

Building LLM Solutions

Module 1: What on Earth is going on!?



Hamza Farooq

Founder & CEO
Traversaal.ai



01

Introduction

Nice to meet you all!

- My name is Hamza Farooq
- Founder @traversaal.ai
- Ex- Google, Walmart Labs
- 15+ of experience in Machine Learning
- Adjunct Professor at Stanford & UCLA



Nice to meet you, from the TA!

- My name is Giulio Martini, and I am from Florence, Italy
- Master's Student at Florida State
- AI Engineer at RightData
- My primary research interests Financial Machine Learning, and Developmental Robotics



Class Etiquette

- This is an in person virtual class, please keep your video on – as much as possible
- Raise your hand [virtually] or drop a question in the chat
- Take your assignments seriously
- When in doubt, use Slack
- Reach out to Giulio during the week for questions
- All recordings are automatically available

Caution

- This is an extremely technical course in terms of coding
- You will be coding for each assignment and in each class
- You may make teams or submit the assignments on your own

Note: We won't be able to provide feedback to each individual, but a single working solution will be shared for the assignments.

Course Overview

- Module 1: Intro to Everything
- Module 2: Overview of Transformer with overview of Encoders and Decoders Models, Self Attention
- Module 3: Building basic search using fundamental blocks
- Module 4: Full Scale Search
- Module 5: Generation using Decoder Models and API
- Module 6: Fine-tuning

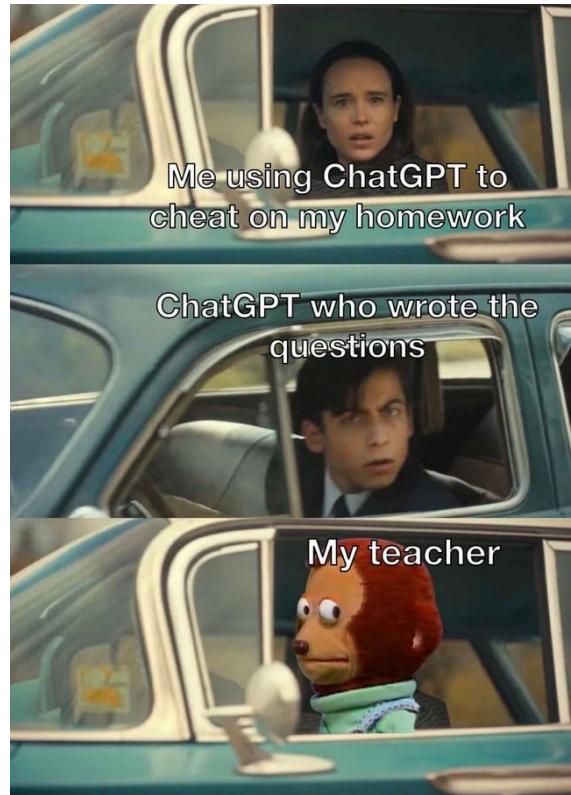
Learning outcomes for Module – 1

- Why on Earth do we need to change?
- Basic intro to Machine Learning
- What are foundational models?
- NLP Overview
- A larger overview of application of LLMs
- Working Prototypes

02

Here we go!

Why are we here?



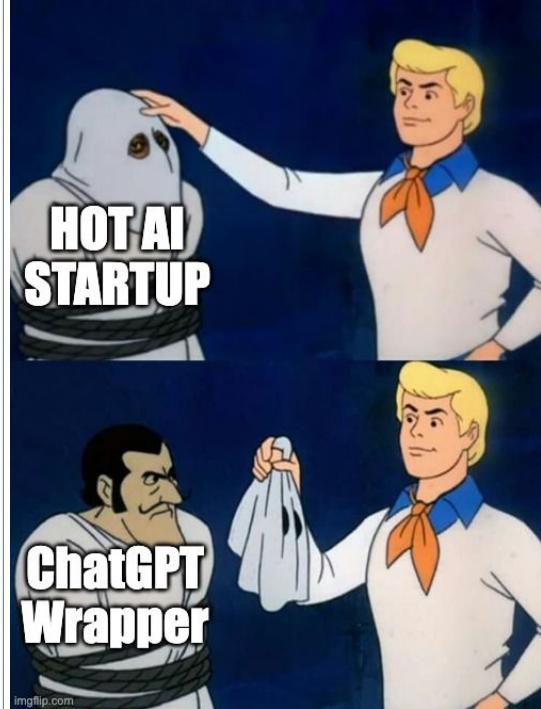
Funny Pictures
@eatliver

Yours sincerely, ChatGPT.

ME LOOKING CLEVER
ON AN EMAIL

CHATGPT

2:19 AM · Jun 13, 2023



The main question: What can Gen AI do for you?

Let's explore our first example

Find me a stylish Nike blue t-shirt specifically designed for golf, with a comfortable fit and moisture-wicking fabric, in size medium and at a reasonable price point

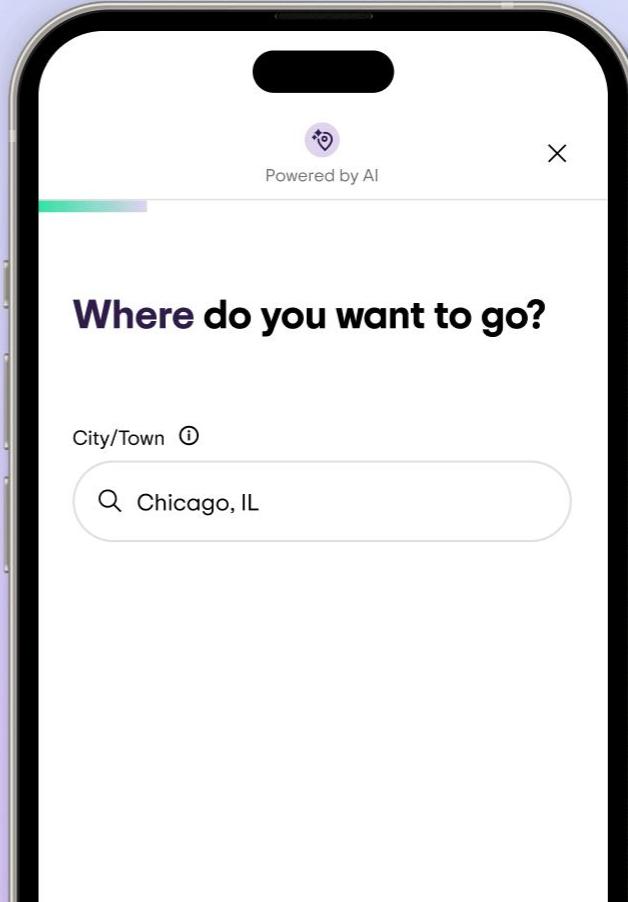
Search

- Query 1: Find me a Nike t-shirt.
- Query 2: Find me a blue t-shirt.
- Query 3: Find me a t-shirt specifically designed for golf.
- Query 4: Find me a stylish t-shirt.
- Query 5: Find me a t-shirt with a comfortable fit.
- Query 6: Find me a t-shirt with moisture-wicking fabric.
- Query 7: Find me a t-shirt in size medium.
- Query 8: Find me a t-shirt at a reasonable price point.
- Combine Results:
 - "Find me a stylish Nike blue t-shirt specifically designed for golf, with a comfortable fit and moisture-wicking fabric, in size medium and at a reasonable price point."

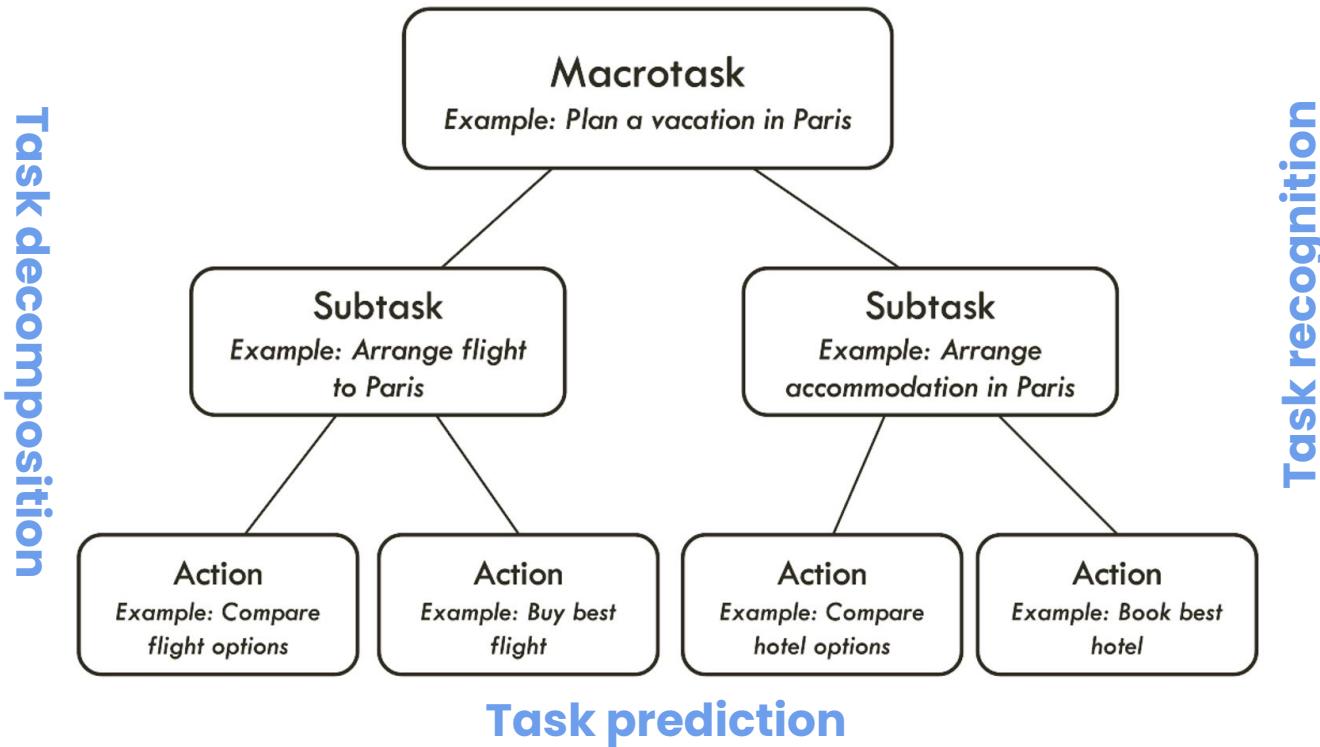


Tripadvisor[®]

**Kick-start your
travel planning**



Dividing a complex search into subtasks, to then retrieve the results from them, and combine them into **one answer**, by searching the internet is quite a hard task



And...

Complex search is difficult, multi-faceted and requires deeper engagement

- There are 10 billion search queries a day, **an estimated half of them go unanswered.**
- That's because, **people are using search to do things it wasn't originally designed to do.**
- It's great for finding a website, but **for more complex questions or tasks, too often it falls short.**
- Complex search tasks involve **multiple queries, multiple sessions or requires deep engagement with search**

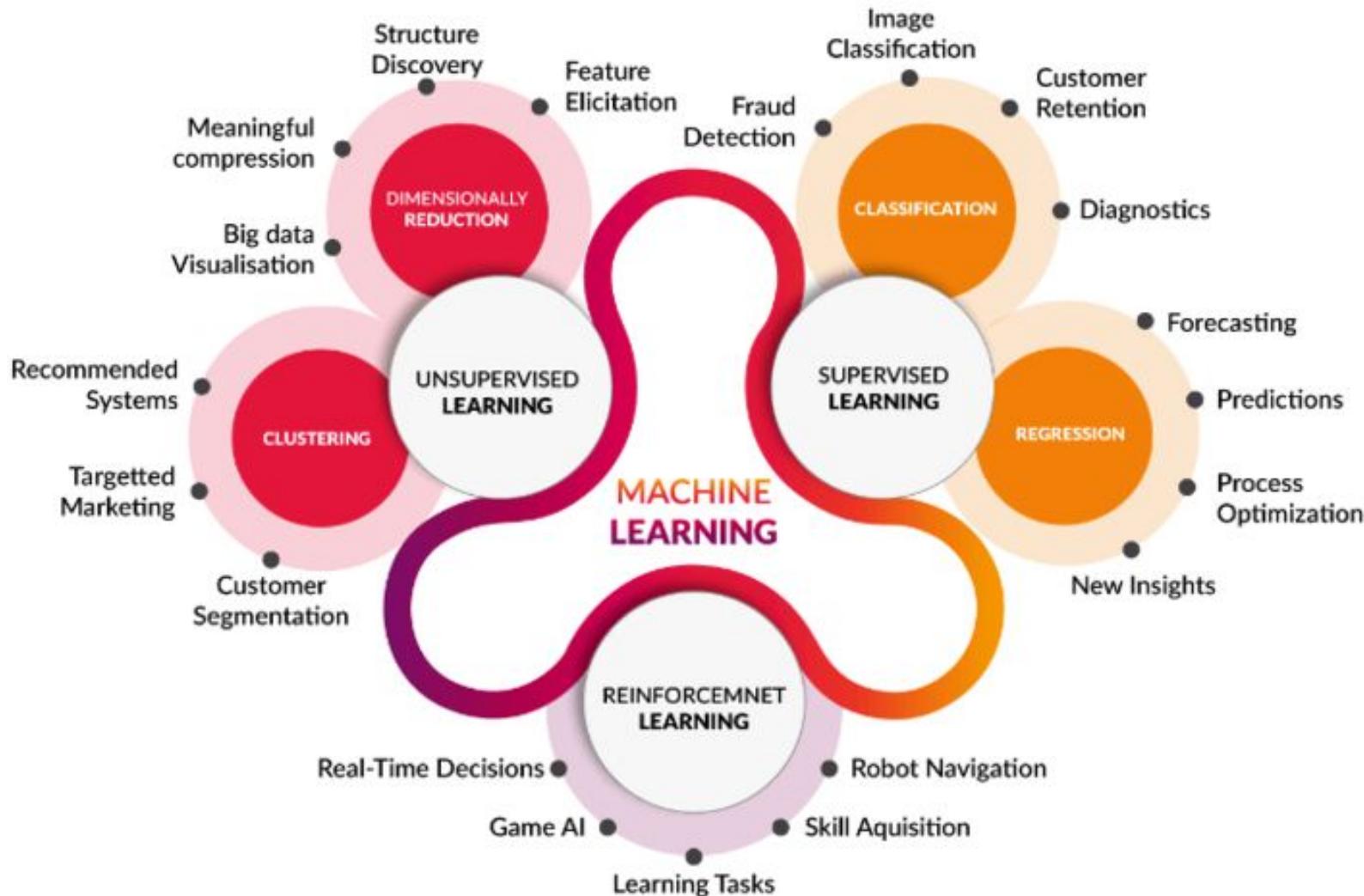
**Let's revisit
some basic
terminologies**

Machine Learning Models

Machine learning is a discipline that focuses on creating algorithms and software capable of “learning” from the information we provide, and effectively perform specific types of tasks, as a result of this training.

Some examples:

- Netflix Recommendation Models
- Sales Prediction
- Self Driving Cars



Generative AI

Generative artificial intelligence (generative AI) is a type of AI that can create new content and ideas, including conversations, stories, images, videos, and music.

AI technologies attempt to mimic human intelligence in nontraditional computing tasks like image recognition, natural language processing (NLP), and translation.

Generative AI market in focus

Use case

Text



Applications

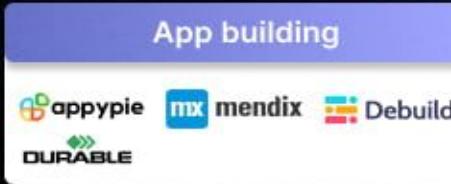
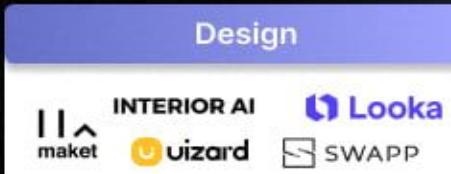
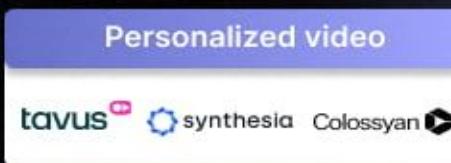
Video



Image



Code



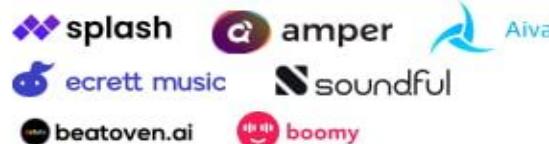
pixelplex

Use case

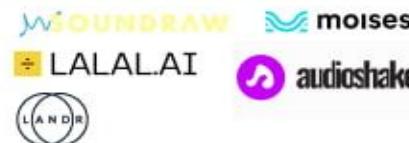
Applications

Music

- Generation & editing



- Music customization



Speech

- Text-speech / speech-text

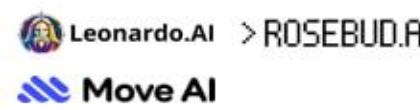


3D

- Art



- Gaming



Foundational Models

A foundation model (also called base model) is a large machine learning (ML) model trained on a vast quantity of data at scale (often by self-supervised learning or semi-supervised learning), such that it can be adapted to a wide range of downstream tasks.

Data

Text



Images



Speech



Structural Data



3D Signals



Training

Foundation model

Adaption

Tasks

Question
answering



Sentient
Analysis



Information
Extraction



Image
Captioning



Object
Recognition



Instruction
Following



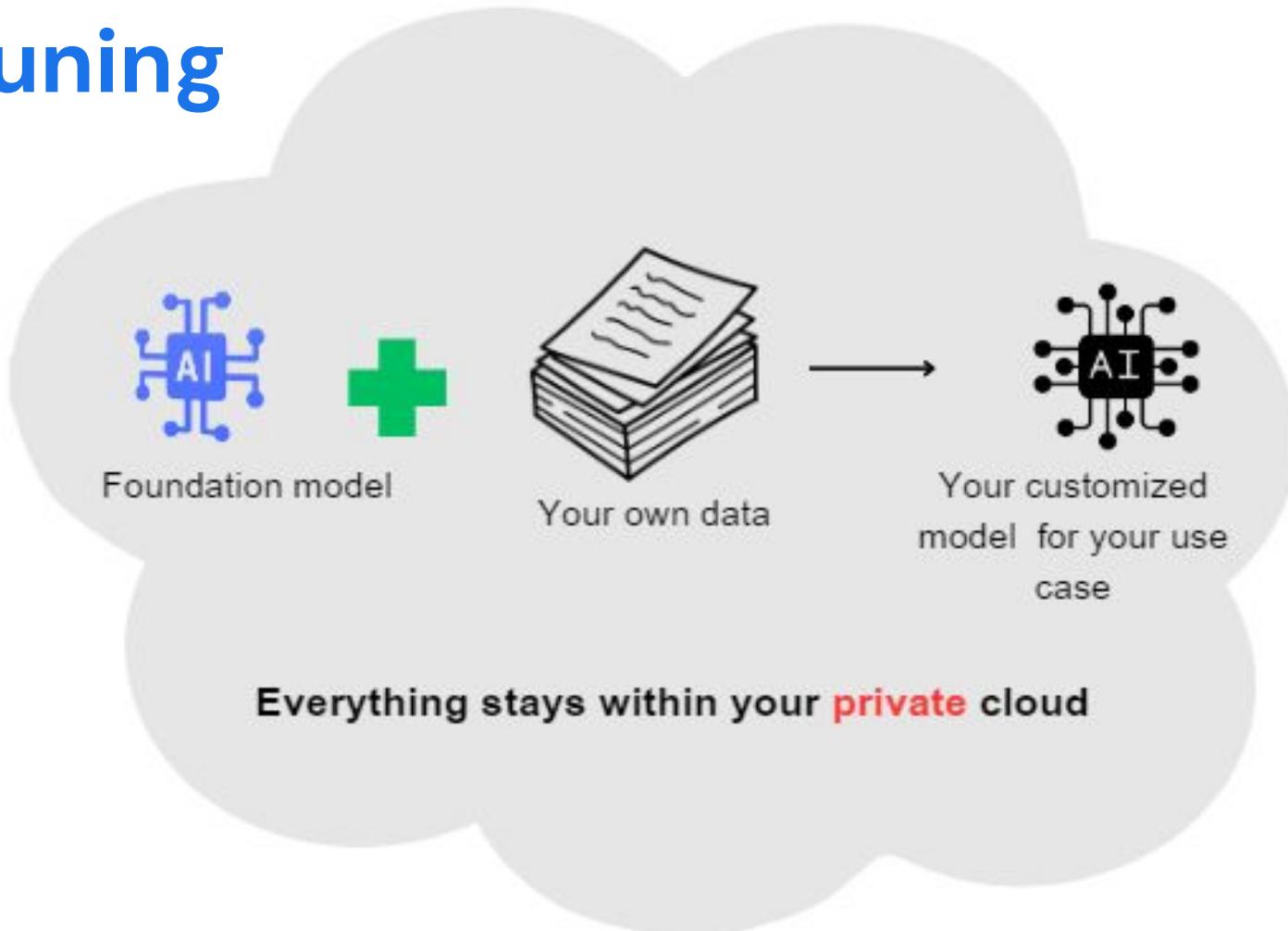
Foundation Models (FMs)

- * Pretrained
- * Generalized
- * Adaptable
- * Large
- * Self-supervised

Large Language Models (LLMs)
ex: ChatGPT, Chinchilla, GPT-3

FMs are models trained on broad data (using self-supervision at scale) that can be adapted to a wide range of downstream tasks.
<https://hai.stanford.edu/news/reflections-foundation-models>

fine-tuning



Foundational Models vs. LLMs

Large Language Models (LLMs) differ from Foundational Models in their scope of language understanding.

LLMs are specifically focused on language-related systems, while Foundational Models are attempting to stake out a broader function-based concept, which could accommodate new types of systems in the future

What is a language model?

A language model refers to a type of model specifically designed to generate human-like text or predict the probability of a sequence of words. Language models learn patterns and statistics from large amounts of text data, enabling them to generate sensible and contextually appropriate sentences.

In short...

The cat

Enter, Large Language Models

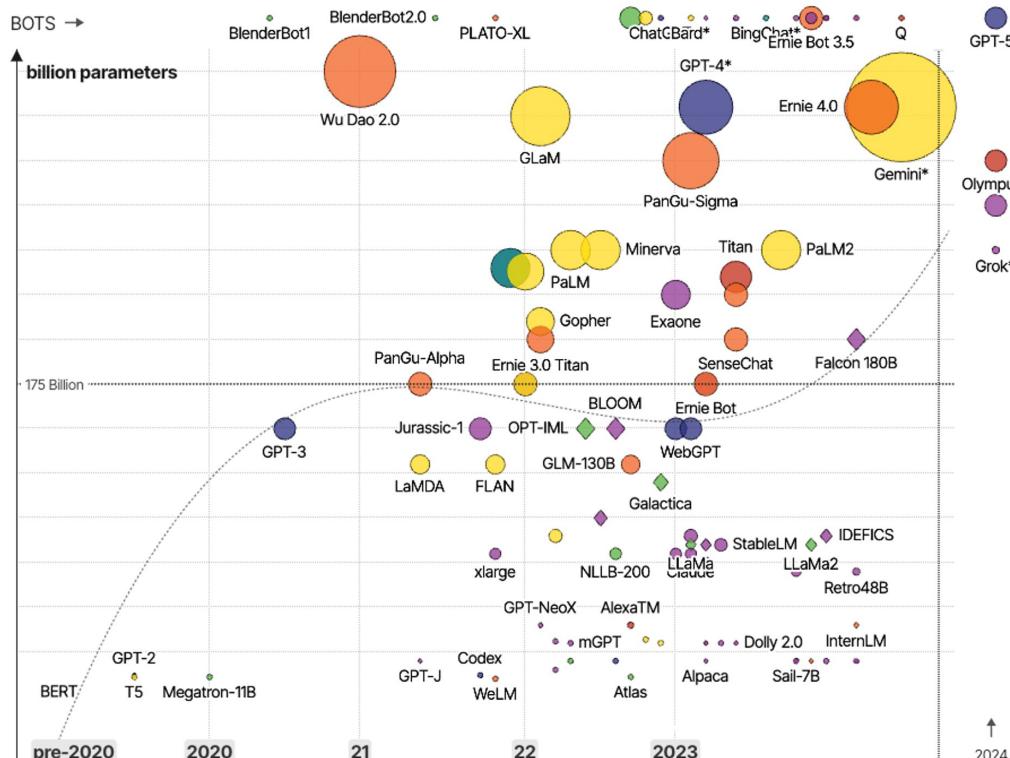
Extensive Training Data: Large language models are trained on vast amounts of text data, often comprising billions or even trillions of words. This extensive training data helps the models learn patterns, grammar, context, and a wide range of language nuances, enabling them to generate coherent and contextually appropriate responses.

Complex Architectures: Large language models employ complex architectures, such as transformer networks, that contain numerous layers and millions or even billions of parameters. These architectures enable the models to capture intricate language structures, understand semantics, and generate high-quality text by leveraging the vast amount of training data they have been exposed to. The large number of parameters allows the models to learn fine-grained details and provide nuanced responses.

The Rise and Rise of A.I. Large Language Models (LLMs) & their associated bots like ChatGPT

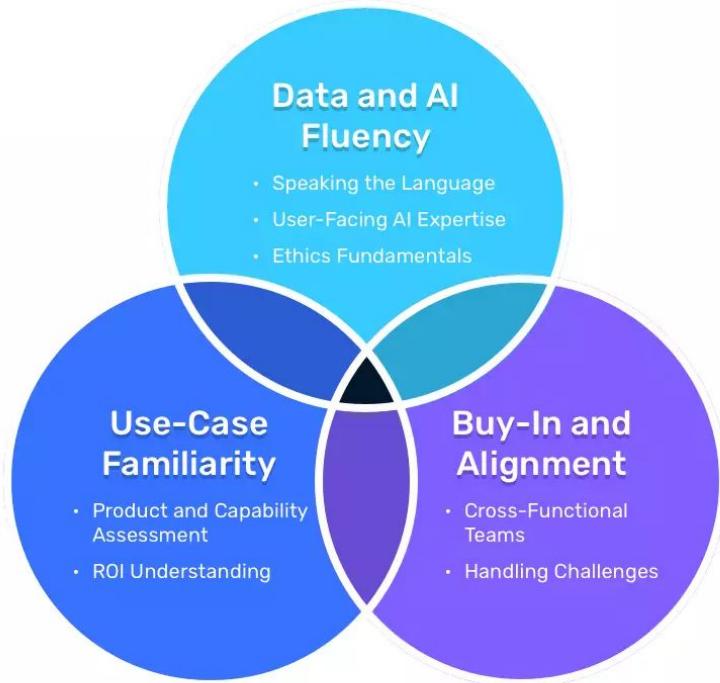
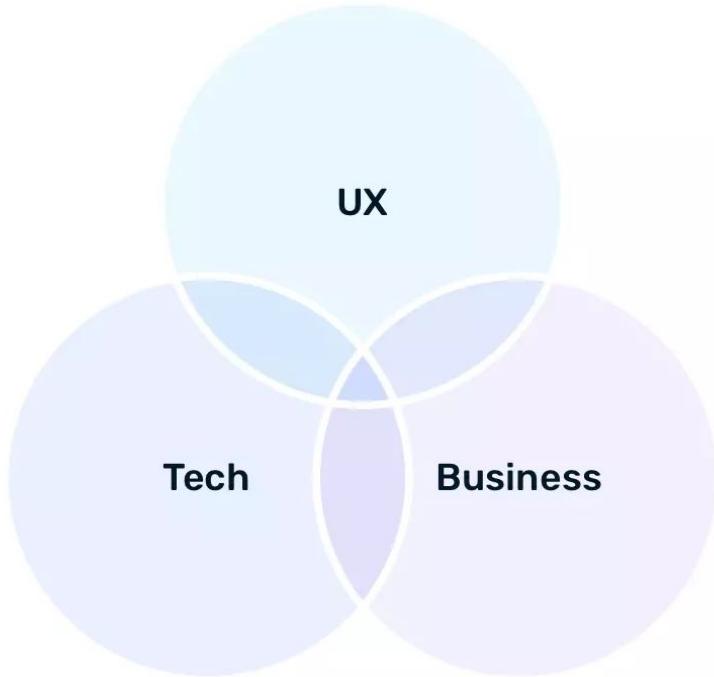
size = no. of parameters open-access

● Amazon-owned ● Chinese ● Google ● Meta / Facebook ● Microsoft ● OpenAI ● Other

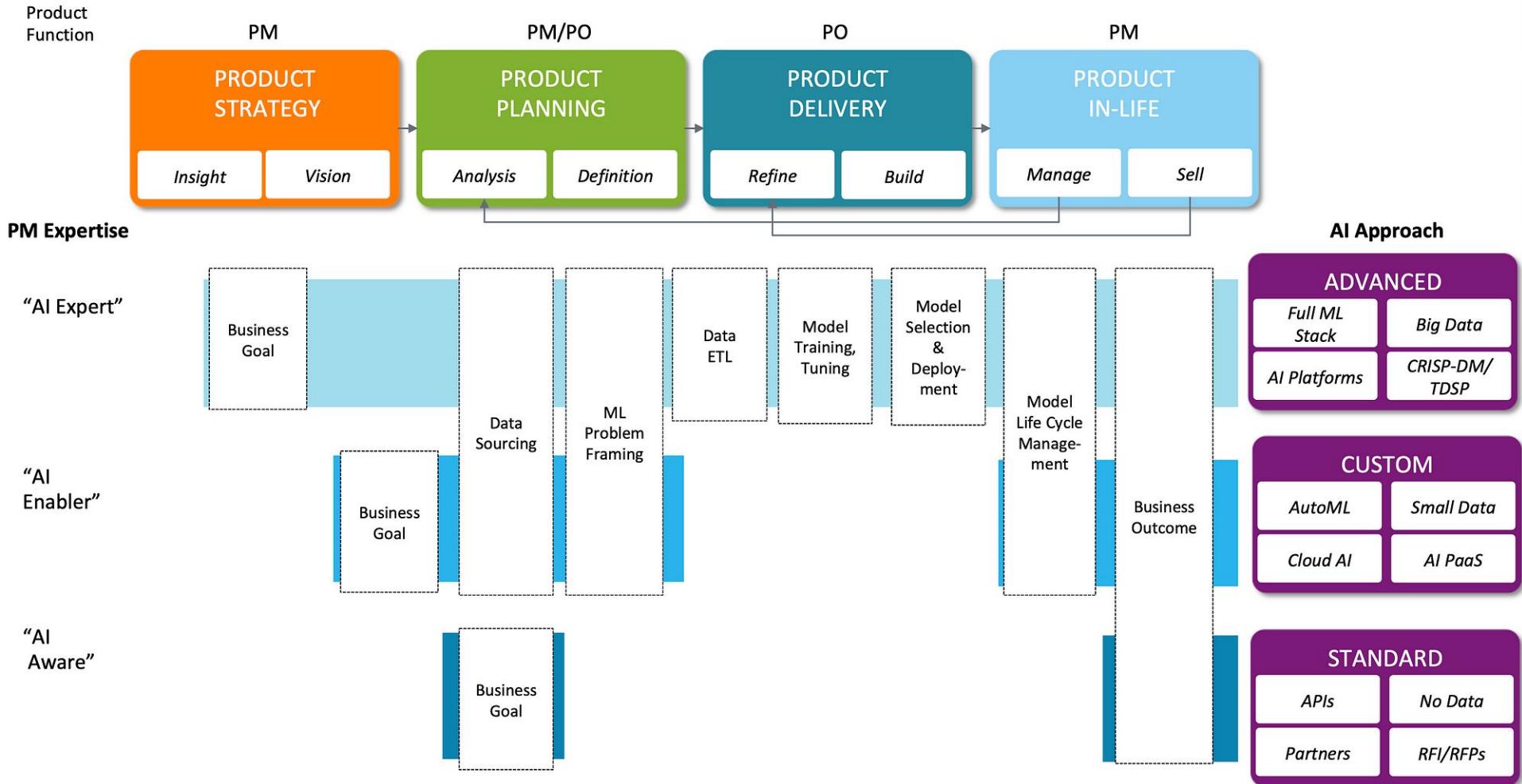




**How can things
change for you?**



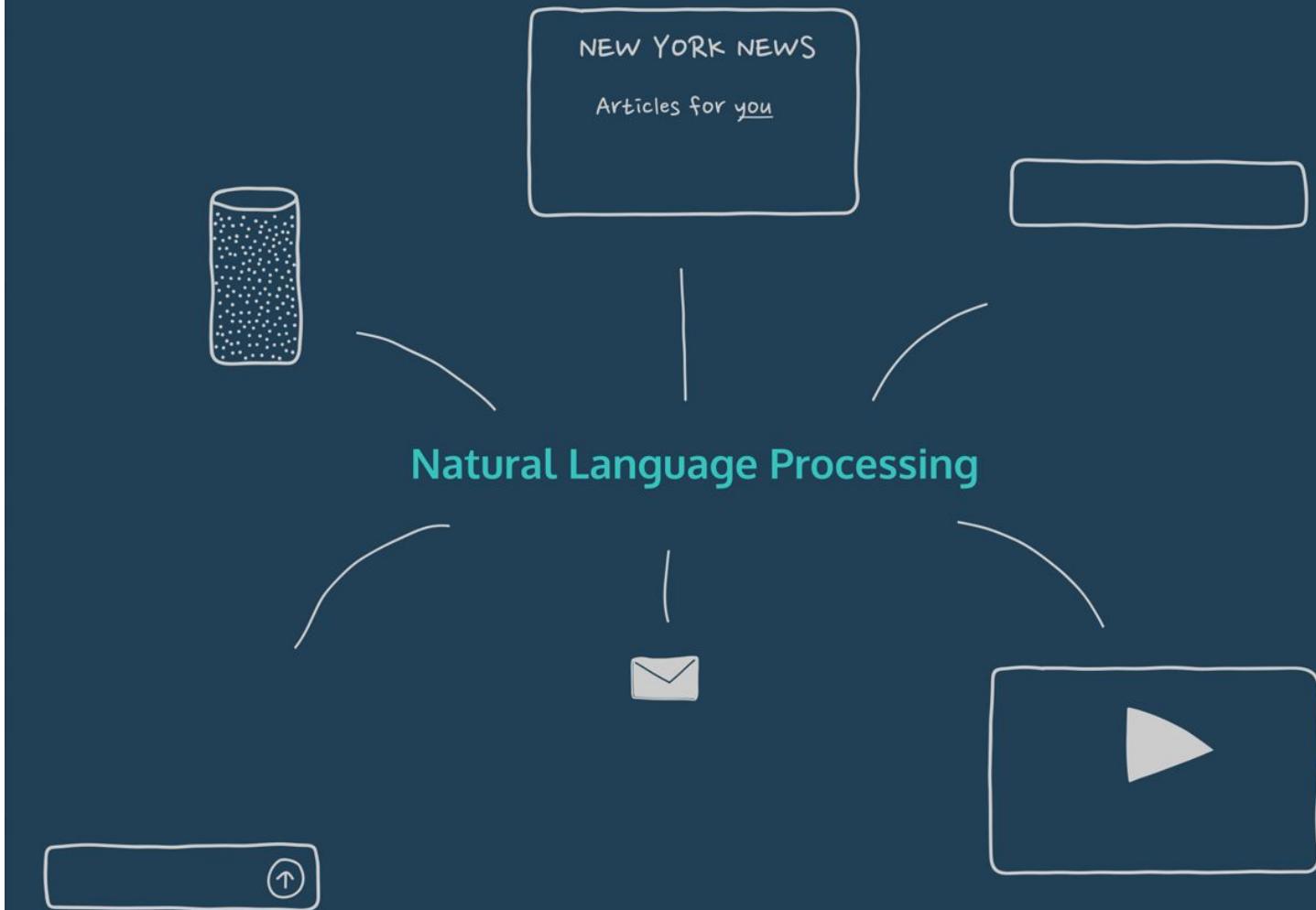
AI Product Lifecycle



03

**Let's go back in
time**

We live in a
world of NLP



What is NLP anyways?

Natural Language Processing (NLP) is defined as the branch of Artificial Intelligence that provides computers with the capability of understanding text and spoken words, in the same way a human being can.

It incorporates machine learning models, statistics, and deep learning models into computational linguistics i.e. rule-based modeling of human language to allow computers to understand text, spoken words and understands human language, intent, and sentiment.

Applications – 1

- Information retrieval
- Information extraction
- Question answering

Google list of good sushi restaurant in nyc

About 505,000,000 results (1.29 seconds)

The search results page shows a map of New York City with several red dots indicating sushi restaurants. Below the map are three cards for Sushi Nakazawa, Sushi Yasuda, and Blue Ribbon Sushi, each with a photo, rating, address, closing time, and a short description.

Sushi Nakazawa
4.7 ★★★★★ (1,038) - \$\$\$ - Sushi
23 Commerce St
Closes soon · 11PM
Dine-in · No takeout · No delivery

Sushi Yasuda
4.4 ★★★★★ (1,119) - \$\$\$ - Japanese
204 E 43rd St
Closes soon · 11PM
Good sushi, but over priced

Blue Ribbon Sushi
4.5 ★★★★★ (1,193) - \$\$ - Sushi
119 Sullivan St
Closes soon · 11PM
Good sushi, extensive menu.

View all

why is the sky blue

All Books Videos Images News More Tools

Child Bill Nye Adult Daddy

The diagram illustrates light dispersion through a prism, showing how white light is broken down into a spectrum of colors. It then shows how sunlight passing through Earth's atmosphere is scattered by particles, with blue light being scattered more than red, which is why the sky appears blue.

Thus, as sunlight of all colors passes through air, the blue part causes charged particles to oscillate faster than does the red part. ... More of the sunlight entering the atmosphere is blue than violet, however, and our eyes are somewhat more sensitive to blue light than to violet light, so the sky appears blue. Apr 7, 2003

<https://www.scientificamerican.com/article/why-is-the-sky-blue/>

Why is the sky blue? - Scientific American

About featured snippets · Feedback

People also ask

Why is the sky blue short answer?

Is the sky blue because of the ocean?

Why is the sky blue explain to a child?

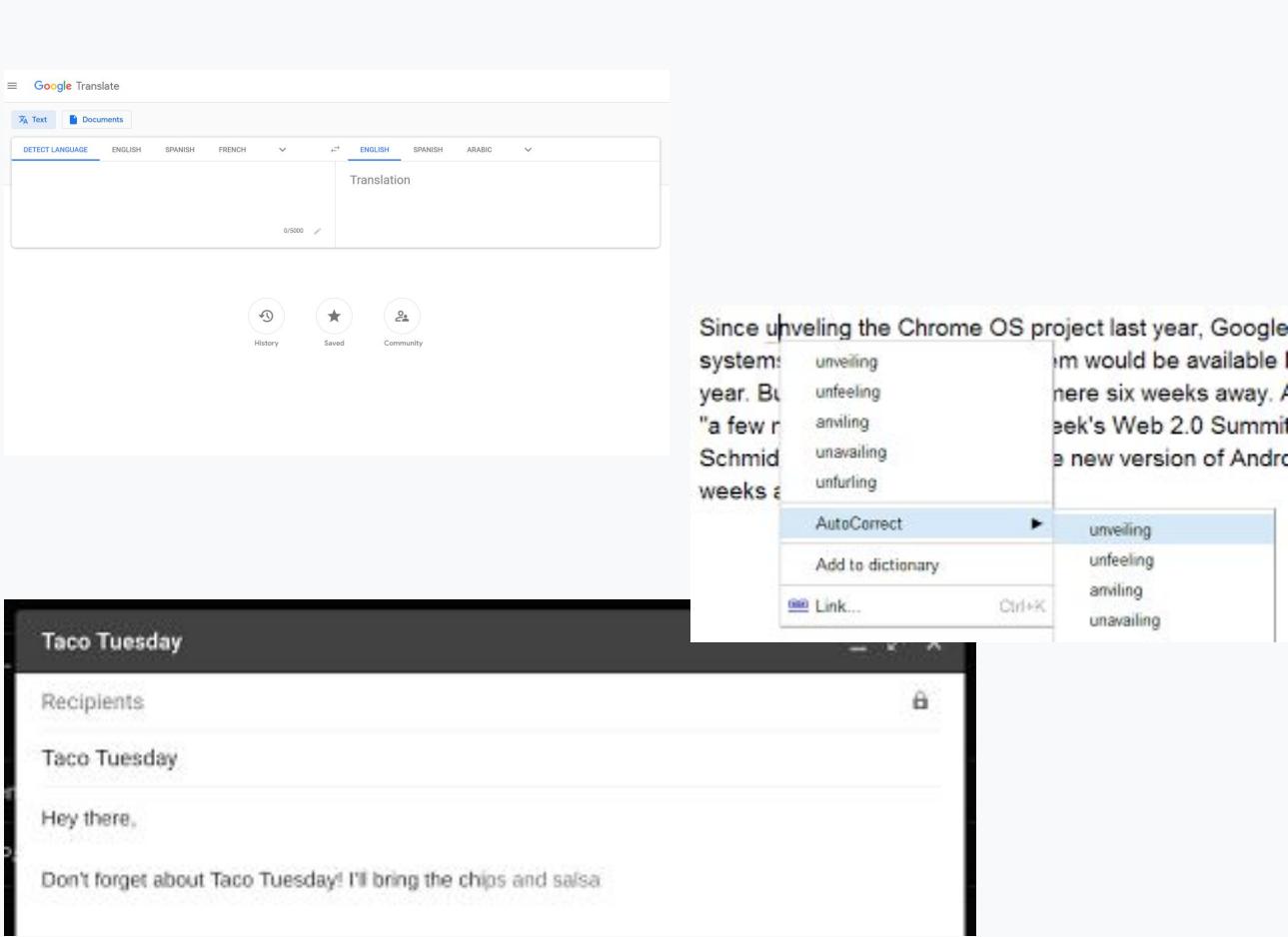
What is the reason the sky looks blue?

Feedback

Applications –2

- Machine Translation
- Summarization
- Auto Completion
- Spell Correction

Many More...



NLP Ambiguities

There are different types of ambiguities present in natural language:

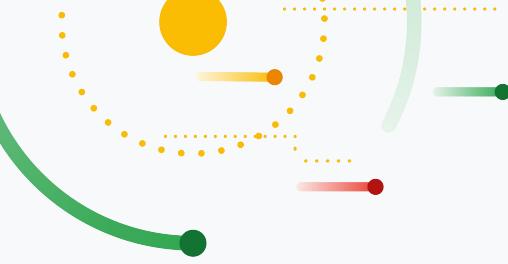
1. Lexical Ambiguity: It is defined as the ambiguity associated with the meaning of a single word. A single word can have different meanings. Also, a single word can be a noun, adjective, or verb. For example, The word “bank” can have different meanings. It can be a financial bank or a riverbank. Similarly, the word “clean” can be a noun, adverb, adjective, or verb.

NLP Ambiguities

2. Syntactic Ambiguity: It is defined as the ambiguity associated with the way the words are parsed. For example, The sentence “Visiting relatives can be boring.” This sentence can have two different meanings. One is that visiting a relative’s house can be boring. The second is that visiting relatives at your place can be boring.

NLP Ambiguities

3. Semantic Ambiguity: It is defined as ambiguity when the meaning of the words themselves can be ambiguous. For example, The sentence “Mary knows a little french.” In this sentence the word “little french” is ambiguous. As we don’t know whether it is about the language french or a person.



Common NLP tasks

Common NLP tasks

NLP systems

- Natural language understanding
- Natural language generation and summarization
- Natural language translation

Natural language understanding

- Extract information (e.g. about entities or events) from text
- Translate raw text into a meaning representation
- Reason about information given in text
- Execute NL instructions

Natural language generation and summarization

- Translate database entries or meaning representations to raw natural language text
- Produce (appropriate) utterances/responses in a dialog
- Summarize (newspaper or scientific) articles, describe images

Natural language translation

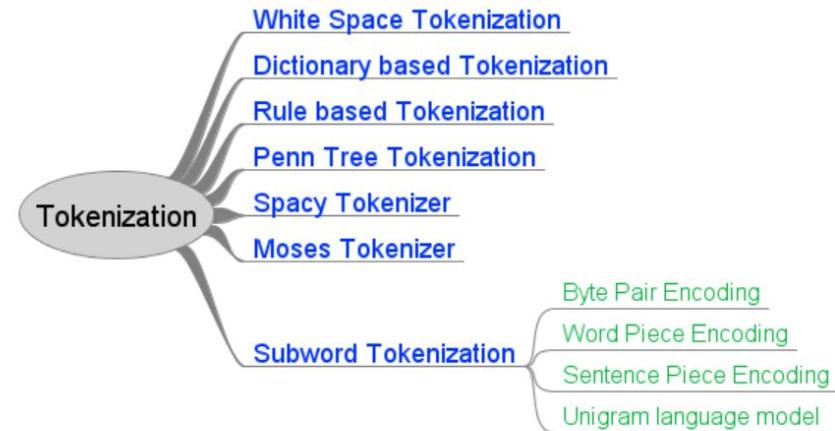
- Translate one natural language to another

Common NLP tasks

- Tokenization
 - POS tagging
 - Word sense disambiguation
 - Dependency Parsing
 - Syntactic parsing
 - Semantic analysis
 - Coreference resolution
 - Named Entity Recognition (NER)
 - Text representation
 - Text classification
 - Natural language generation
 - Multimodal NLP
-
- Tokenization is the process of breaking down a text into individual units called tokens.
 - Tokens are typically words, but can also be phrases or even individual characters, depending on the application.
 - Tokenization is a crucial step in natural language processing tasks such as machine translation, sentiment analysis, and named entity recognition.

Common NLP tasks

- Tokenization
- POS tagging
- Word sense disambiguation
- Dependency Parsing
- Syntactic parsing
- Semantic analysis
- Coreference resolution
- Named Entity Recognition (NER)
- Text representation
- Text classification
- Natural language generation
- Multimodal NLP



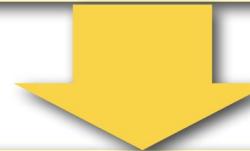
Common NLP tasks

- Tokenization
- **POS tagging**
 - POS stands for Part-of-Speech, which is a linguistic term used to describe the grammatical category of a word in a sentence.
 - POS tagging is the process of assigning each word in a text with its corresponding POS category, such as noun, verb, adjective, or adverb.
 - POS tagging is a critical component in various natural language processing tasks, including text-to-speech conversion, information retrieval, and machine translation.
- Word sense disambiguation
- Dependency Parsing
- Syntactic parsing
- Semantic analysis
- Coreference resolution
- Named Entity Recognition (NER)
- Text representation
- Text classification
- Natural language generation
- Multimodal NLP

Common NLP tasks

- Tokenization
- POS tagging
- Word sense disambiguation
- Dependency Parsing
- Syntactic parsing
- Semantic analysis
- Coreference resolution
- Named Entity Recognition (NER)
- Text representation
- Text classification
- Natural language generation
- Multimodal NLP

Open the pod door, Hal.



Verb Det Noun Noun , Name .
Open the pod door , Hal .

open:
verb, adjective, or noun?
Verb: ***open the door***
Adjective: ***the open door***
Noun: ***in the open***

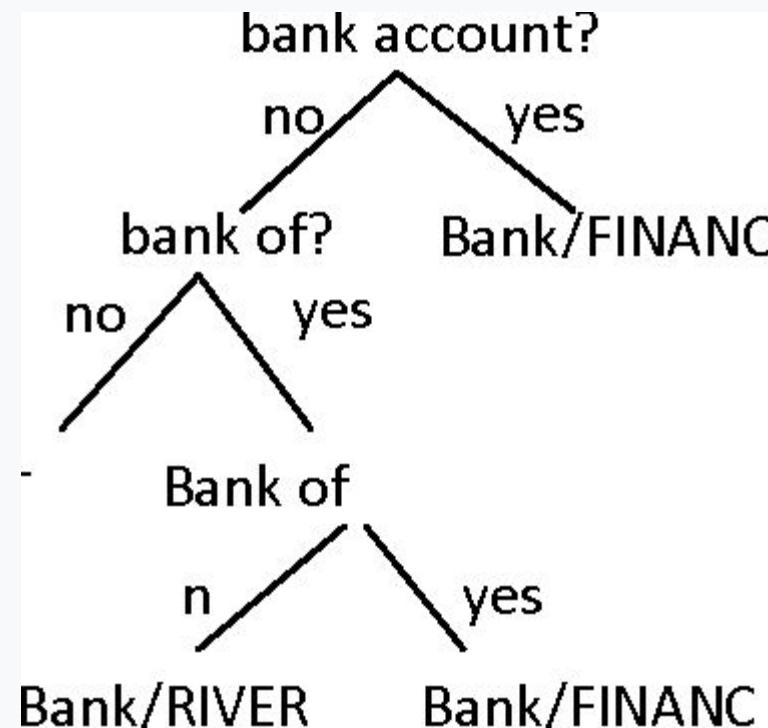
Common NLP tasks

- Tokenization
- POS tagging
- **Word sense disambiguation**
- Dependency Parsing
- Syntactic parsing
- Semantic analysis
- Coreference resolution
- Named Entity Recognition (NER)
- Text representation
- Text classification
- Natural language generation
- Multimodal NLP

- Word sense disambiguation is the process of identifying the correct meaning of a word with multiple possible meanings based on the context in which it appears.
- This is a crucial task in natural language processing because words often have different meanings depending on the context in which they are used.
- Word sense disambiguation is used in various applications, including information retrieval, machine translation, and question answering systems.

Common NLP tasks

- Tokenization
- POS tagging
- **Word sense disambiguation**
- Dependency Parsing
- Syntactic parsing
- Semantic analysis
- Coreference resolution
- Named Entity Recognition (NER)
- Text representation
- Text classification
- Natural language generation
- Multimodal NLP



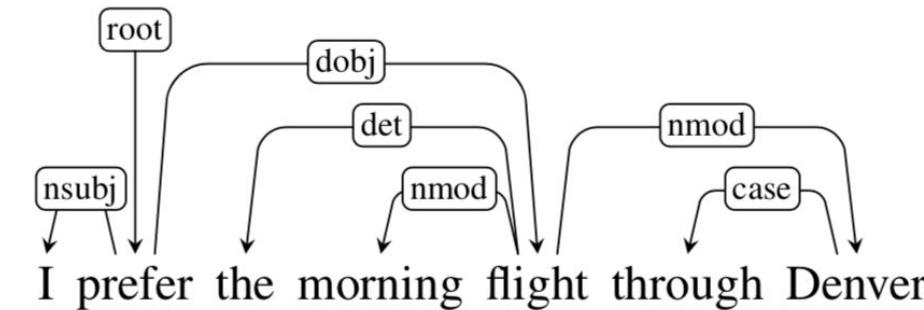
Common NLP tasks

- Tokenization
- POS tagging
- Word sense disambiguation
- **Dependency Parsing**
- Syntactic parsing
- Semantic analysis
- Coreference resolution
- Named Entity Recognition (NER)
- Text representation
- Text classification
- Natural language generation
- Multimodal NLP

- Dependency parsing is the process of analyzing the grammatical structure of a sentence by identifying the relationships between words in a sentence.
- It involves identifying the subject, object, and other dependent clauses and phrases, and representing them as a tree-like structure known as a dependency tree.
- Dependency parsing is used in various natural language processing applications, including sentiment analysis, named entity recognition, and machine translation.

Common NLP tasks

- Tokenization
- POS tagging
- Word sense disambiguation
- **Dependency Parsing**
- Syntactic parsing
- Semantic analysis
- Coreference resolution
- Named Entity Recognition (NER)
- Text representation
- Text classification
- Natural language generation
- Multimodal NLP



- **Head-Dependent:** In the arrows representing relationship, the origin word is the Head & the destination word is Dependent.
- **Root:** Word which is the root of our parse tree. It is 'prefer' in the above example.
- **Grammar Functions and Arcs:** Tags between each Head-Dependent pair is a grammar function determining the relation between the Head & Dependent. The arrowhead carrying the tag is called an Arc.

Clausal Argument Relations	Description
NSUBJ	Nominal subject
DOBJ	Direct object
IOBJ	Indirect object
CCOMP	Clausal complement
XCOMP	Open clausal complement

Nominal Modifier Relations	Description
NMOD	Nominal modifier
AMOD	Adjectival modifier
NUMMOD	Numeric modifier
APPOS	Appositional modifier
DET	Determiner
CASE	Prepositions, postpositions and other case markers

Other Notable Relations	Description
CONJ	Conjunct
CC	Coordinating conjunction

Common NLP tasks

- Tokenization
- POS tagging
- Word sense disambiguation
- Dependency Parsing
- **Syntactic parsing**
- Semantic analysis
- Coreference resolution
- Named Entity Recognition (NER)
- Text representation
- Text classification
- Natural language generation
- Multimodal NLP

- Syntactic parsing is the process of analyzing the grammatical structure of a sentence to determine its syntactic components, such as nouns, verbs, adjectives, and adverbs.
- It involves identifying the parts of speech of each word in the sentence and grouping them together into phrases and clauses based on their syntactic relationships.
- Syntactic parsing is used in various natural language processing applications, including text-to-speech conversion, machine translation, and information retrieval.

Common NLP tasks

- Tokenization
- POS tagging
- Word sense disambiguation
- Dependency Parsing
- **Syntactic parsing**
- Semantic analysis
- Coreference resolution
- Named Entity Recognition (NER)
- Text representation
- Text classification
- Natural language generation
- Multimodal NLP

- POS tagging is the process of labeling individual words in a sentence with their part of speech, such as noun, verb, adjective, or adverb, while syntactic parsing involves analyzing the relationships between the words to determine the overall grammatical structure of the sentence.
- For example, consider the sentence "John eats pizza." POS tagging would label "John" as a proper noun and "eats" as a verb, while syntactic parsing would identify "John" as the subject of the verb "eats" and "pizza" as the object of the verb.
- In short, POS tagging is concerned with the individual words, while syntactic parsing focuses on the overall sentence structure.

Common NLP tasks

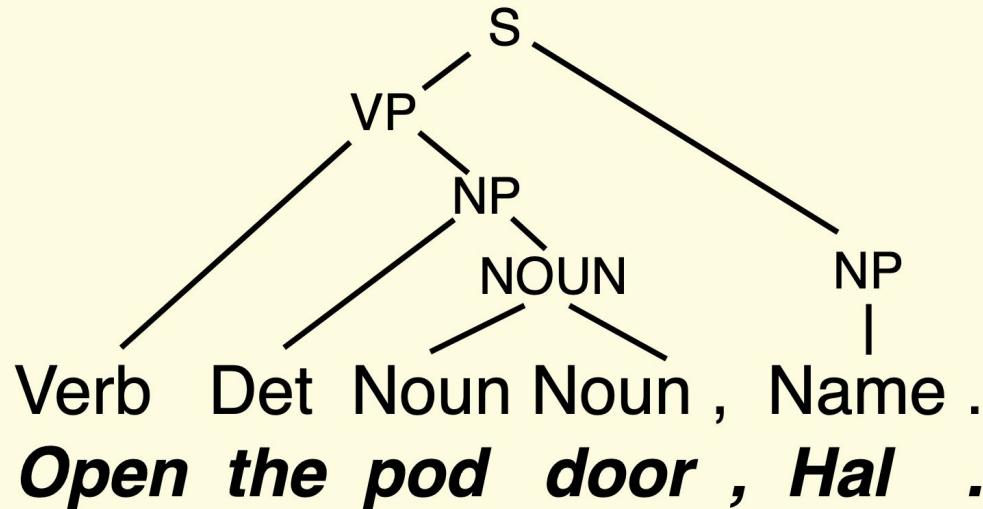
- Tokenization
- POS tagging
- Word sense disambiguation
- Dependency Parsing
- Syntactic parsing
- **Semantic analysis**
- Coreference resolution
- Named Entity Recognition (NER)
- Text representation
- Text classification
- Natural language generation
- Multimodal NLP

- Semantic analysis is the process of extracting the meaning of a text by analyzing the relationships between words and phrases in a sentence.
- It involves identifying the underlying concepts and ideas conveyed by the text and representing them in a structured form, such as a knowledge graph or ontology.
- Semantic analysis is used in various natural language processing applications, including question answering, information retrieval, and chatbots, to enable more accurate and intelligent responses.

Common NLP tasks

- Tokenization
- POS tagging
- Word sense disambiguation
- Dependency Parsing
- Syntactic parsing
- **Semantic analysis**
- Coreference resolution
- Named Entity Recognition (NER)
- Text representation
- Text classification
- Natural language generation
- Multimodal NLP

$\exists x \exists y (\text{pod_door}(x) \ \& \ \text{Hal}(y) \\ \& \ \text{request}(\text{open}(x, y)))$



Common NLP tasks

- Tokenization
- POS tagging
- Word sense disambiguation
- Dependency Parsing
- Syntactic parsing
- **Semantic analysis**
- Coreference resolution
- Named Entity Recognition (NER)
- Text representation
- Text classification
- Natural language generation
- Multimodal NLP

We need a **meaning representation language**.

“Shallow” semantic analysis: **Template-filling**
(Information Extraction)

Named-Entity Extraction: Organizations, Locations, Dates,...
Event Extraction

“Deep” semantic analysis: (Variants of) **formal logic**
 $\exists x \exists y (\text{pod_door}(x) \& \text{Has}(y) \& \text{request}(\text{open}(x, y)))$

We also distinguish between
Lexical semantics (the meaning of words) and
Compositional semantics (the meaning of sentences)

Common NLP tasks

- Tokenization
- POS tagging
- Word sense disambiguation
- Dependency Parsing
- Syntactic parsing
- Semantic analysis
- **Coreference resolution**
- Named Entity Recognition (NER)
- Text representation
- Text classification
- Natural language generation
- Multimodal NLP

More than a decade ago, Carl Lewis stood on the threshold of what was to become the greatest athletics career in history. He had just broken two of the legendary Jesse Owens' college records, but never believed he would become a corporate icon, the focus of hundreds of millions of dollars in advertising. His sport was still nominally amateur. Eighteen Olympic and World Championship gold medals and 21 world records later, Lewis has become the richest man in the history of track and field -- a multi-millionaire.

Who is Carl Lewis?
Did Carl Lewis break any world records?
(and how do you know that?)

Common NLP tasks

- Tokenization
- POS tagging
- Word sense disambiguation
- Dependency Parsing
- Syntactic parsing
- Semantic analysis
- **Coreference resolution**
 - Coreference resolution is the task of identifying all the expressions (e.g., pronouns, names) in a text that refer to the same entity, and linking them together.
 - It is a crucial task in natural language processing as it enables a system to maintain a consistent representation of entities throughout a document, enabling more accurate information extraction and text understanding.
- Named Entity Recognition (NER)
- Text representation
- Text classification
- Natural language generation
- Multimodal NLP

Common NLP tasks

- Tokenization
- POS tagging
- Word sense disambiguation
- Dependency Parsing
- Syntactic parsing
- Semantic analysis
- Coreference resolution
- **Named Entity Recognition (NER)**
- Text representation
- Text classification
- Natural language generation
- Multimodal NLP

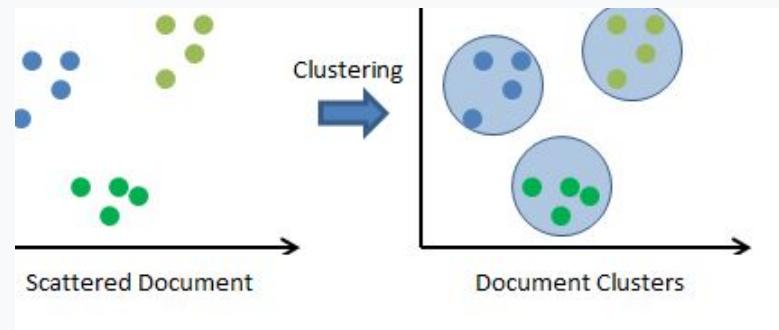
Named entity recognition (NER) is the process of identifying and categorizing named entities in a text, such as people, organizations, locations, and dates.

When Sebastian Thrun PERSON started at Google ORG in 2007 DATE, few people outside of the company took him seriously. "I can tell you very senior CEOs of major American NORP car companies would shake my hand and turn away because I wasn't worth talking to," said Thrun PERSON, now the co-founder and CEO of online higher education startup Udacity, in an interview with Recode ORG earlier this week DATE.

A little less than a decade later DATE, dozens of self-driving startups have cropped up while automakers around the world clamor, wallet in hand, to secure their place in the fast-moving world of fully automated transportation.

Common NLP tasks

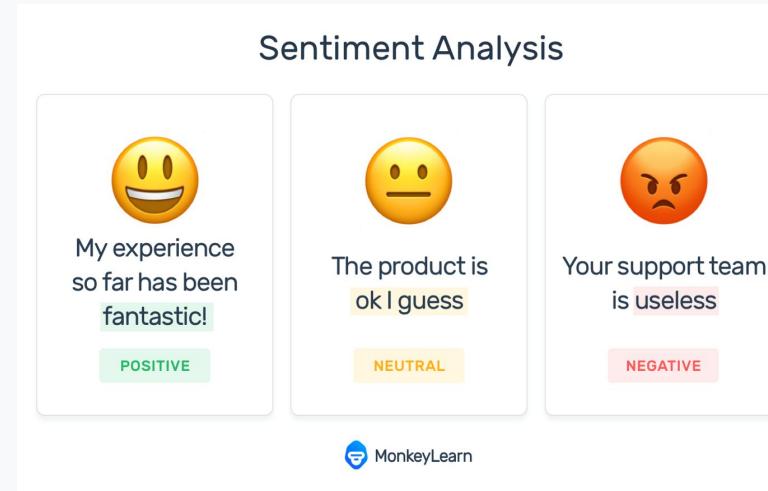
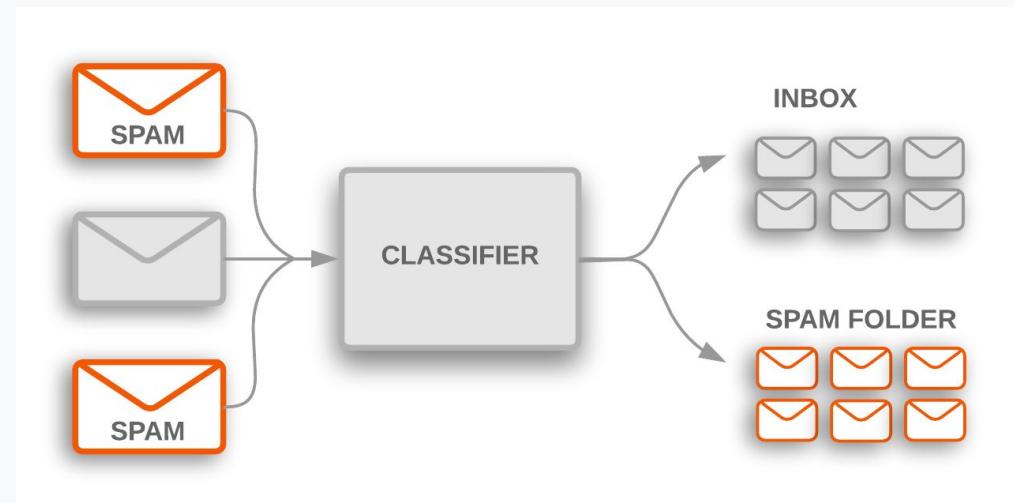
- Tokenization
- POS tagging
- Word sense disambiguation
- Dependency Parsing
- Syntactic parsing
- Semantic analysis
- Coreference resolution
- Named Entity Recognition (NER)
- **Text representation**
- Text classification
- Natural language generation
- Multimodal NLP



- Text representation is the process of converting unstructured text data into a structured format that can be used for natural language processing tasks.
- It involves selecting a suitable representation scheme, such as bag-of-words, word embeddings, or topic models, to capture the key features and characteristics of the text data in a numerical form that can be processed by machine learning algorithms.

Common NLP tasks

- Tokenization
- POS tagging
- Word sense disambiguation
- Dependency Parsing
- Syntactic parsing
- Semantic analysis
- Coreference resolution
- Named Entity Recognition (NER)
- Text representation
- Text classification**
- Natural language generation
- Multimodal NLP



Common NLP tasks

- Tokenization
- POS tagging
- Word sense disambiguation
- Dependency Parsing
- Syntactic parsing
- Semantic analysis
- Coreference resolution
- Named Entity Recognition (NER)
- Text representation
- Text classification
- **Natural language generation**
- Multimodal NLP

Input Article

Marseille, France (CNN) The French prosecutor leading an investigation into the crash of Germanwings Flight 9525 insisted Wednesday that he was not aware of any video footage from on board the plane. Marseille prosecutor Brice Robin told CNN that " so far no videos were used in the crash investigation ." He added, " A person who has such a video needs to immediately give it to the investigators ." Robin 's comments follow claims by two magazines, German daily Bild and French Paris Match, of a cell phone video showing the harrowing final seconds from on board Germanwings Flight 9525 as it crashed into the French Alps . All 150 on board were killed. Paris Match and Bild reported that the video was recovered from a phone at the wreckage site. ...

Abstractive summarization

Generated summary

Prosecutor : " So far no videos were used in the crash investigation "

Text
Summarization
Models

Extractive summarization

Extractive summary

marseille prosecutor brice robin told cnn that " so far no videos were used in the crash investigation ." robin 's comments follow claims by two magazines , german daily bild and french paris match , of a cell phone video showing the harrowing final seconds from on board germanwings flight 9525 as it crashed into the french alps . paris match and bild reported that the video was recovered from a phone at the wreckage site .

Sentence having the right answer

'context': 'Beyoncé Giselle Knowles-Carter (/bi: 'jpnsei/ bee-YON-say) (born September 4, 1981) is an American singer, songwriter, record producer and actress. Born and raised in Houston, Texas, she performed in various singing and dancing competitions as a child, and rose to fame in the late 1990s as lead singer of R&B girl-group Destiny's Child. Managed by her father, Mathew Knowles, the group became one of the world's best-selling girl groups of all time. Their hiatus saw the release of Beyoncé's debut album, Dangerously in Love (2003), which established her as a solo artist worldwide, earned five Grammy Awards and featured the Billboard Hot 100 number-one singles "Crazy in Love" and "Baby Boy".',
'text': 'in the late 1990s'
'question': 'When did Beyonce start becoming popular?'

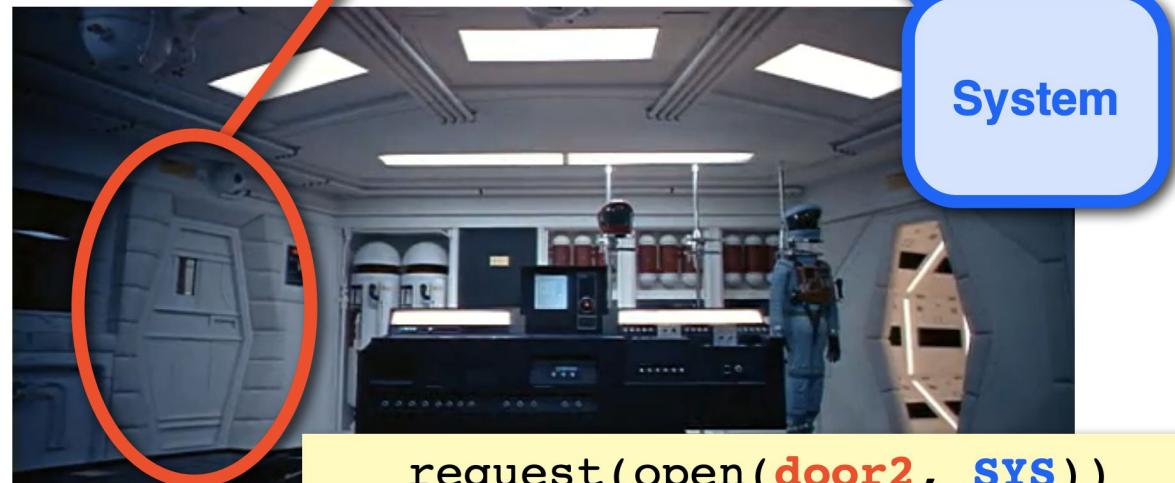
Exact Answer

Common NLP tasks

- Tokenization
- POS tagging
- Word sense disambiguation
- Dependency Parsing
- Syntactic parsing
- Semantic analysis
- Coreference resolution
- Named Entity Recognition (NER)
- Text representation
- Text classification
- Natural language generation
- Multimodal NLP

Multimodal NLP: mapping from language to the world

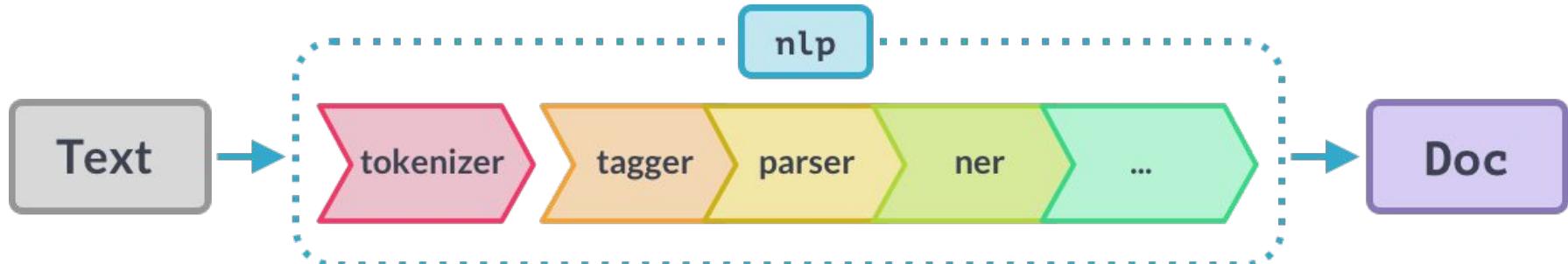
$\exists x \exists y (\text{pod_door}(x) \And \text{Hal}(y) \And \text{request}(\text{open}(x, y)))$



spaCy Package

spaCy is an open-source library used for natural language processing in python. It is extremely popular for processing a large amount of unstructured data generated at a vast scale in the industry and generate useful and meaningful insights from the data.

spaCy NLP Pipeline



Let's code

Colab

Appendix