Natural Language Processing and Large Language Models

Corso di Laurea Magistrale in Ingegneria Informatica

Lesson 1

NLP Overview

Nicola Capuano and Antonio Greco

DIEM – University of Salerno

Outline

- What is Natural Language Processing
- Applications of NLP
- History of NLP

What is Natural Language Processing

NLP in the Press

New powerful AI bot creates angst among users: Are robots ready to take our jobs?

The New York Times

A Smarter Robot

A new chatbot shows rapid advances in artificial intelligence.

The Washington Post

What is ChatGPT, the viral social media AI?

This AI chatbot is dominating social media with its frighteningly good essays

BUSINESS INSIDER

ChatGPT may be coming for our jobs. Here are the 10 roles that AI is most likely to replace.

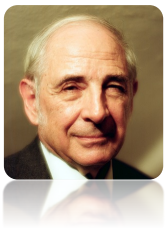
REUTERS

Microsoft co-founder Bill Gates: ChatGPT 'will change our world'

Importance of NLP

Natural language is the most important part of Artificial Intelligence

John Searle, Philosopher



Natural language processing is a cornerstone of artificial intelligence, allowing computers to read and understand human language, as well as to produce and recognize speech

Ginni Rometty, IBM CEO

Natural language processing is one of the most important fields in artificial intelligence and also one of the most difficult

Dan Jurafsky, Professor of Linguistics and Computer Science at Stanford University

Definitions

Natural language processing is the set of methods for making **human language accessible to computers**

(Jacob Eisenstein)

Natural language processing is the field at the **intersection of computer science and linguistics**

(Christopher Manning)

Make **computers to understand natural language** to do certain task humans can do such as **translation**, **summarization**, **questions answering**

(Behrooz Mansouri)

Definitions

Natural language processing is an area of research in **computer science** and **artificial intelligence** concerned with **processing natural languages** such as English or Mandarin.

This processing generally involves **translating natural language into data that a computer can use** to learn about the world.

And this understanding of the world is sometimes used to **generate natural language text** that reflects that understanding.

(Natural Language Processing in Action)

Natural Language Understanding

A subfield of NLP focused on **transforming human language in a way that machines can process**

- Involves extracting **meaning**, **context**, and **intent** from text
- Text is transformed into a **numerical representation (embedding)**

Who uses Embeddings:

- **Search Engines…** to interpret the meaning behind search queries
- **Email Clients…** to detect spam and classify emails as important or not
- **Social Media…** to moderate posts and understand user sentiment
- **CRM Tools…** to analyze customer inquiries and route them
- **Recommender Systems…** to suggest articles, products, or content

Natural Language Generation

A subfield of NLP focused on **generating human-like text**

- Involves creating **coherent**, **contextually appropriate text**
- Based on a numerical representation of the **meaning and sentiment** you would like to convey

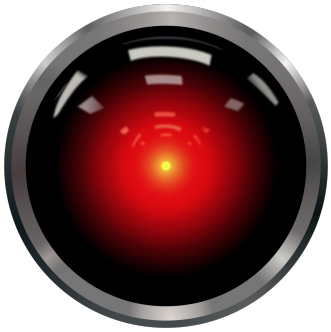
Applications:

- **Machine Translation…** translates text from one language to another
- **Text Summarization…** creation of concise summaries of long documents preserving key information
- **Dialogue Processing…** powers chatbots and virtual assistants to provide relevant responses in conversations
- **Content Creation…** generation of articles, reports, stories, poetry, …

Example: Conversational Agents

Conversational agents include:

- **Speech recognition**
- **Language analysis**
- **Dialogue processing**
- **Information retrieval**
- **Text to speech**



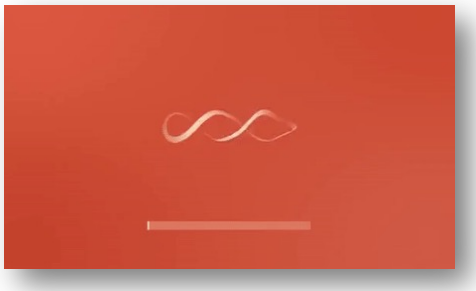
Open the pod bay doors, Hal.

I'm sorry, Dave, I'm afraid I can't do that.

What are you talking about, Hal?

I know that you and Frank were planning to disconnect me, and I'm afraid that's something I cannot allow to happen.

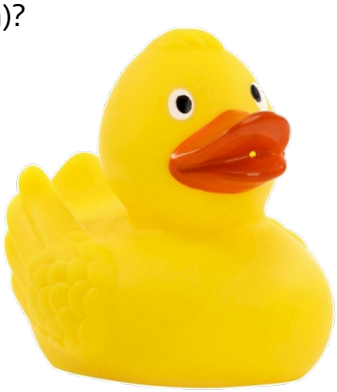
Conversational Agents in Movies



NLP is Hard

I made her duck... what does it means?

- Duck: **noun** (waterfowl) or **verb** (getting down)?
- Make: **cook X** or **cause X to do Y**?
- Her: **for her** or **belonging to her**?



Possible meanings:

- I cooked waterfowl for her
- I cooked waterfowl belonging to her
- I created the (plaster?) duck she owns
- I caused her to quickly lower her head or body
- I waved my magic wand and turned her into undifferentiated waterfowl

Ambiguity

Natural language is **extremely rich in form and structure** and **very ambiguous**

- One input can mean many **different things**
- Many input can mean the **same thing**

Levels of ambiguity

- **Lexical ambiguity**: different meanings of words
- **Syntactic ambiguity**: different ways to parse the sentence
- **Interpreting partial information**: how to interpret pronouns
- **Contextual information**: context of the sentence may affect the meaning of that sentence

Ambiguity

I saw bats… ?

Call me a cab… ?

NLP and Linguistics

NLP techniques draw on various aspects of linguistics:

- **Phonetics:** understanding the physical sounds of speech and how they are produced and perceived
- **Morphology:** Knowledge of the structure and formation of words, including their meaningful components (morphemes)
- **Syntax:** Understanding the rules and structures governing the arrangement of words in sentences
- **Semantics:** Insight into the meaning of words, phrases, and sentences
- **Pragmatics:** Understanding how context influences the interpretation of meaning

NLP vs Linguistics

Linguistics:

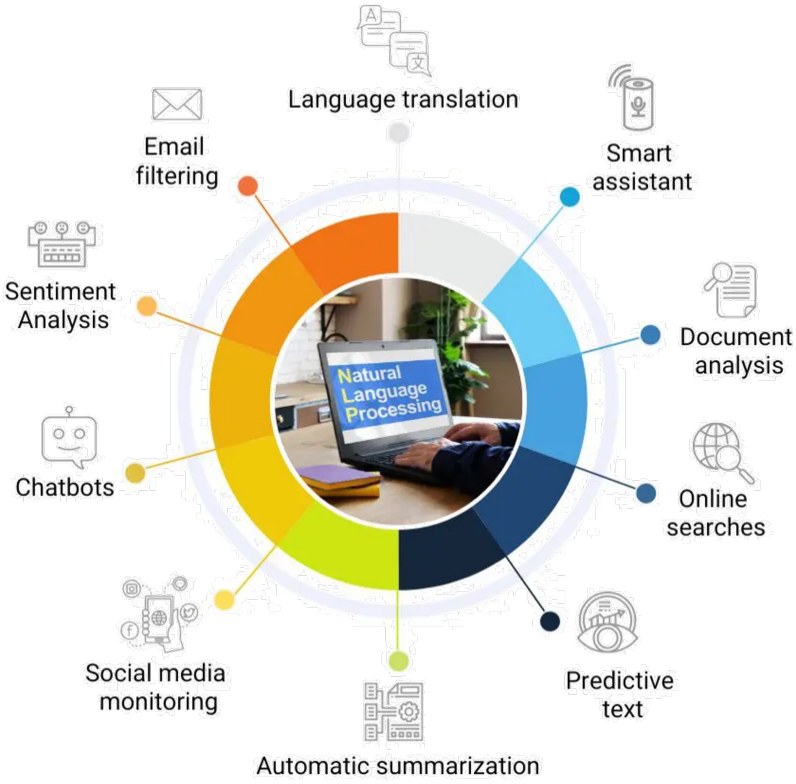
- Focused on the **study of language**
- Explores the **structure**, **meaning**, and **use** of language
- May employ computational methods and tools as part of **computational linguistics**

NLP:

- Focused on providing **computational capabilities** that utilize **human language**
- Designs and implements algorithms to **understand** and **generate** human language
- Applies **results from linguistics** to develop **practical applications**

Applications of Natural Language Processing

NLP Killer Applications



Applications by Business Sector

Healthcare:

- Process and interpret patient data, including medical records, to assist in diagnosis, treatment plans, and patient care
- Extract information from unstructured data

Finance:

- Analyze market sentiment, managing risk, detecting fraudulent activities
- Generate insights from financial reports and news

E-commerce and Retail:

- Personalized recommendations, improved search functionalities, and customer service chatbots
- Sentiment analysis to gauge customer satisfaction and market trends

Legal:

- Automate document analysis, aiding in legal research
- Streamlining the review process for contracts and legal documentation

Applications by Business Sector

Customer Service:

- Automate responses, guide users, and analyze feedback, improving efficiency

Education:

- Automatic grading, provision of learning tools
- Summarization and generation of educational materials

Automotive:

- Intelligent navigation systems and voice-activated controls

Technology:

- Assists in software development by generating code snippets and completing code
- Enhances code quality through automated reviews and suggestions

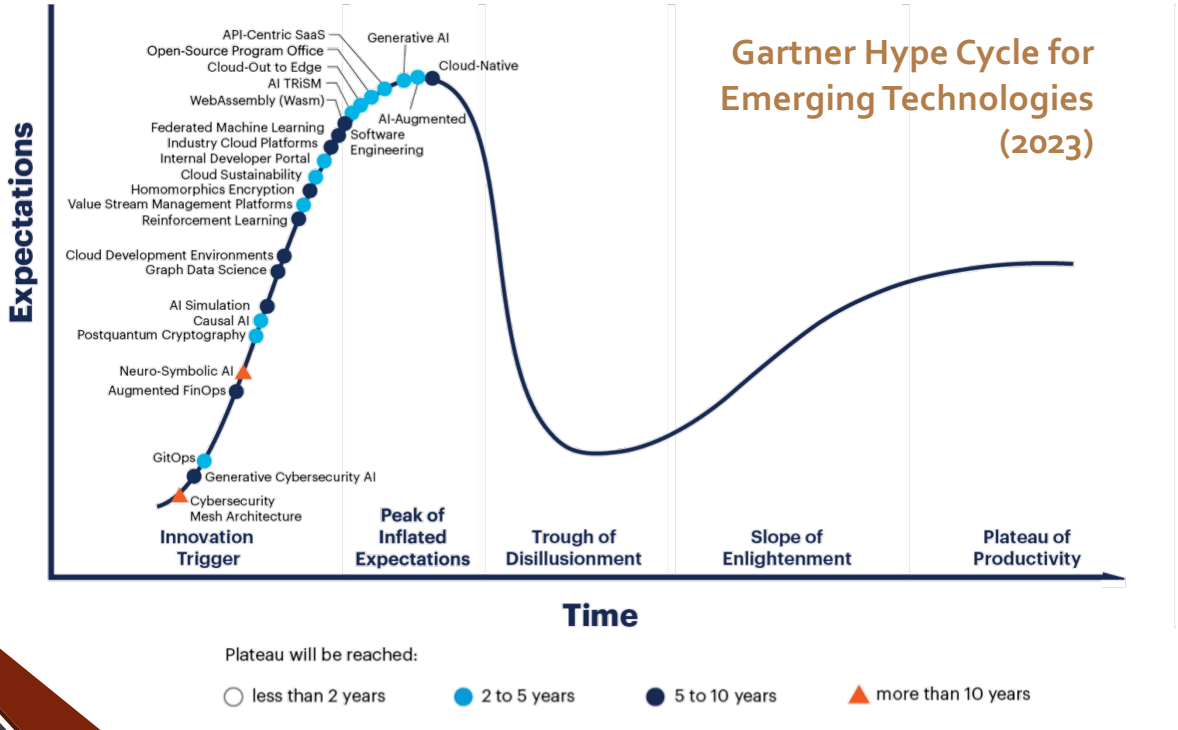
Media and Entertainment:

- Assist in generating scripts, articles, and creative writing
- Enhance user engagement with interactive storytelling and personalized media experiences

Many Other Applications…

Search	Web	Documents	Autocomplete
Editing	Spelling	Grammar	Style
Dialog	Chatbot	Assistant	Scheduling
Writing	Index	Concordance	Table of contents
Email	Spam filter	Classification	Prioritization
Text mining	Summarization	Knowledge extraction	Medical diagnoses
Law	Legal inference	Precedent search	Subpoena classification
News	Event detection	Fact checking	Headline composition
Attribution	Plagiarism detection	Literary forensics	Style coaching
Sentiment analysis	Community morale monitoring	Product review triage	Customer care
Behavior prediction	Finance	Election forecasting	Marketing
Creative writing	Movie scripts	Poetry	Song lyrics

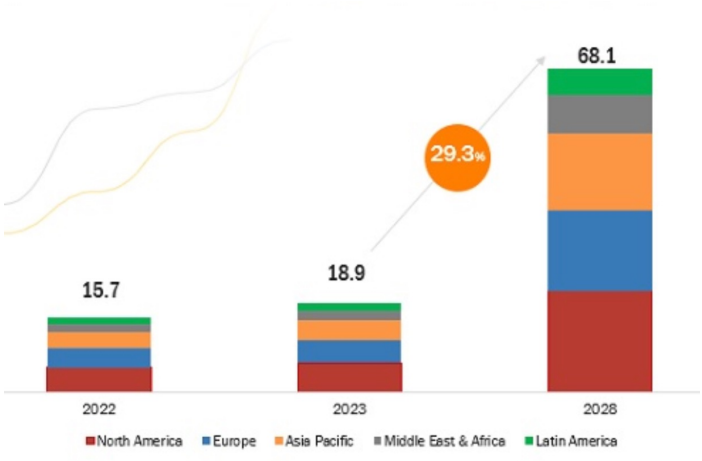
Hype Cycle



NLP Market

NLP is a **promising career option**

- **Growing demand for NLP applications**
- Projected **employment growth of 22%** between 2020 and 2030



NLP market global forecast in USD Billions

History of NLP

First Steps of NLP

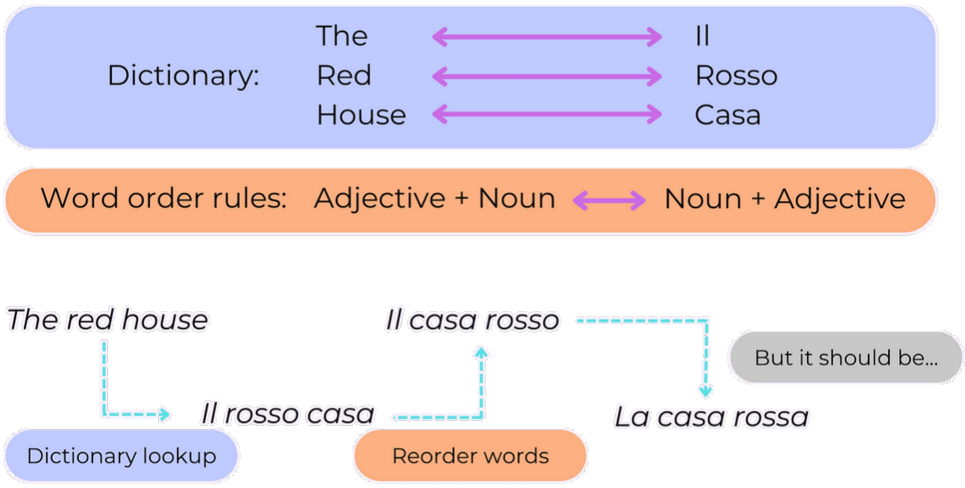
NLP has had a **history of ups and downs**

- Influenced by the **growth of computational resources** and **changes in approaches**

1950's and 1960's

- The first application that sparked interest in NLP was **machine translation**
- The first machine translation systems used **dictionary lookup** and **basic word order rules** to produce translations
- The 1950s saw a **lot of excitement**: researchers predicted that machine translation can be solved in **3 years or so**

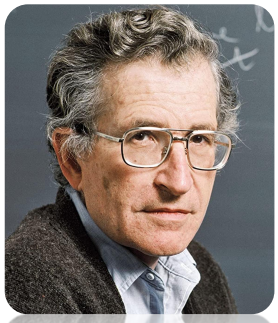
Machine Translation in 50s



How to deal with **language ambiguity**?

Generative Grammars

1957: Chomsky's Generative Grammar



- A system of rules for **generating all possible sentences** in a language
- Enabled prediction of **grammatical correctness**
- Understanding of **language structure**
- **Influenced research** in machine translation

1966: The Reality Check

- Early translation systems **fell short in effectiveness**
- Limited by their **inability to handle the ambiguity and complexity** of natural language

ALPAC Report

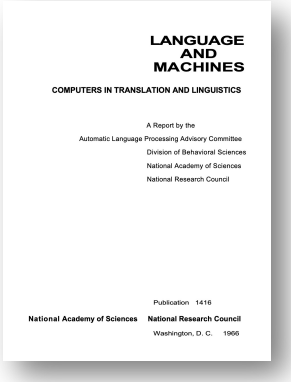
Automatic Language Processing Advisory Committee

- Established to assess advancements in **computational linguistics**

The **1966 ALPAC report** recommended **halting research into machine translation**

- Shift focus from developing **end-to-end machine translation** systems to enhancing **tools that assist human translators**
- It significantly impacted NLP and AI research, contributing to the **first AI winter**

https://www.mt-archive.net/50/ALPAC-1966.pdf



LANGUAGE AND MACHINES

COMPUTERS IN TRANSLATION AND LINGUISTICS

A Report by the
Automatic Language Processing Advisory Committee
Division of Behavioral Sciences
National Academy of Sciences
National Research Council

Publication 1416

National Academy of Sciences National Research Council
Washington, D. C. 1966

ELIZA

A pioneering conversational agent

- Created by **Joseph Weizenbaum** in the 1960s
- Designed to simulate a **conversation between a psychotherapist and a patient**

Features and Limitations:

- Demonstrated the **potential of computer-based conversation**
- Utilized **pattern matching** and **substitution** to generate responses
- Limited in handling **complex conversations**
- Could not maintain **context** beyond a few exchanges
- Often produced **irrelevant or repetitive responses**

ELIZA

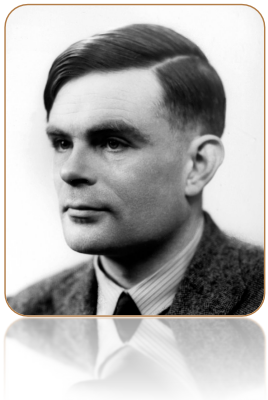
```
Welcome to
                EEEEEE  LL      IIII   ZZZZZZ   AAAAA
                EE      LL       II       ZZ   AA   AA
                EEEEE   LL       II     ZZZ    AAAAAAA
                EE      LL       II    ZZ      AA   AA
                EEEEEE  LLLLLL  IIII ZZZZZZ    AA   AA

  Eliza is a mock Rogerian psychotherapist.
  The original program was described by Joseph Weizenbaum in 1966.
  This implementation by Norbert Landsteiner 2005.


ELIZA: Is something troubling you ?
YOU:   Men are all alike.
ELIZA: What is the connection, do you suppose ?
YOU:   They're always bugging us about something or other.
ELIZA: Can you think of a specific example ?
YOU:   Well, my boyfriend made me come here.
ELIZA: Is it important to you that your boyfriend made you come here ?
YOU:   He says I'm depressed much of the time.
ELIZA: I am sorry to hear that you are depressed.
YOU:   It's true. I am unhappy.
ELIZA: Can you explain what made you unhappy ?
YOU:
```

https://psych.fullerton.edu/mbirnbaum/psych101/eliza.htm

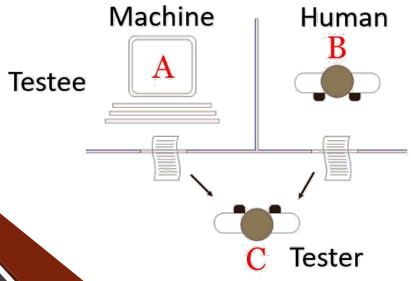
The Touring Test

*I propose to consider the question: **can machines think?** … We can only see a **short distance ahead**, but we can see **plenty there that needs to be done***

*Alan Touring, Computing Machinery and Intelligence, **1950***



Turing Test aka The Imitation game:



- A **human**, a **computer**, an **interrogator** in a different room communicate via **written messages**
- The interrogator should classify the **human** and the **machine**

The Touring Test

Capabilities for **passing the Turing Test**

- **Natural Language Understanding** to interpret user input
- **Knowledge Representation** to draw on relevant information
- **Automated Reasoning** to generate appropriate and logical responses
- **Natural Language Generation** to produce human-like textual responses
- **Context Management** to maintain and utilize context across multiple exchanges in a conversation
- **Adaptability and Learning** to adapt responses based on user behavior and feedback

The Touring Test

Successes with Turing test

- A (controversial) **success in 2014**: a chatbot mimicking the answer of a 13 years old boy
- Since then, other (controversial) successes

Limitations of Turing Test

- **Not reproducible**
- Is **emulating humans** necessary for achieving intelligence?
- Many AI researchers have shifted focus to **other benchmarks**
- **Less commonly used today**

Raise of Symbolic Approaches

1970's and 1980's:

- Programmers started creating **structured representations of real-world information** for computer understanding (**ontologies**)
- Complex **rule-based systems** were developed for various NLP tasks, including parsing, morphology, semantics, reference, …
- Main **applications** were:
  - **Expert Systems**: mimicked human expertise in specific domains
  - **Information Retrieval**: enhanced search and data extraction
- Main **limitations** were:
  - **Flexibility**: challenges in adapting to new or ambiguous contexts
  - **Scalability**: difficulty handling large-scale or diverse data

Statistical Revolution

1990's:

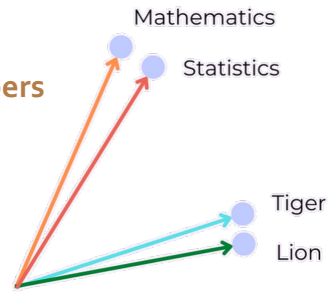
- The **computing power** increased substantially
- **Statistical models** with simple representations started to outperform complex hand-coded linguistic rules
  - Learn **patterns from data**
  - Can handle **variations** and **complexities** in natural language
  - **Large corpora** became essential
- **Long Short-Term Memory (LSTM) networks** was invented by **Hochreiter** and **Schmidhuber** in 1997

Whenever I fire a linguist, our machine translation performance improves (Fred Jelinek, IBM)

Advances in NLP

2000's

- Increased Use of **Neural Networks**
- Introduction of **Word Embeddings**
  - Words are represented as **dense vectors of numbers**
  - Words with **similar meanings** are associated with similar vectors
  - Early algorithms struggled to **efficiently learn** these representations



2006: launch of Google Translate

- The **first commercially successful NLP system**
- Utilized **statistical models** to automatically translate documents

Deep Learning Era

2010's:

- **LSTM** and **CNN** became widely adopted for NLP
- The availability of **large text corpora** enabled the training of increasingly complex models

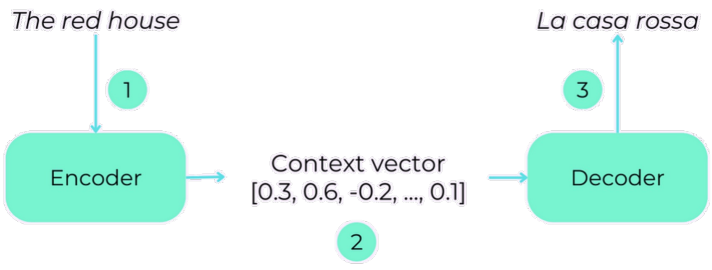
Word2Vec (2013):

- ***Efficient Estimation of Word Representations in Vector Space***
- The first algorithm to **efficiently learn word embeddings**
- Enables **semantic operations with word vector**
- Paved the way for more advanced models such as **GloVe**, **fastText**, **ELMo**, **BERT**, **COLBERT**, **GPT**, ...

Deep Learning Era

Sequence-to-Sequence Models (2014):

- Introduction of the encoder-decoder architecture:
  - **Encoder:** Encodes the input into a **context vector**
  - **Decoder:** Decodes the output from the context vector
- Useful for **automatic translation**, **question answering**, **text summarization**, **text generation**, ...



Virtual Assistants

A Virtual Assistant performs a range of tasks or services based on **user input in natural language**

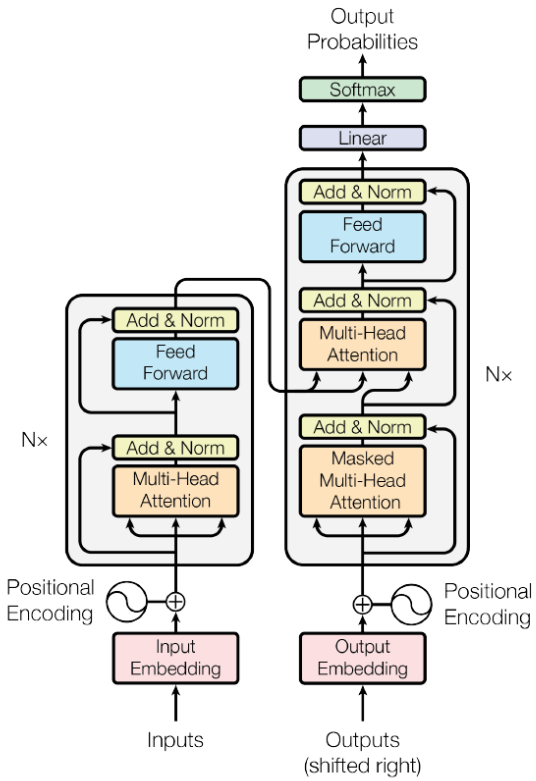
Many VA were launched in **2010's:**

- **2011: Siri** launched by Apple on iOS devices
- **2014: Cortana** introduced by Microsoft for Windows Phone
- **2014: Alexa** launched by Amazon with the Echo, pioneering voice-controlled smart home
- **2015: Google Assistant** introduced, integrating voice interaction with Android and Google Home

Deep Learning Era

Transformer (2017):

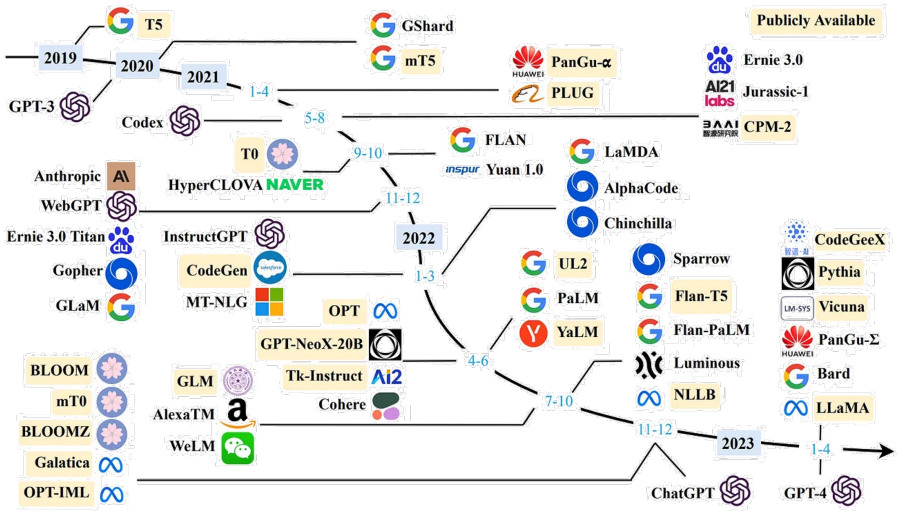
- ***Attention Is All You Need***
- Integration of **attention mechanisms**
- Allows a greater passage of information between the decoder and the encoder
- Defined and adopted by **Google** for the translator
- It remains the **dominant architecture in NLP** today



Large Language Models

After **transformers**, the next step was **scaling**…

- LLM leverage **extensive data** and **computational power** to understand and generate **human-like text**



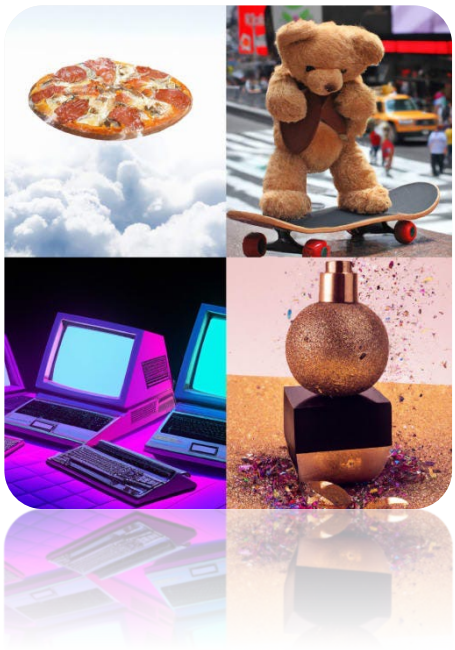
LLM Applications

- **Text Generation:** Producing articles, stories, and creative writing
- **Machine Translation:** Translating between languages
- **Chatbots:** Engaging in human-like conversations for customer support and interaction
- **Code Generation:** Generating and suggesting code snippets, completing code, and assisting with programming tasks
- **Question Answering:** Providing answers based on a given context or database
- **Text Summarization:** Condensing long documents into concise summaries
- **Writing Assistance:** Generating and completing text, improving grammar, and enhancing style

Multimodal LLM

Integrate and process **multiple types of data**

- **Image-to-Text:** generating descriptive text from images (**CLIP**)
- **Text-to-Image:** creating images based on textual descriptions (**DALL-E**)
- **Audio-to-Text:** converting spoken language into written text (**Whisper**)
- **Text-to-Audio:** composing or generating audio, such as music, from textual descriptions (**Jukebox**)



Multimodal LLM

Integrate and process **multiple types of data**

- **Video-to-Text:** Generating textual descriptions or summaries from video content (**VideoBERT**)
- **Text-to-Video:** Video content from textual descriptions (**Sora**)

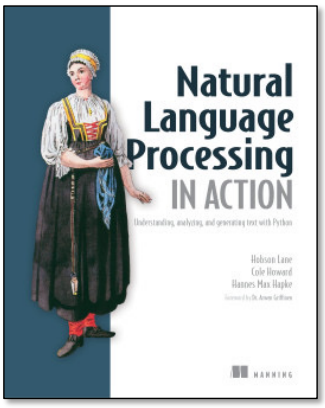
Photorealistic closeup video of two pirate ships battling each other as they sail inside a cup of coffee

References

Natural Language Processing IN ACTION

Understanding, analyzing, and generating text with Python

Chapter 1



Natural Language Processing and Large Language Models

Corso di Laurea Magistrale in Ingegneria Informatica

Lesson 1

NLP Overview

Nicola Capuano and Antonio Greco

DIEM – University of Salerno