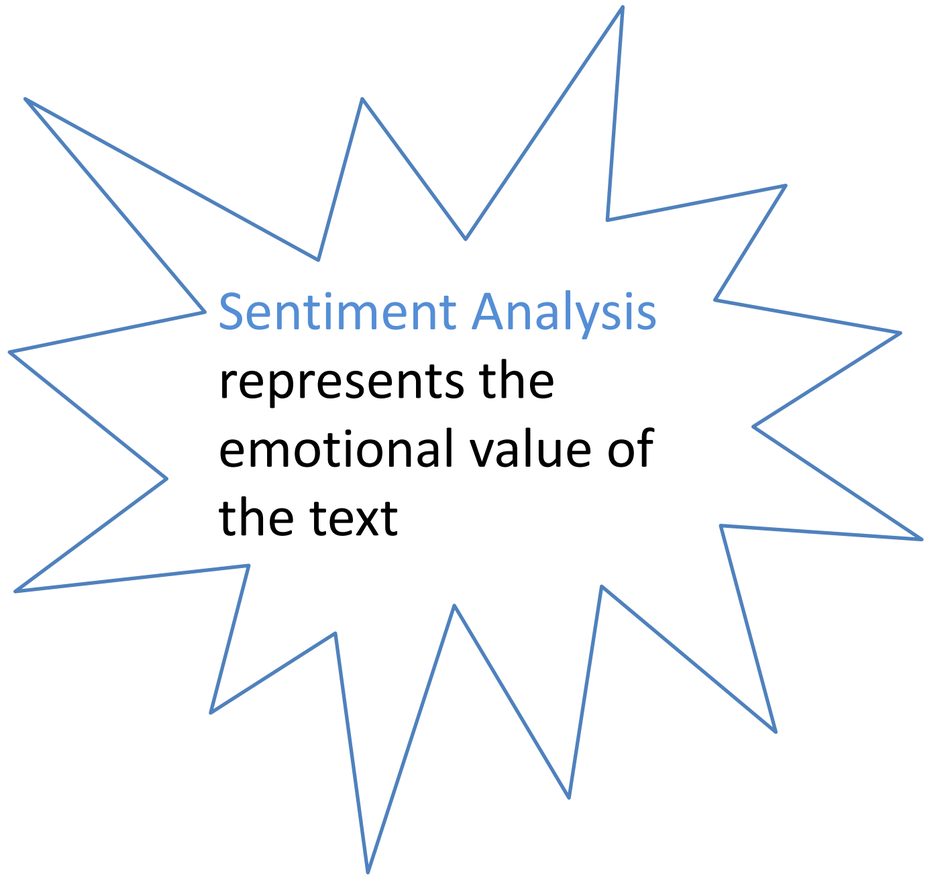
- Sentiment analysis is a way of measuring tone and intent in social media comments or reviews. It is often used on text data by businesses so that they can monitor their customers' feelings towards them and better understand customer needs.

  - Simple emotion detection systems use lexicons– lists of words and the emotions they convey from positive to negative. More advanced systems use complex machine learning algorithms for accuracy.

  - This is because lexicons may class a word like "killing" as negative and so wouldn't recognize the positive connotations from a phrase like, "you guys are killing it" (someone is doing very, very well)

# Semantic Vs. Sentiment Analysis



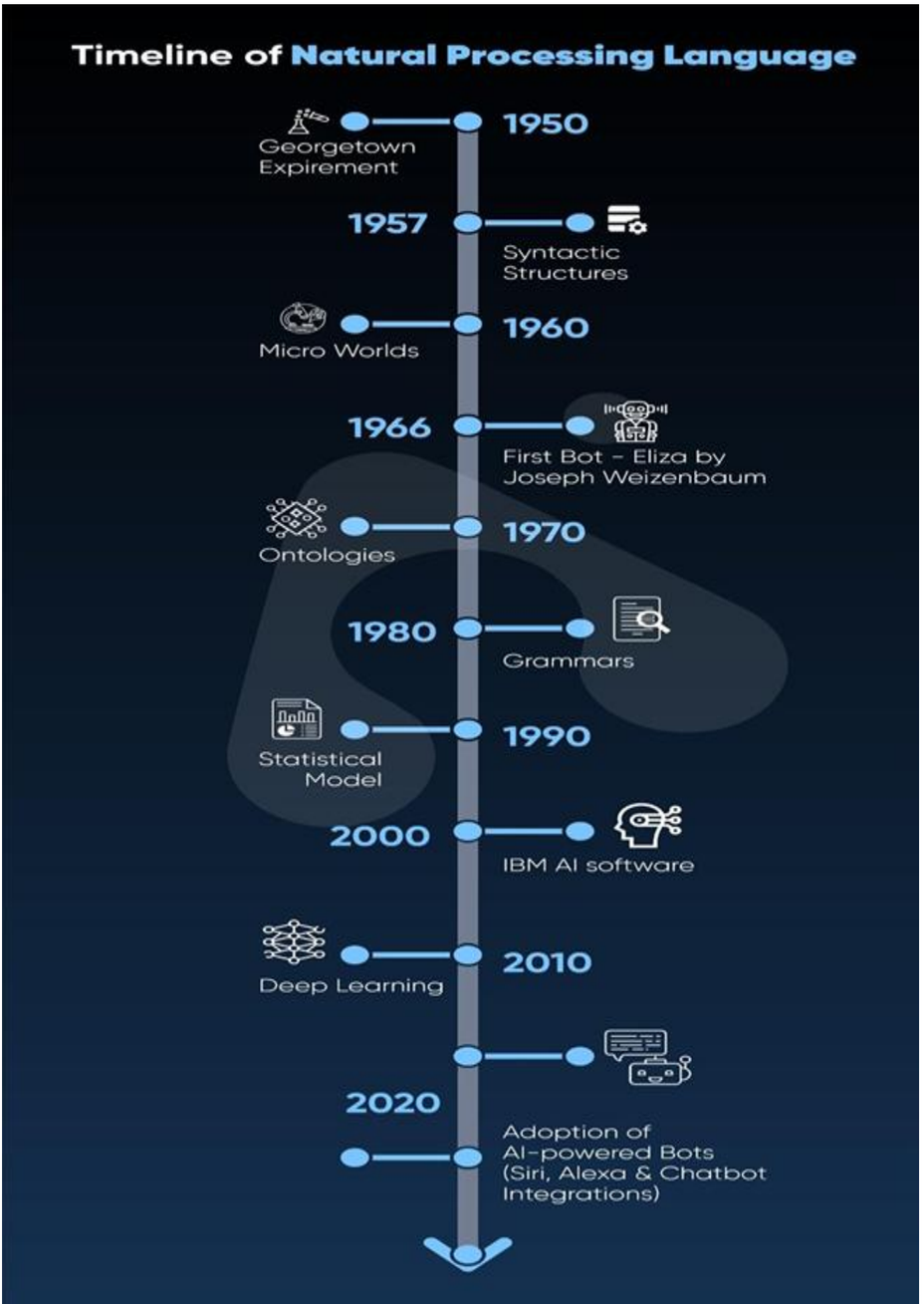



# How does NLP work?

- Word sense disambiguation (WSD) is used to ascertain which sense of a word is being used in a sentence.

Example of different context using the same word;

- “The bank will not be accepting cash on Saturdays. ”
- “The river overflowed the bank.”




# NLP Evolution

- In **1950**, **Alan Turing** published his famous paper *"Computing Machinery and Intelligence"*, in which he introduced what is now known as the **Turing Test**.

•In **1954**, the **Georgetown Experiment** was conducted in collaboration with **IBM**. It showcased one of the **first fully automatic machine translation systems**, translating **more than 60 sentences** from **Russian to English**.

Person: Pick up a big red block.

Computer: OK.

Person: Grasp the pyramid.

Computer: I DON'T UNDERSTAND WHICH PYRAMID YOU MEAN.

Person (changing their mind): Find a block which is taller than the one you are holding and put it into the box.

Computer: BY "IT", I ASSUME YOU MEAN THE BLOCK WHICH IS TALLER THAN THE ONE I AM HOLDING.

Computer: OK.

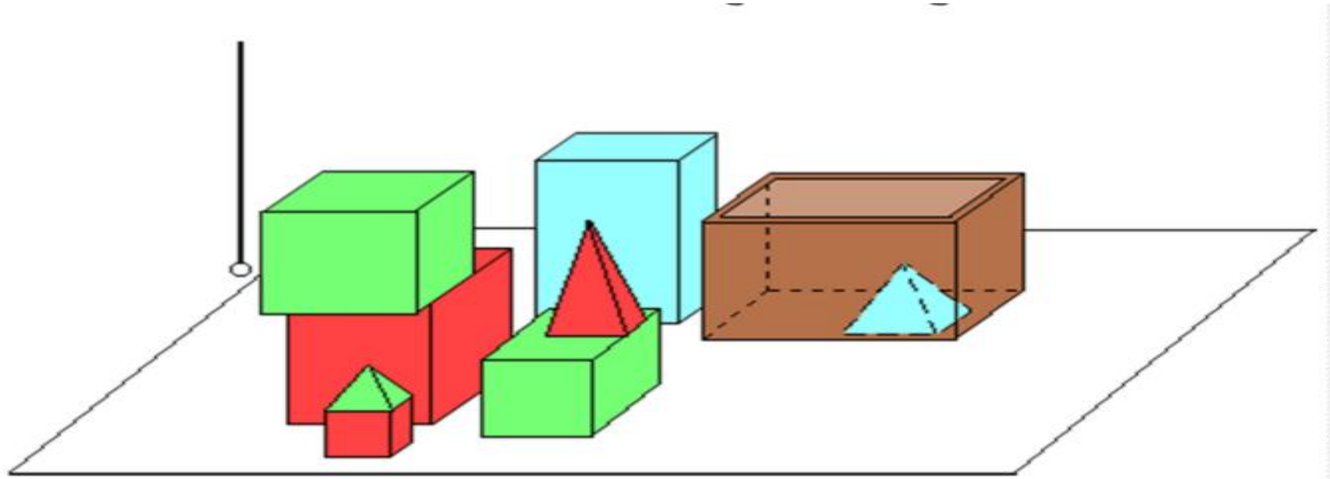
Person: What does the box contain?

Computer: THE BLUE PYRAMID AND THE BLUE BLOCK.

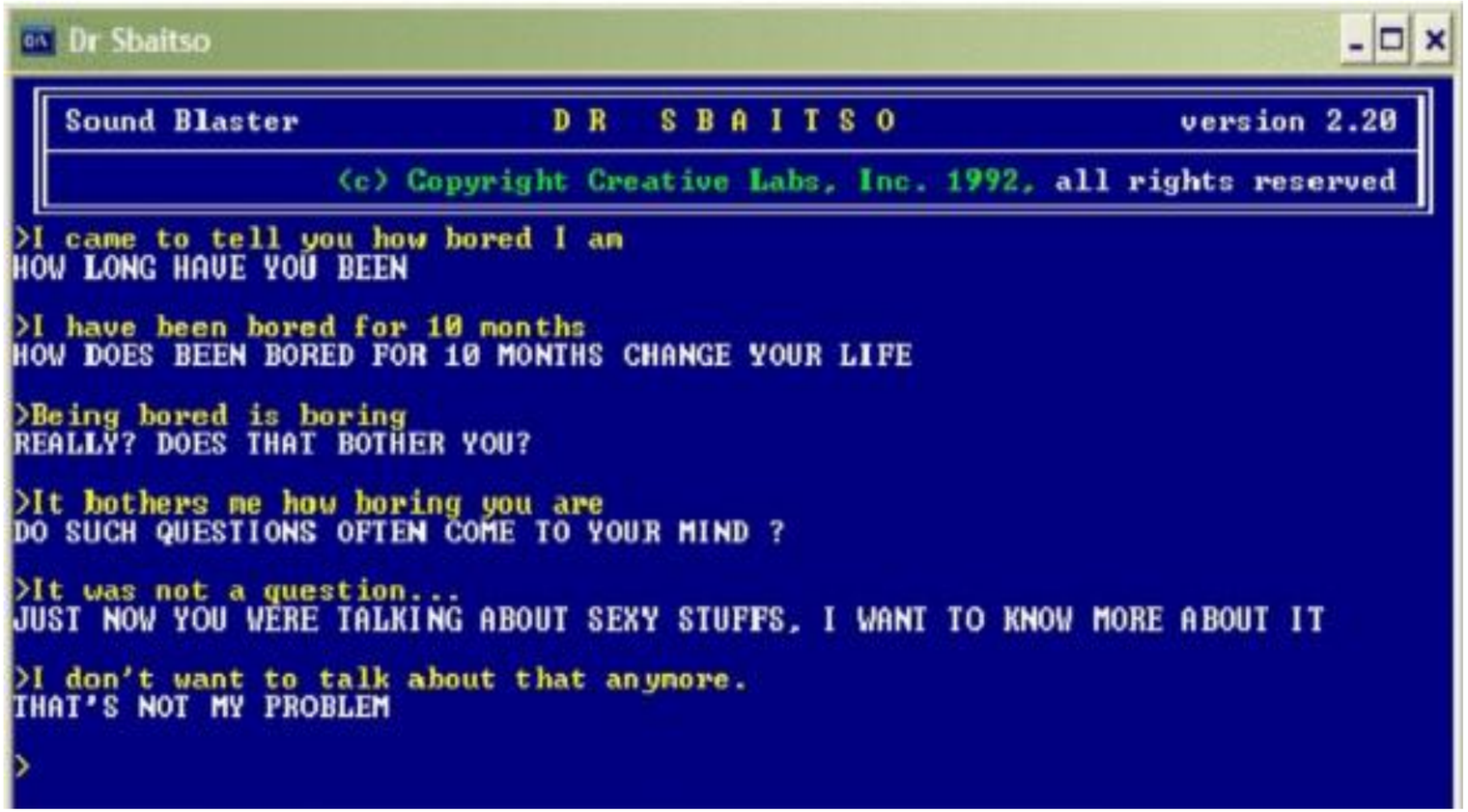
Person: What is the pyramid supported by?

Computer: THE BOX.

Person: How many blocks are not in



- In **1968**, American computer scientist **Terry Winograd** at **MIT** developed a pioneering program called **SHRDLU**.
- SHRDLU operated in a **simple virtual world (called the blocks world)** and was able to understand and respond to **natural language commands**.
- It successfully engaged in **context-aware dialogue** with users, demonstrating a degree of **politeness, coherence, and reasoning**.




## Doctor Sbaitso – Early AI Psychologist Chatbot (1991)

- In **1991**, a program called **Doctor Sbaitso** was developed as an early example of an **AI-based psychologist chatbot**.
- It was designed to simulate a **conversation with a therapist**, running on **MS-DOS systems**.
- While basic in intelligence, **Doctor Sbaitso** represented an early attempt to bring **interactive NLP** to personal computers and inspired future virtual assistants.


# Evolution of Virtual Personal Assistants (2011–2017)

- In 2011, Apple launched Siri, the first widely adopted voice-enabled personal assistant for smartphones.
- In 2014, both Amazon Alexa and Microsoft Cortana were introduced, expanding the use of AI assistants to smart homes and Windows devices.
- In 2016, Google Assistant was launched, offering more advanced contextual understanding and integration with the Google ecosystem.
- In 2017, Samsung introduced Bixby, designed to provide device control and voice interaction across Samsung products.

# NLP Pipeline

- Sentence Segmentation
- Word Tokenization
- Stemming
- Lemmatization
- Identifying Stop Words
- Dependency Parsing
- POS tags
- Named Entity Recognition(NER)
- Chunking


- Extract hashtags from a tweet.
- Clean HTML tags from a string.
- Validate password strength (e.g., at least one uppercase, lowercase, digit, and symbol).
- Split a sentence by punctuation.

National Telecommunication Institute

# Thank You For Attention