NEW YORK NEWS

Articles for you

Natural Language Processing

Language Processing

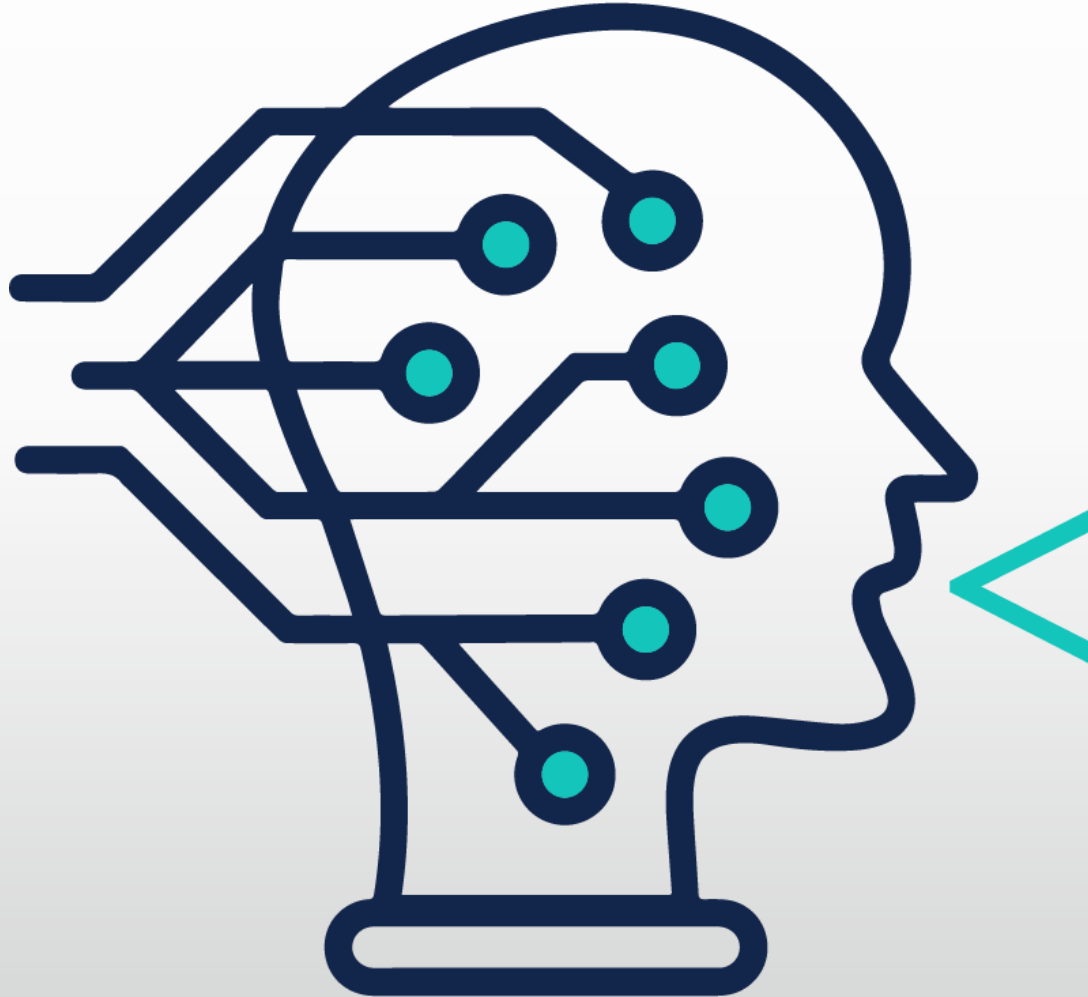Ability to understand Language

Machine Instructions

High Level Programming Language

Highly Structured and Error Free

Natural Language

✓ A natural language or ordinary language is any language that has evolved naturally in humans through use and repetition without conscious planning or premeditation. [source:Wikipedia]

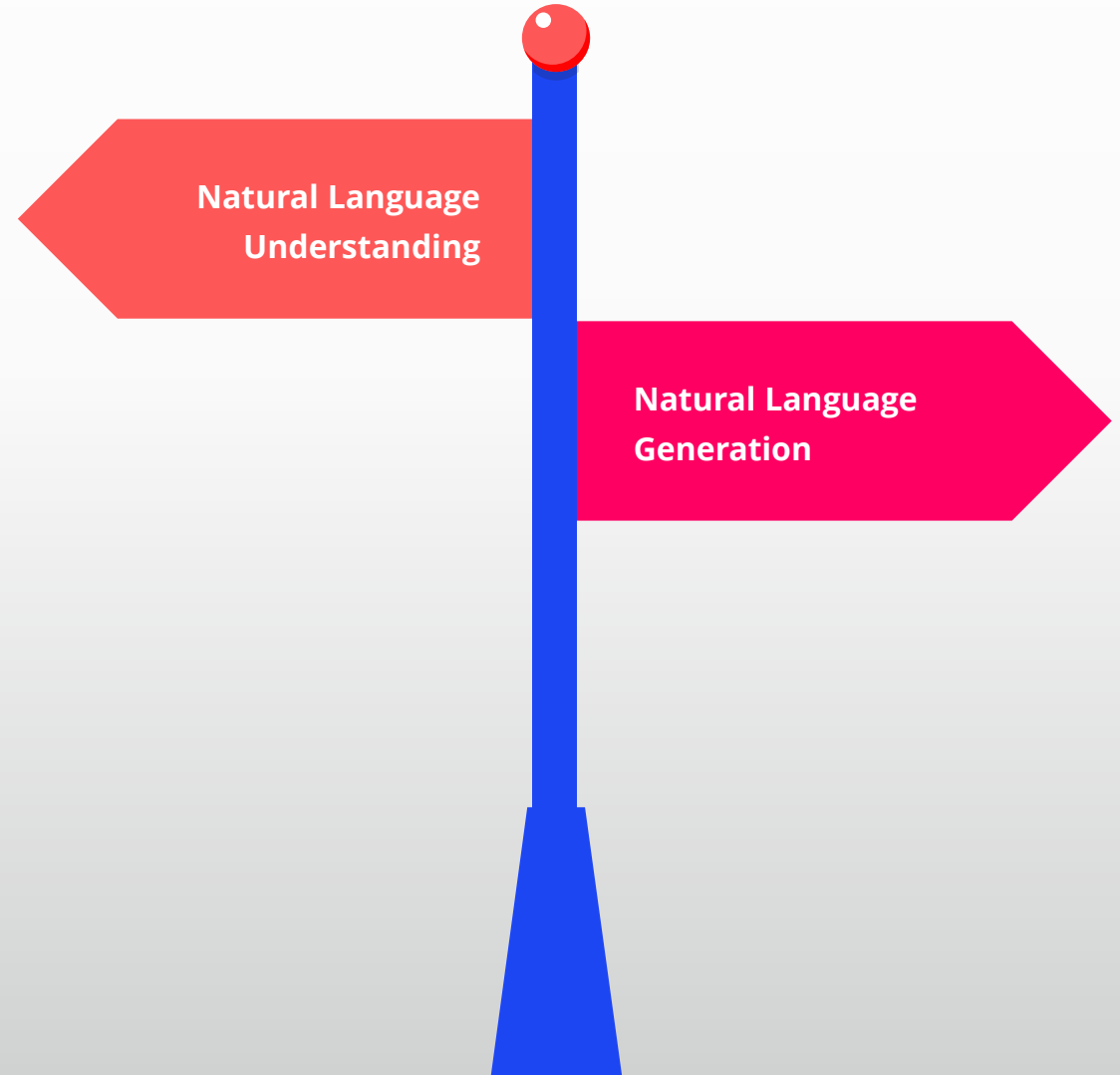✓ Natural languages can take different forms, such as speech or signing.

voodoo

Natural Language Issues

✓ **Human Language Vocabulary is very large.**

✓ **Words have multiple meaning attached to it.**

✓ **People speak with varied accents.**

✓ **Word Plays like comma, full stop, exclamation mark!!!**

 e.g. *I didn't say he stole the money.*

✓ **Mispronunciation of words.**

✓ **Missing out key details in conversations.**

01

Vocab

02

Accent

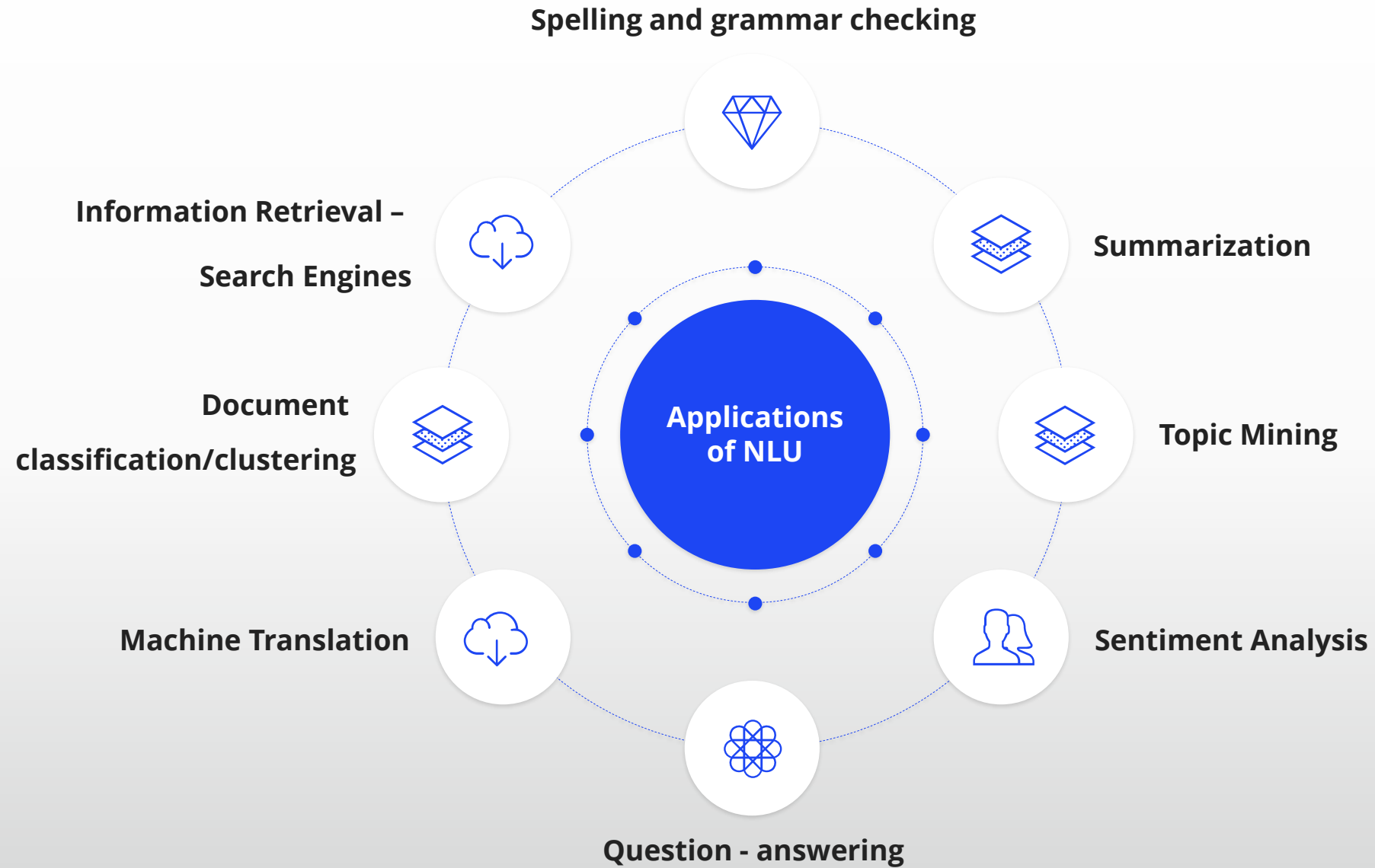**Conversations**

03

voodoo
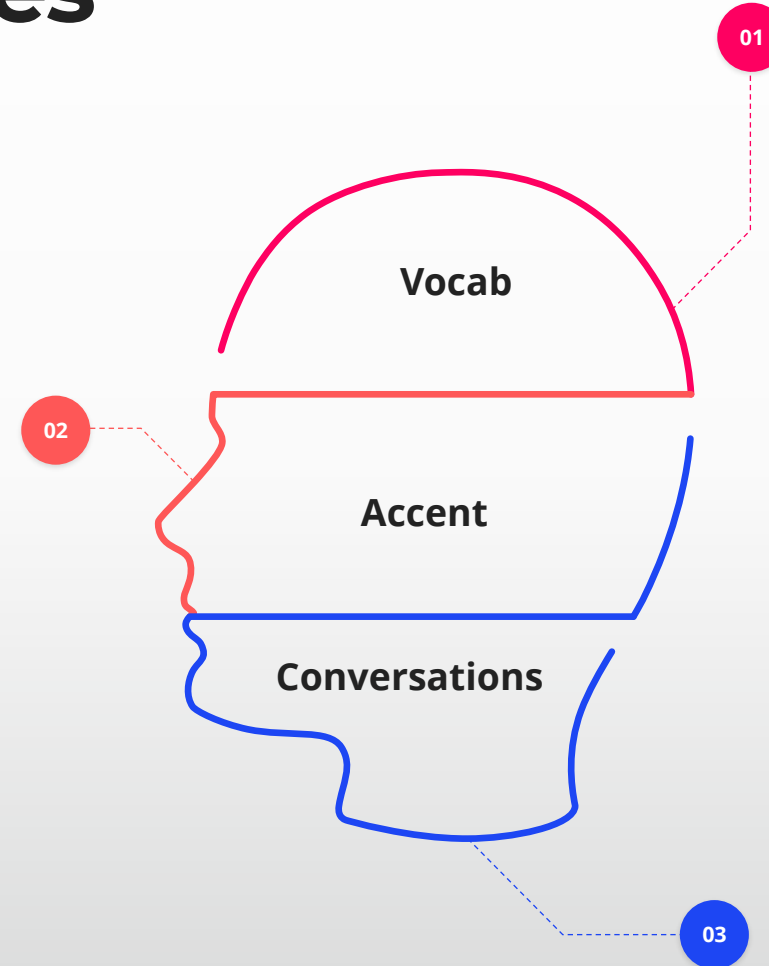
Natural Language Processing

NLP is a branch of data science that consists of systematic processes for **analysing, understanding, and deriving information** from the text data in a smart and efficient manner.

Natural Language Understanding

Natural Language Generation

5

Natural Language Understanding

Spelling and grammar checking

Summarization

Topic Mining

Sentiment Analysis

Question - answering

Machine Translation

Document classification/clustering

Information Retrieval – Search Engines

Applications of NLU

voodoo

Basic Preprocessing Pipeline for NLP

Sentence Segmentation

Hello world. This blog post is about sentence segmentation. It is not always easy to determine the end of a sentence. One difficulty of segmentation is periods that do not mark the end of a sentence. An ex. is abbreviations.

- Hello world.
- This blog post is about sentence segmentation.
- It is not always easy to determine the end of a sentence.
- One difficulty of segmentation is periods that do not mark the end of a sentence.
- An ex. is abbreviations.

Text
"The cat sat on the mat."

↓

Tokens
"the", "cat", "sat", "on", "the", "mat", "."

White Space Tokenization

Dictionary based Tokenization

Rule based Tokenization

Penn Tree Tokenization

Spacy Tokenizer

Moses Tokenizer

Tokenization

Subword Tokenization

Byte Pair Encoding

Word Piece Encoding

Sentence Piece Encoding

Unigram language model

PARTS OF SPEECH

Noun
A noun is the name of a person, place, or thing.

Person
girl
boy
teacher

Place
school
home
store

Thing
pencil
jacket
dog

Adjective
An adjective describes a noun or a pronoun. An adjective tells what kind, how many, or which one.

What Kind
happy
brave

How Many
more
two

Which One
this
that

Verb
A verb can tell what action someone or something is doing. A verb can also express a state of being.

Action
run   jump
sit   ask
think   talk

State of Being
am   is
are   was
were

Pronoun
A pronoun is used in place of a noun in a sentence. A pronoun may take the place of the name of a person, place, or thing.

I   she   it   they
you   he   we   me

Article
The words *a*, *an*, and *the* belong to a special group of adjectives called articles. An article can be used before a noun in a sentence.

a   an   the
a dog   an apple   the boy
a rabbit   an ant   the bird

Adverb
An adverb describes a verb, adjective, or another adverb. An adverb tells how, when, where, or to what degree.

How
quickly

When
today

Where
outside

To What Degree
barely

Preposition
A preposition combines with a noun or pronoun to form a phrase that tells something about another word in a sentence.

from   to   until
over   with   after

Conjunction
A conjunction joins together single words or groups of words in a sentence.

and   but   or   nor

Interjection
An interjection expresses strong feeling or emotion. An interjection can be a single word or a phrase.

Help!   Oh!   Ouch!
Ugh!   Whew!   Ah!
Wow!   Look out!   Oh dear!

The boy yelled, "Help!" and he saw a brave dog quickly run to the rescue.

To solve this ambiguity computer need to learn grammar.

That can be done through phrase structure grammar

Rules in English

PHRASE STRUCTURE RULES

[SENTENCE] ⟶ [NOUN PHRASE] [VERB PHRASE]

[NOUN PHRASE] ⟶ [ARTICLE] [NOUN]

[NOUN PHRASE] ⟶ [ADJECTIVE] [NOUN]

[NOUN PHRASE] ⟶ [NOUN]

[VERB PHRASE] ⟶ [VERB]

[VERB PHRASE] ⟶ [VERB] [NOUN PHRASE]

[VERB PHRASE] ⟶ [VERB] [PREPOSITIONAL PHRASE]

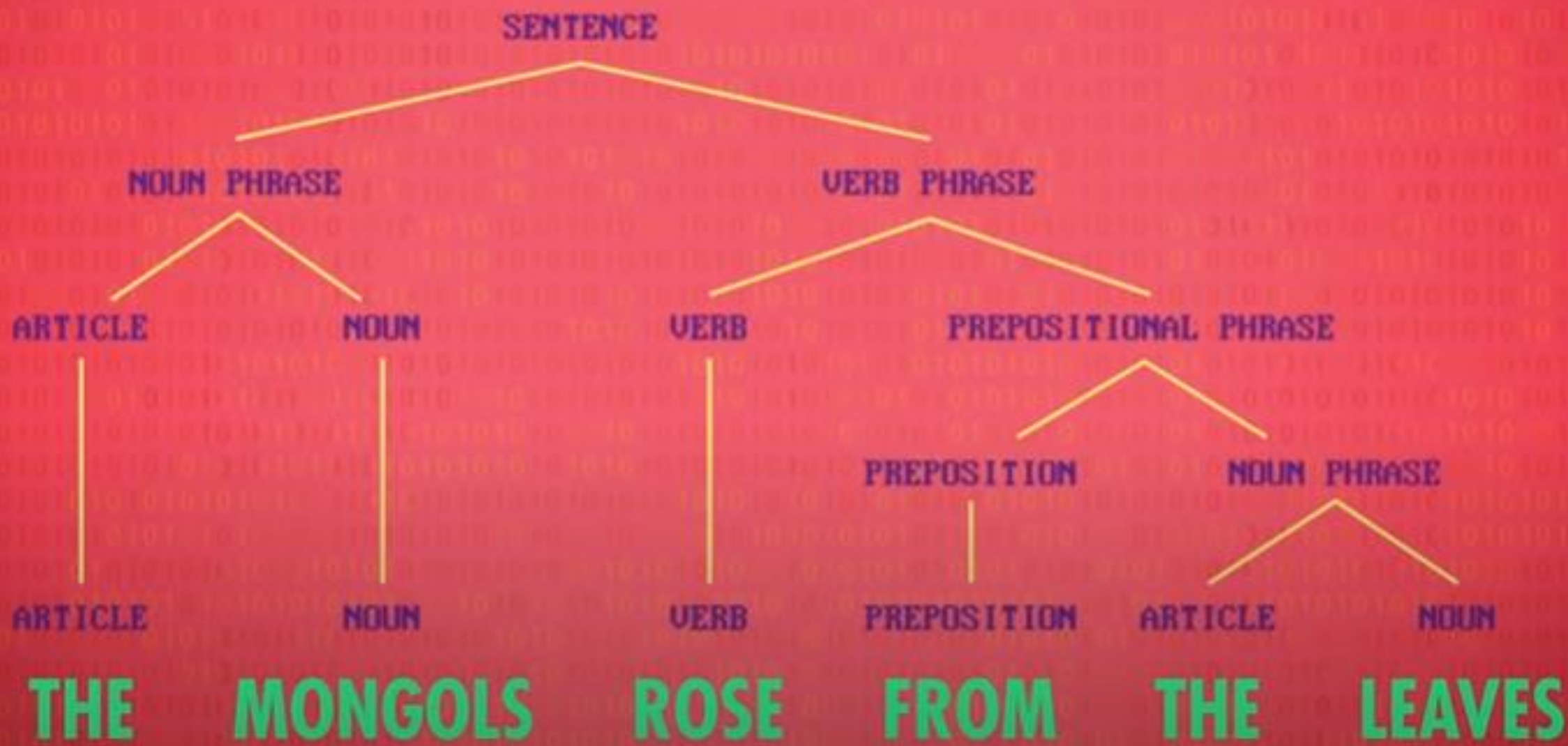[VERB PHRASE] ⟶ [VERB] [NOUN PHRASE] [PREPOSITIONAL PHRASE]

[VERB PHRASE] ⟶ [VERB] [NOUN PHRASE] [ADVERB]

[PREPOSITIONAL PHRASE]   [PREPOSITION] [NOUN PHRASE]

...

PARSE TREE

SENTENCE

NOUN PHRASE · VERB PHRASE

ARTICLE · NOUN · VERB · PREPOSITIONAL PHRASE

PREPOSITION · NOUN PHRASE

ARTICLE · NOUN · VERB · PREPOSITION · ARTICLE · NOUN

THE MONGOLS ROSE FROM THE LEAVES

Sentence Segmentation

Tokenization

Part-of-Speech Tagging

Lemmatization

Stop Words

Dependency Parsing

Noun Phrase

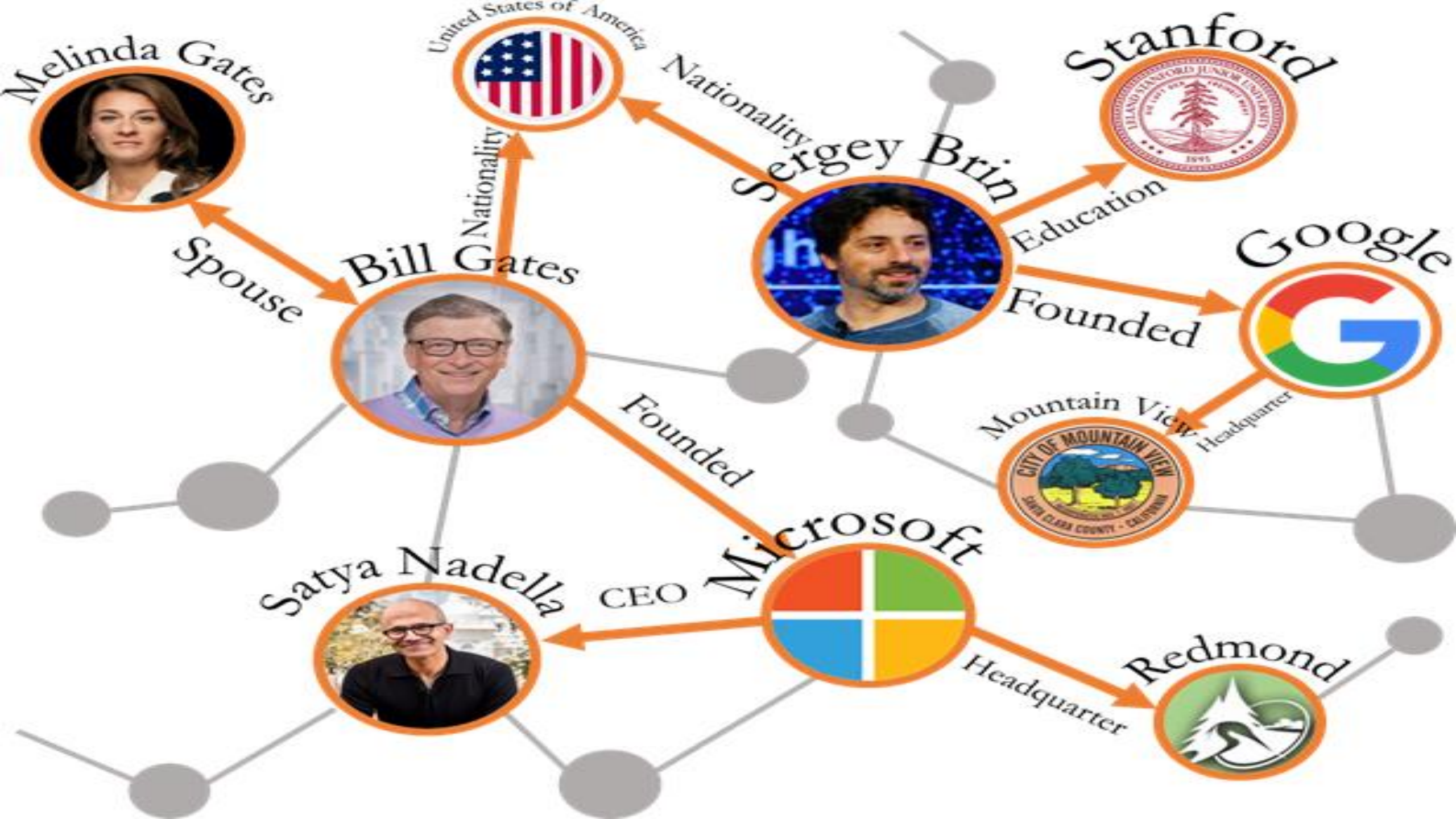**Named Entity Recognition**

Conference Resolution

✓ **Hello alexa. Where is taj mahal located?**

✓ Hello alexa.

✓ **Where is the taj mahal located?**

✓ **Where | is | taj | Mahal | located**

✓ **Where-** interrogative adverb **| is - verb | taj - noun | Mahal - noun | located – verb**

✓ **where | taj | Mahal | locate**

✓ **Named Entity – Taj Mahal**

I hear Berlin [Place] is wonderful in the winter [Time]

The nurse notified the patient that his shift would be ending in an hour.

This works particularly well when data is stored in a web of semantic information, where entities are linked to one another in meaningful relationships, providing all the ingredients you need to craft informational sentences.
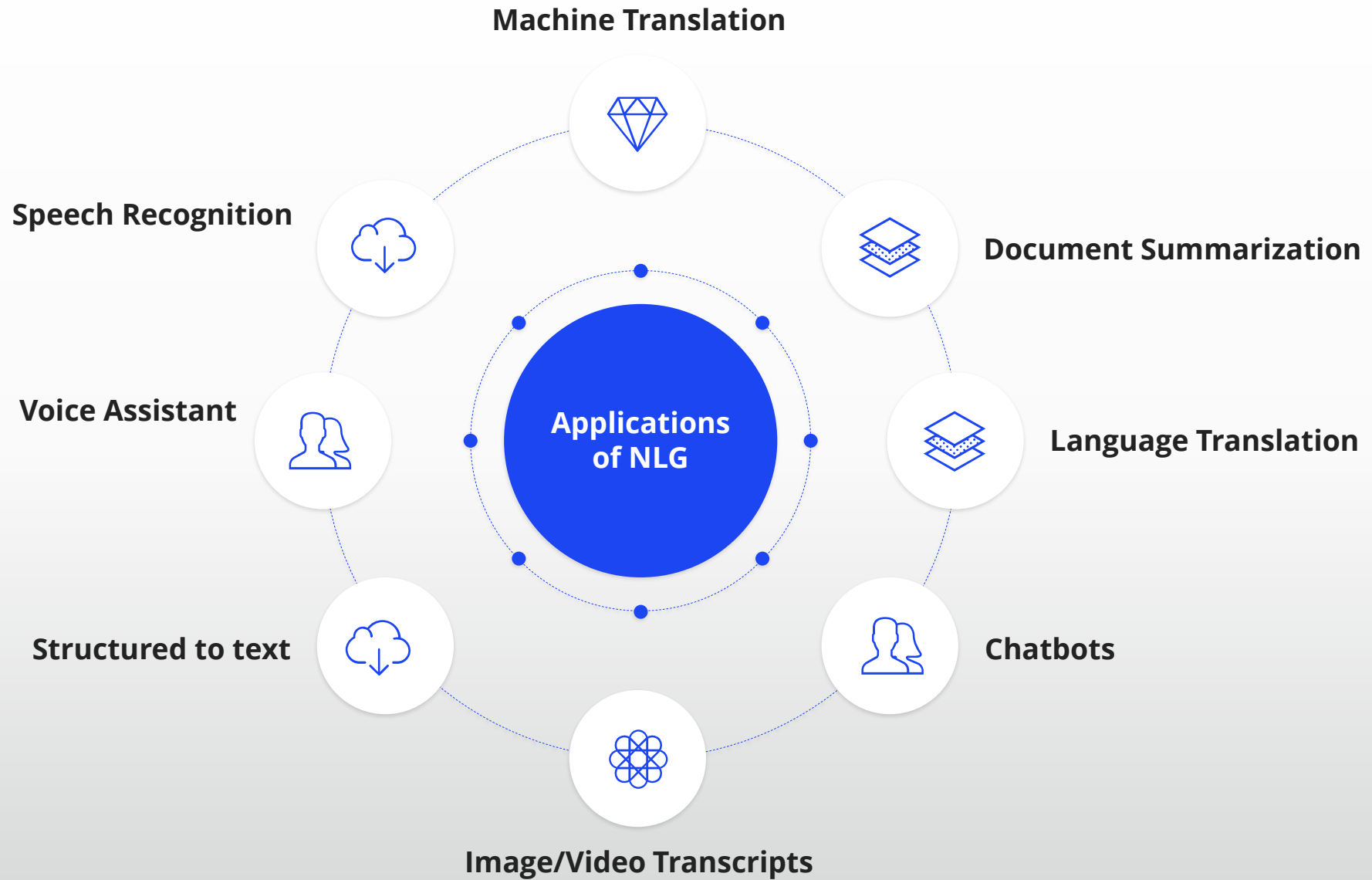
Natural Language Generation

We were driving on a dirt road up a hill when a motorcycle suddenly hit us. The rider was hit with the motorcycle and the rider did not have what would have been considered fatal injuries or serious damage to his body. The doctor told the man to let go of the bike. He had taken off his helmet and the doctor said that that is when he was supposed to be stopped. The doctor said that what had happened was not right and he was told that the rider must have suffered serious injuries to his head, back and arms. The doctor went on and told the man that you couldn't go on to the street. He had to get out of your car to get out of the truck. He pulled out a knife and stabbed in the side of his head. He had to crawl back out of the truck and get out of there. The doctor asked for him to put his hand up to his head, but said no. The man said that he was going to a church or a church and the doctor put on his headband. He said that the helmet was already attached and he was going to be out there with the knife and knife. The doctor then went into the car and stopped him. The doctor said that he had to put a piece of cardboard on top of the knife just to get it out and he was going to be put into the van. The man said that, at first, he thought that he woul...

We were driving on a dirt road up a hill when a motorcycle suddenly hit us. The rider was hit with the motorcycle and the rider did not have what would have been considered fatal injuries or serious damage to his body. The doctor told the man to let go of the bike. He had taken off his helmet and the doctor said that that is when he was supposed to be stopped. The doctor said that what had happened was not right and he was told that the rider must have suffered serious injuries to his head, back and arms. The doctor went on and told the man that you couldn't go on to the street. He had to get out of your car to get out of the truck. He pulled out a knife and stabbed in the side of his head. He had to crawl back out of the truck and get out of there. The doctor asked for him to put his hand up to his head, but said no. The man said that he was going to a church or a church and the doctor put on his headband. He said that the helmet was already attached and he was going to be out there with the knife and knife. The doctor then went into the car and stopped him. The doctor said that he had to put a piece of cardboard on top of the knife just to get it out and he was going to be put into the van. The man said that, at first, he thought that he woul...
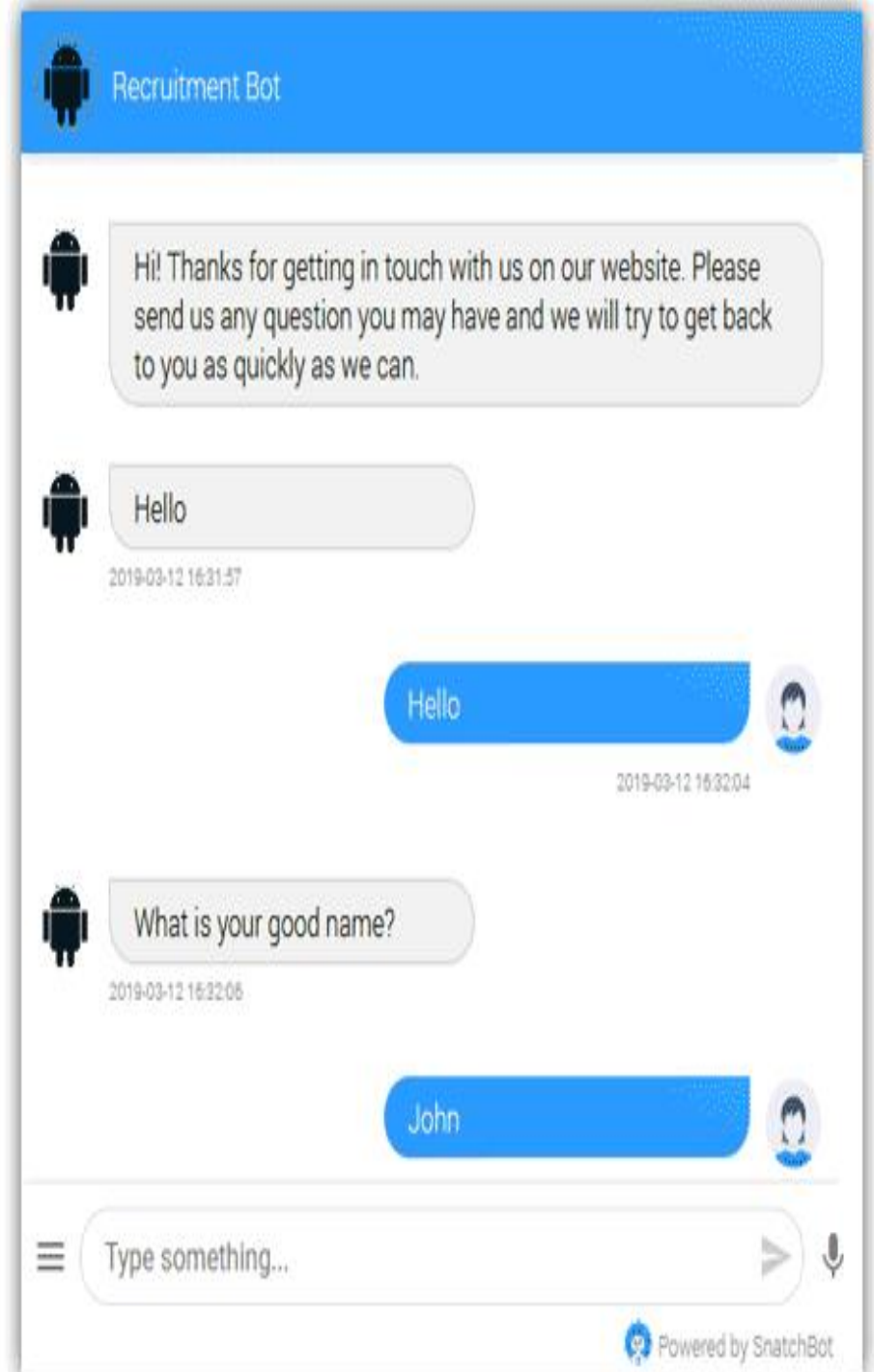
voodoo

voodoo

Machine Translation

Speech Recognition

Document Summarization

Voice Assistant

Applications of NLG

Language Translation

Structured to text

Chatbots

Image/Video Transcripts

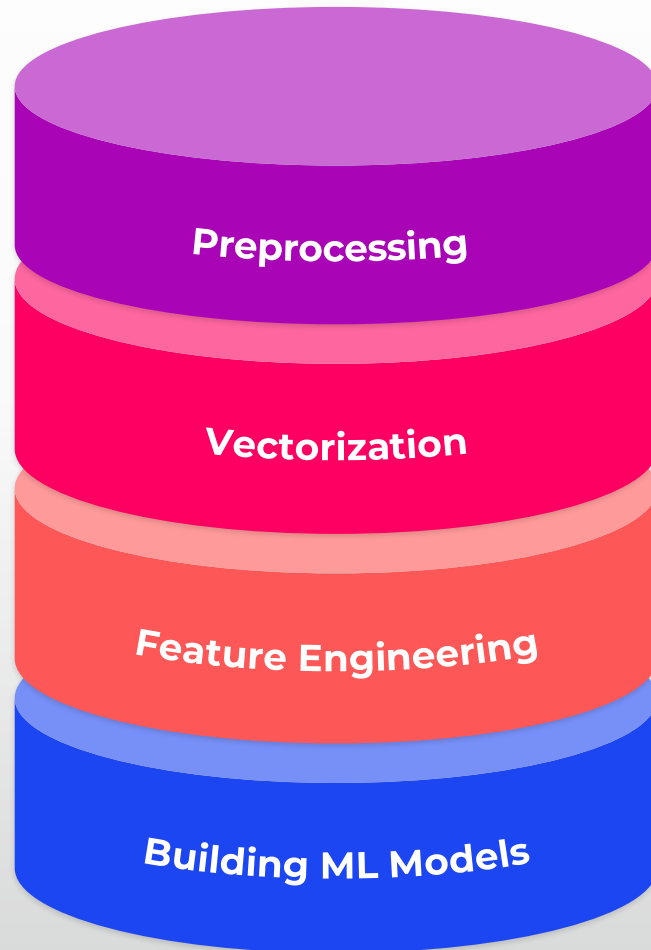**Talking back!**

What to say or text planning

➤ **flight(AA,london,boston,$560,2pm),**

➤ **flight(BA,london,boston,$640,10am),**

How to say it

"There are two flights from London to Boston. The first one is with American Airlines, leaves at 2 pm, and costs $560 …"
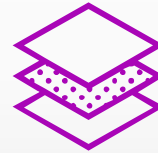
Steps in Natural Language Processing

voodoo

Preprocessing

Vectorization

Feature Engineering

Building ML Models

Preprocessing – Removing stop words, tokenization, Stemming

Vectorizing Data: Bag-Of-Words, N-Grams, TF-IDF

Feature Engineering – word Embedding, Name-Entity

ML Models – Word2Vec, skip-gram, SVM, Logistic regression, GRU, LSTM, RNN , Attention Mechanism

Examples in Natural Language Processing

➤ **Natural language processing helps the Livox app be a communication device for people with disabilities.**

➤ **Another tool enabled by natural language processing is SignAll that converts sign language into text.**

➤ **Google Translate is used by 500 million people every day to understand more than 100 world languages.**

➤ **Natural language processing technology is even being applied for aircraft maintenance.**

➤ **While the issue is complex, there's even work being done to have natural language processing assist with predictive police work to specifically identify the motive in crimes.**

Resources in Natural Language Processing

- The Stanford Natural Language Processing Group - **NLP.Stanford.edu**

- google ai blog - **https://ai.googleblog.com/**

- NeuroML - **https://neuroml.org/**

- IBM Research - **https://www.ibm.com/blogs/research/**

- Facebook AI Research - **https://ai.facebook.com/blog/?page=1**

- OpenAI

- IEEE Transactions on Audio, Speech and Language Processing
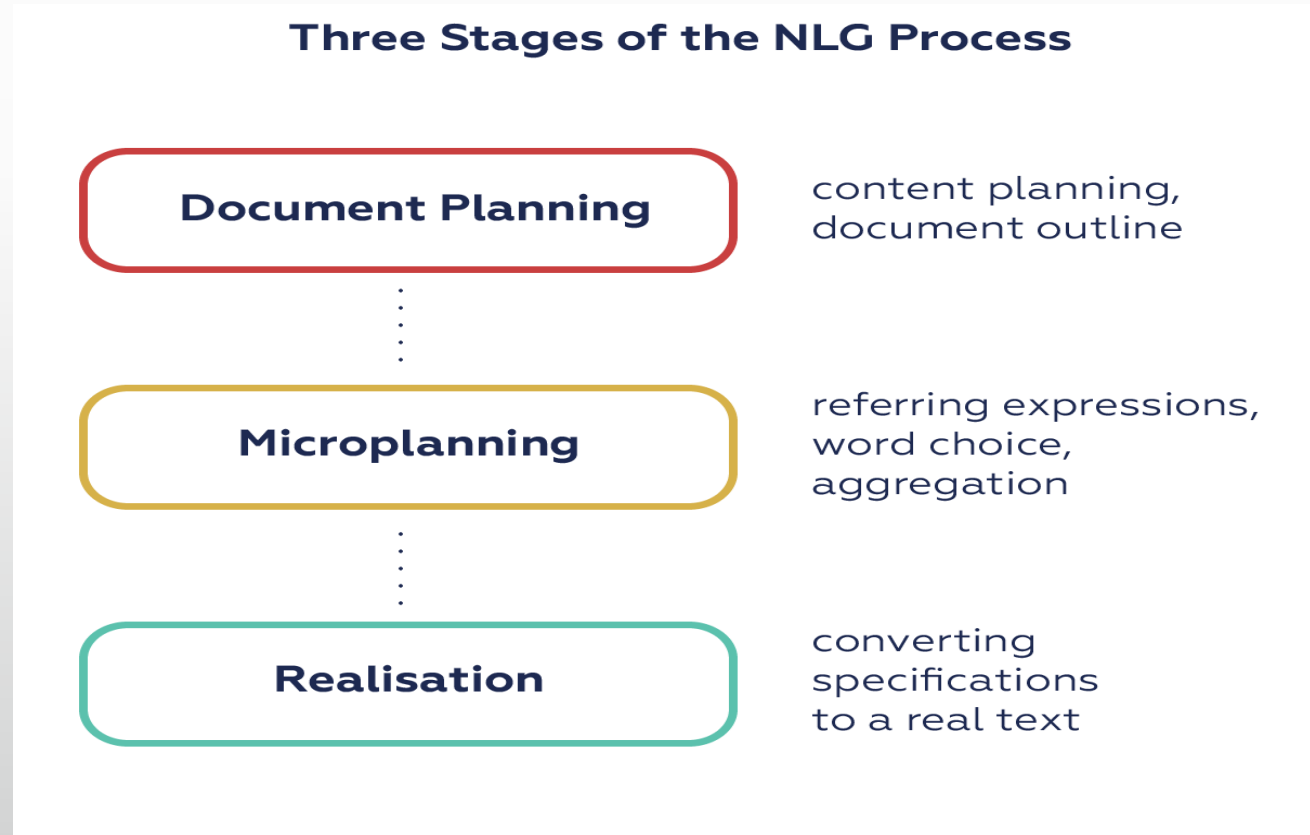
Natural Language Generation

Natural Language Generation

voodoo

➢ **Natural Language Generation, as defined by Artificial Intelligence: Natural Language Processing Fundamentals, is the "process of producing meaningful phrases and sentences in the form of natural language."**

➢ **In its essence, it automatically generates narratives that describe, summarize or explain input structured data in a human-like manner at the speed of thousands of pages per second.**

➢ **However, while NLG software can write, it can't read. The part of NLP that reads human language and turns its unstructured data into structured data understandable to computers is called Natural Language Understanding.**

Natural Language Generation

➤ In 2000, Reiter and Dale pipelined NLG architecture distinguishing **three stages** in the NLG process

Three Stages of the NLG Process

Document Planning	content planning, document outline
Microplanning	referring expressions, word choice, aggregation
Realisation	converting specifications to a real text

Natural Language Generation

➤ There are two major approaches to language generation: using templates and dynamic creation of documents.

➤ While only the latter is considered to be "real" NLG, there was a long and multistage way from basic, straightforward templates to the state-of-the-art and each new approach expanded functionality and added linguistic capacities:

1. Simple Gap-Filling Approach

2. Scripts or Rules-Producing Text

3. Word-Level Grammatical Functions

4. Dynamic Sentence Generation

5. Dynamic Document Creation

Natural Language Generation

NLG Models

- **Markov chain**
- **LSTM**
- **RNN**
- **Transformer**

NLG Tools

Commercial

- Arria NLG PLC
- AX Semantics
- Quill
- Wordsmith

Open-Source

- Simplenlg
- NaturalOWL

Thank You