



DS323: AI in Design (AIID)

Autumn 2023

# Week 05 Lecture 09

## AIID + Text

Wan Fang

Southern University of Science and Technology

---

# Agenda

- Text ( 通向通用人工智能之路? )
  - Intro to Natural language processing (NLP)
  - Challenges in NLP
  - Large Language Model (LLM)
- Generative Adversarial Networks (GAN)
- Practice: In-class Case Review

---

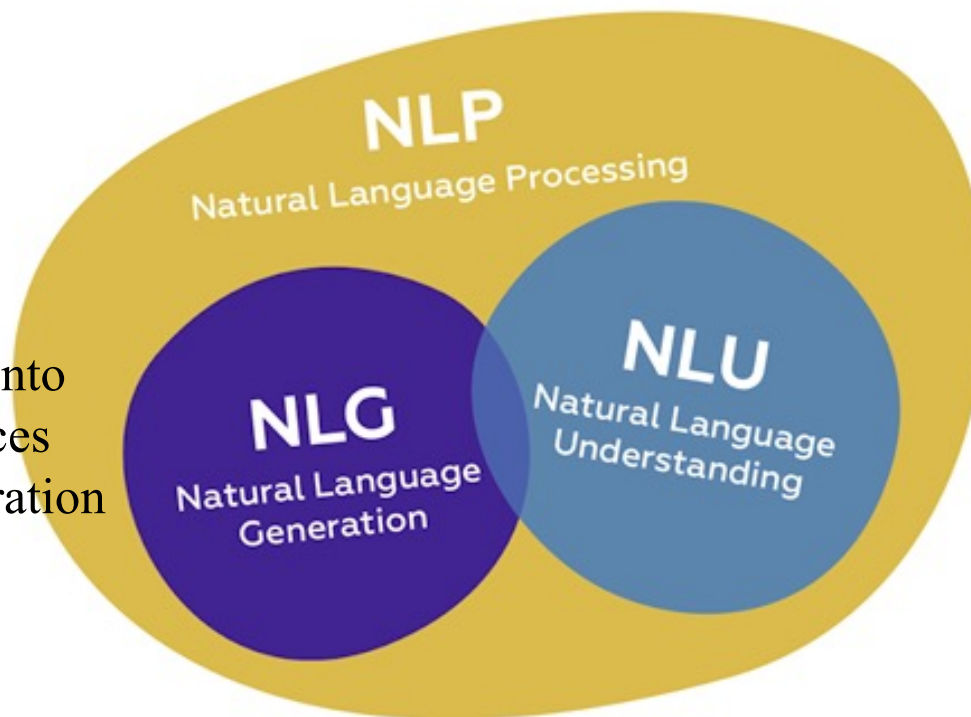
# What is NLP?

- Natural language processing (NLP) is a field of computer science, artificial intelligence, and computational linguistics concerned with the interactions between computers and human (natural) languages – Wikipedia
- Text is the largest repository of human knowledge and is growing quickly.
  - Emails, news articles, web pages, IM, scientific articles, insurance claims, customer complaint letters, transcripts of phone calls, technical documents, government documents, patent portfolios, court decisions, contracts, .....

# Fundamental NLP Tasks

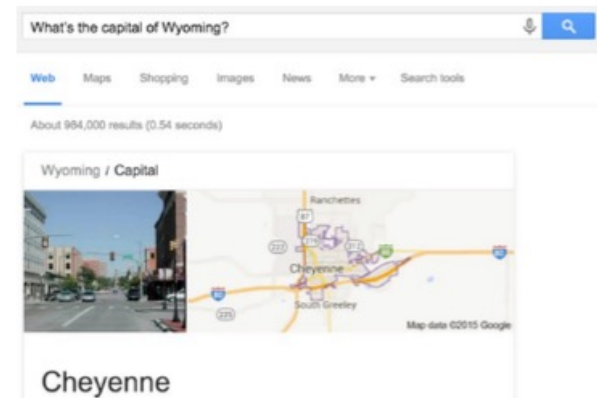
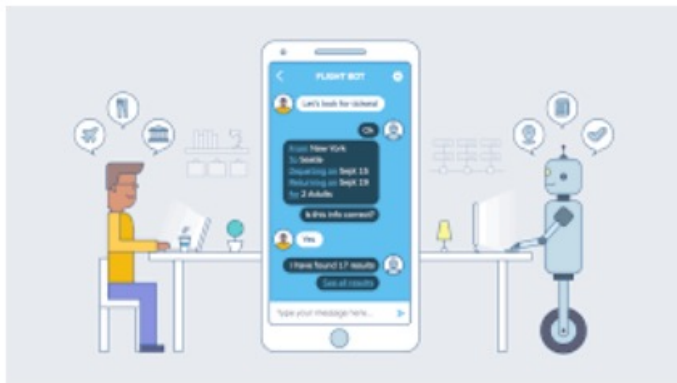
- **NLP** focuses on
  - Understanding the meaning by processing text input
  - Natural language generation

1. Convert input texts into new symbol sequences
2. “Open-ended” generation



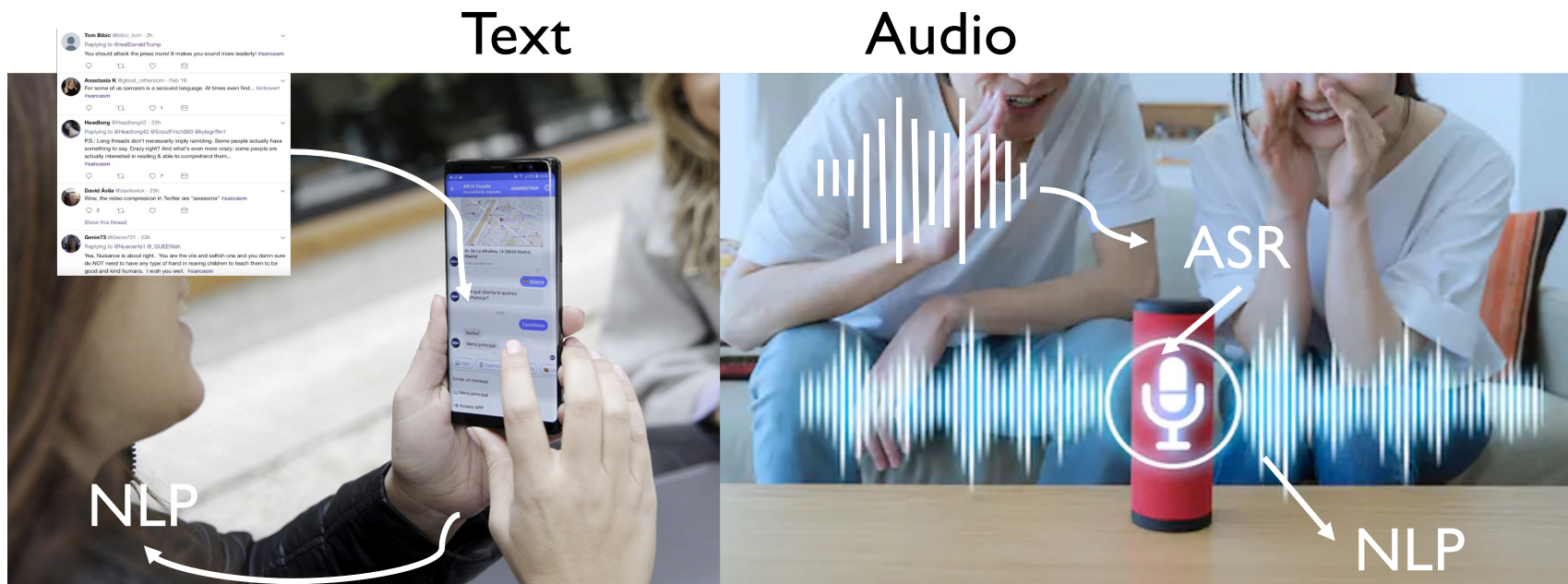
# NLP Applications

- Chatbots and Virtual Assistants: One of the most prominent applications of NLP is in the development of chatbots and virtual assistants.



# Some notes

- **Speech Recognition**, also called speech-to-text, is the task of reliably converting voice data into text data.
- **Conversational Interfaces** = Speech Recognition + NLP + Voice



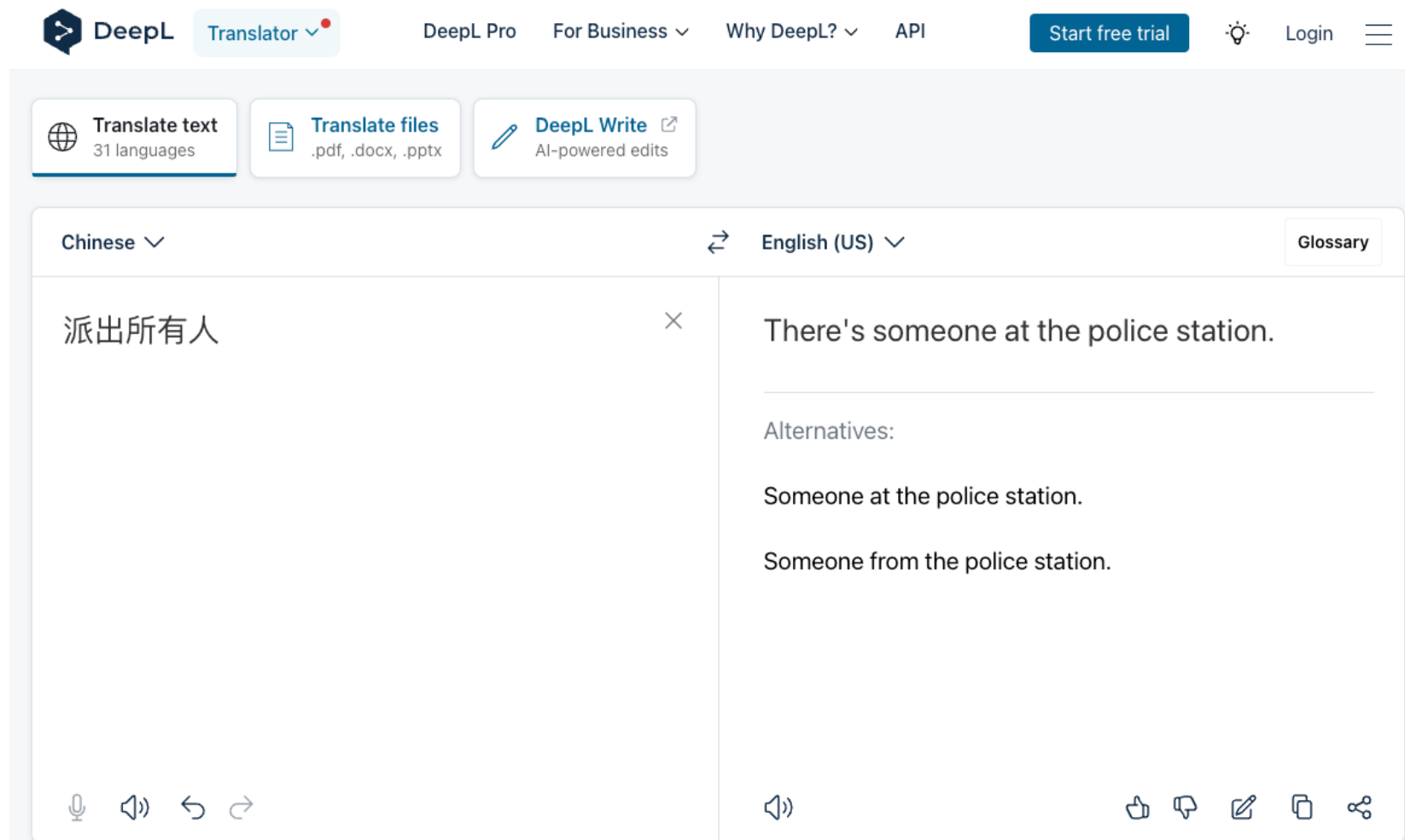
---

# NLP Applications

- **Content Generation:** From news articles to creative writing, NLP algorithms can generate human-like text content. Content creators can leverage these tools to automate routine writing tasks, freeing up time for more complex creative endeavors.

# NLP Applications

- Machine Translation



The screenshot displays the DeepL Translator web interface. At the top, the navigation bar includes the DeepL logo, a 'Translator' dropdown menu, and links for 'DeepL Pro', 'For Business', 'Why DeepL?', and 'API'. A 'Start free trial' button and 'Login' link are also present. Below the navigation bar, three main service buttons are visible: 'Translate text' (31 languages), 'Translate files' (.pdf, .docx, .pptx), and 'DeepL Write' (AI-powered edits). The main translation area is split into two columns. The left column is set to 'Chinese' and contains the text '派出所有人'. The right column is set to 'English (US)' and shows the translation 'There's someone at the police station.' Below the main translation, a section titled 'Alternatives:' lists two other possible translations: 'Someone at the police station.' and 'Someone from the police station.' At the bottom of the interface, there are icons for voice input/output, undo/redo, and other utility functions.



# NLP Applications

- **Information Extraction:** NLP techniques enable computers to extract structured information from unstructured text. This has applications in data mining, knowledge extraction, and content categorization.



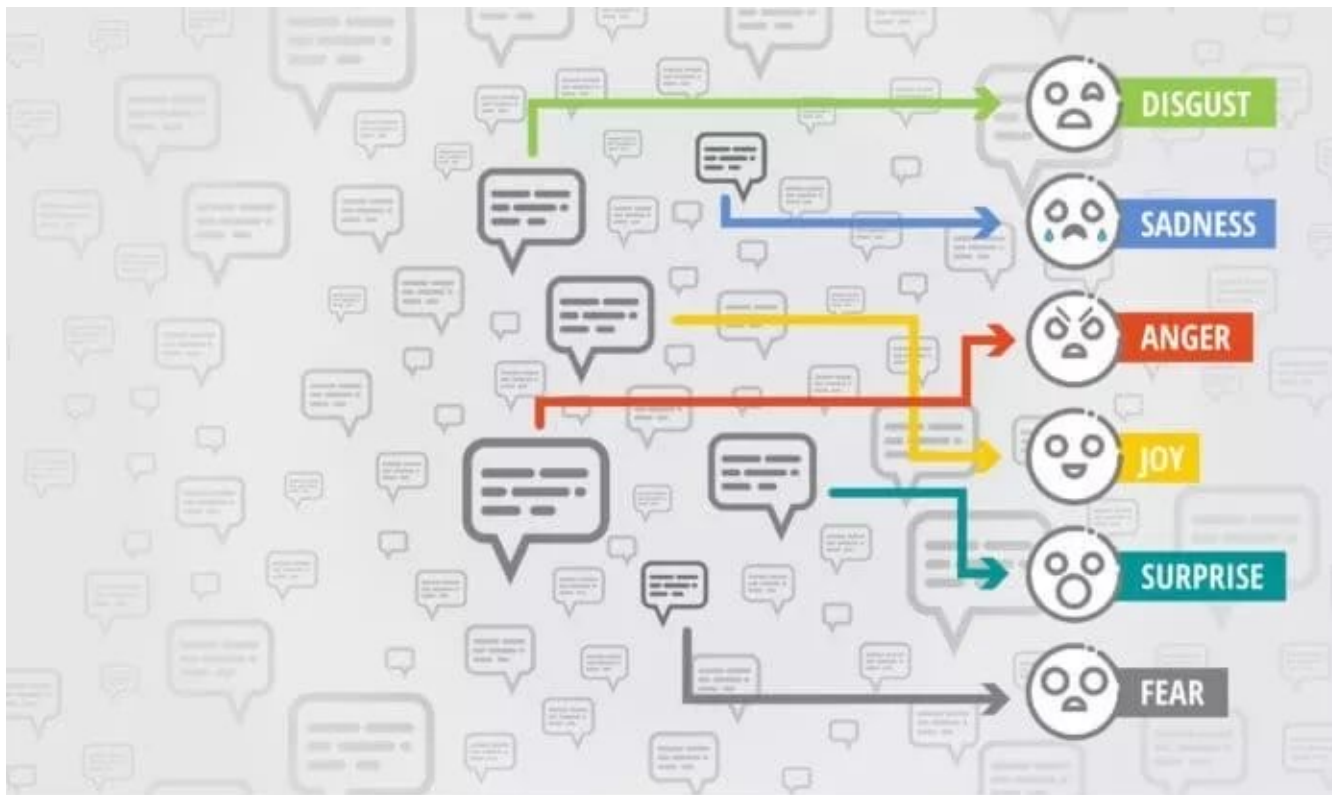
Named-entity recognition



Paragraph summarization

# NLP Applications

- **Sentiment/Opinion Analysis:** NLP has revolutionized sentiment analysis by enabling machines to gauge emotional tones in text data.



# Challenges

- Ambiguity

一语双关

谐音

反语

# Challenges

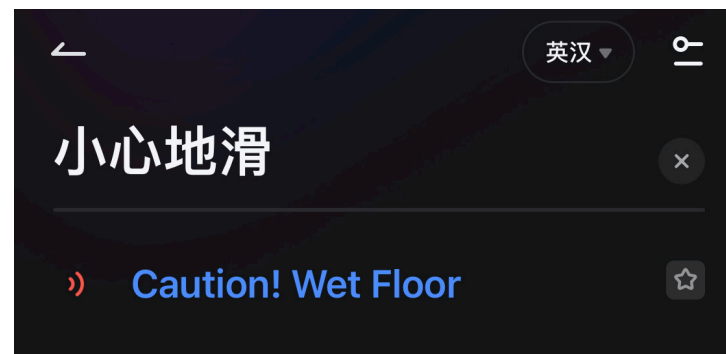
- Ambiguity
- Language is not static

## 2022年网络热梗🔥流行词

- 1、emo——抑郁了
- 2、内卷——非理性内部竞争
- 3、普信男——普通却过于自信的男人
- 4、大聪明——不灵光
- 5、YYDS——永远的神
- 6、破防——心里防线突破
- 7、瑞思拜——厉害了
- 8、集美——姐妹
- 9、999——6翻了
- 10、fyp——上热门
- 11、芭比Q了——完了完了
- 12、kpd——砍拼夕夕
- 13、你礼貌吗——字面意思

# Challenges

- Ambiguity
- Language is not static
- Language is compositional



---

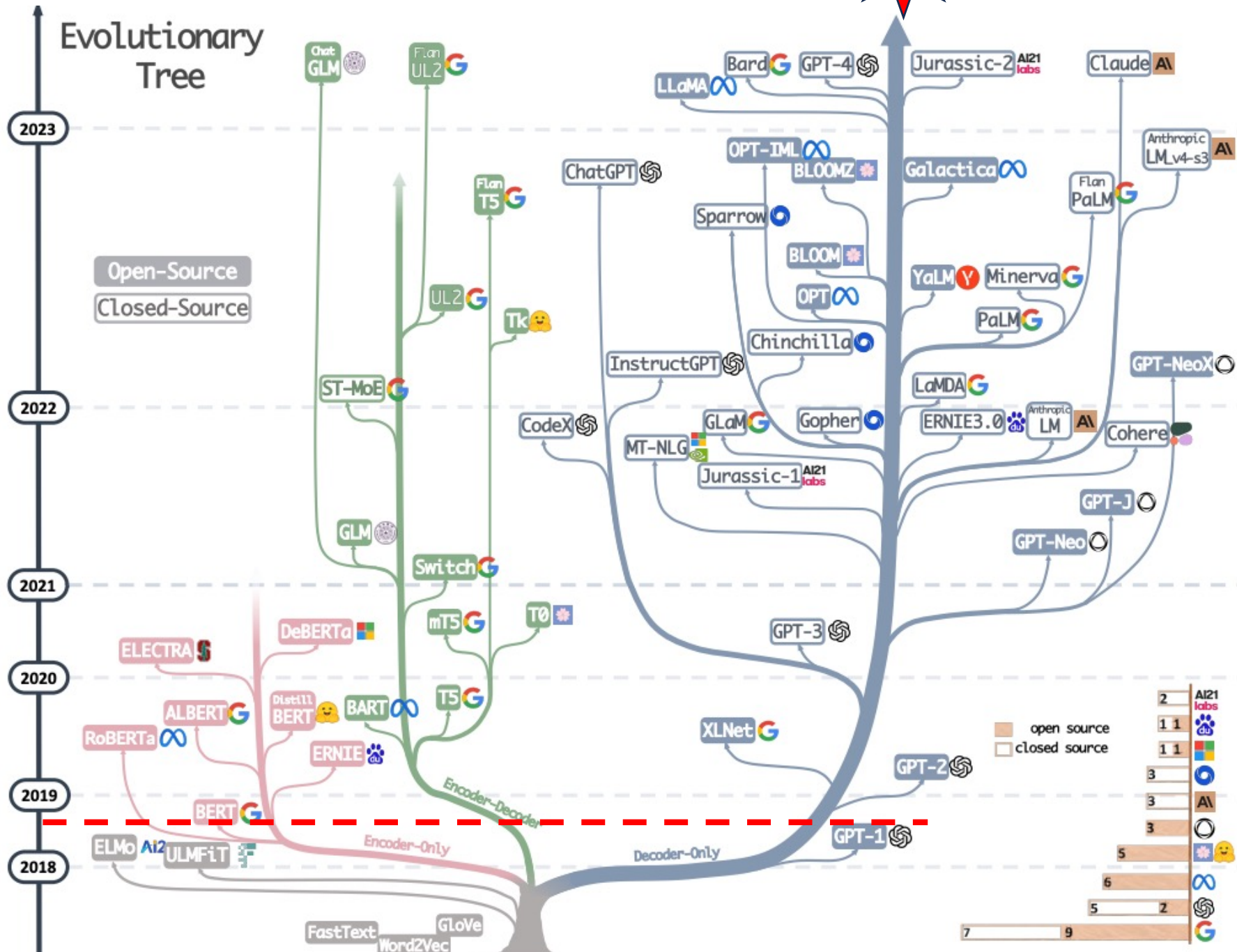
# Challenges

- Ambiguity
- Language is not static
- Language is compositional
- **Scale**
  - ChatGPT was trained on a massive corpus of text data, around 570GB of datasets, including web pages, books, and other sources. To be even more exact, 300 billion words were fed into the system.

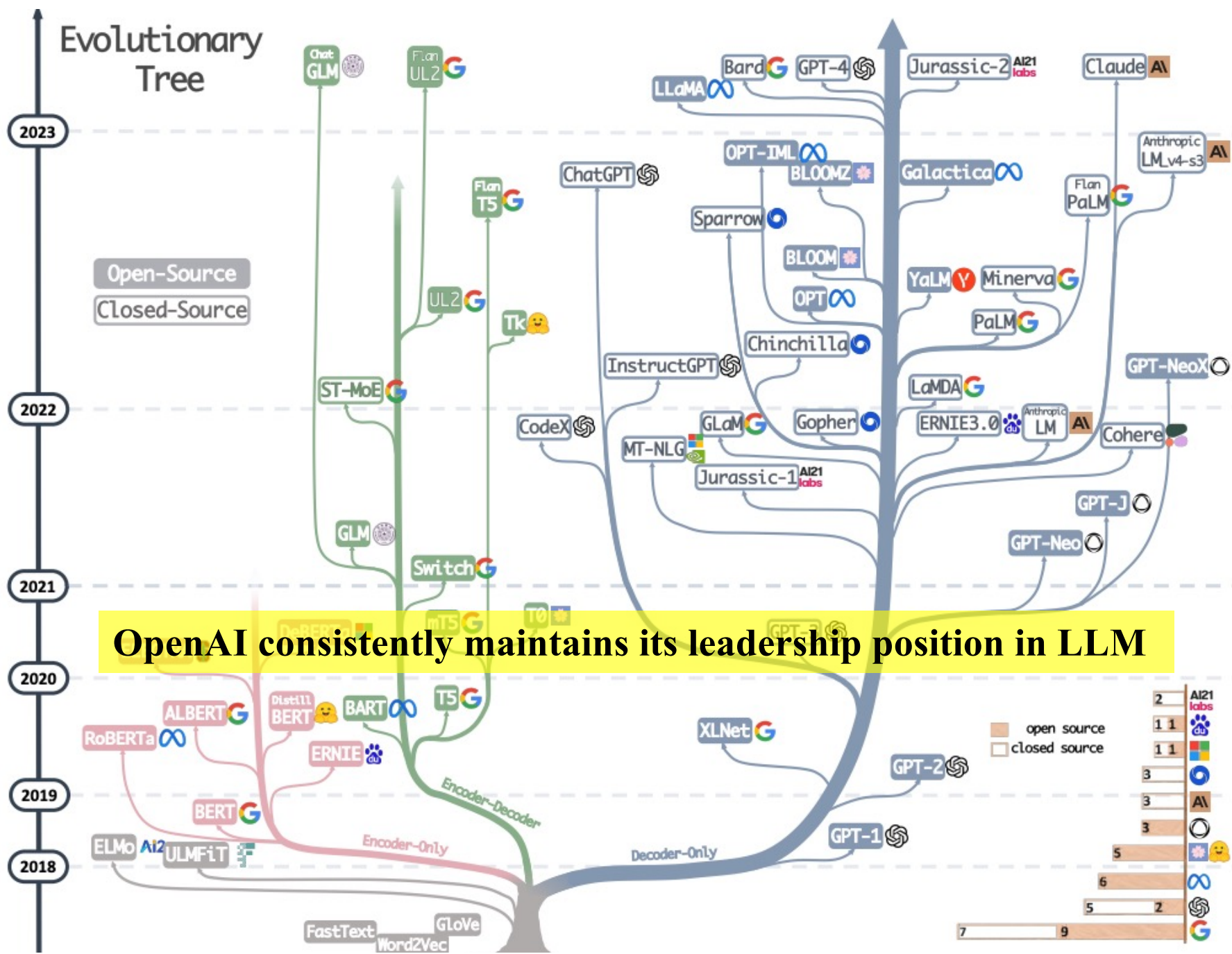
# Large Language Model (LLM)

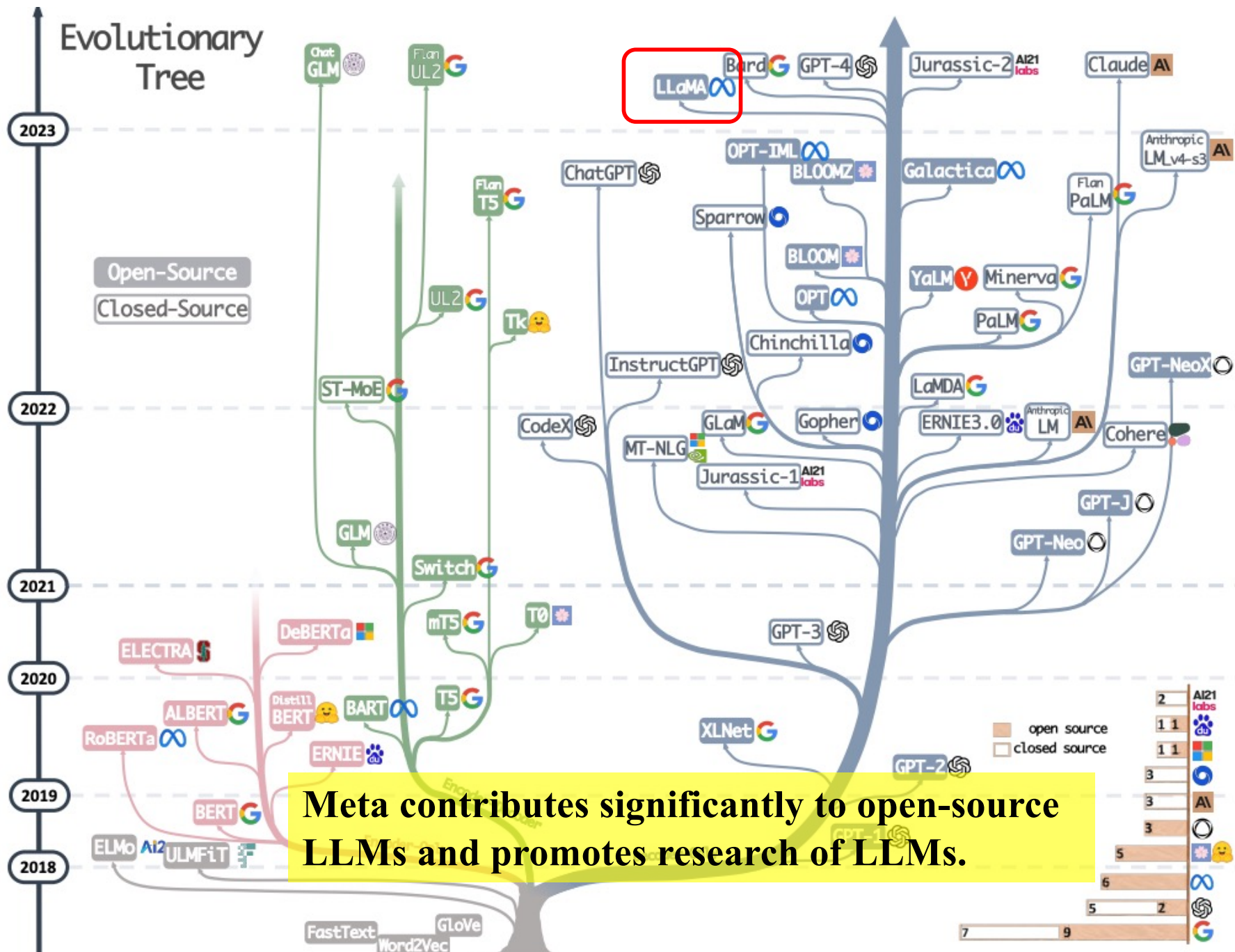
- LLM is a type of language model notable for its ability to achieve general-purpose **language understanding and generation**. Also known as **foundation model**.
- LLMs acquire these abilities by using **massive amounts of data** to learn **billions of parameters** during training and consuming large computational resources during their training and operation.

OPT	175B
BLOOM	176B
GPT 3.0	175B
LaMDA	137B
GLM	130B
YaLM	100B
LLaMA	65B
GPT-NeoX	20B
Falcon	40B
UL2	20B
鹏程·盘古α	13B

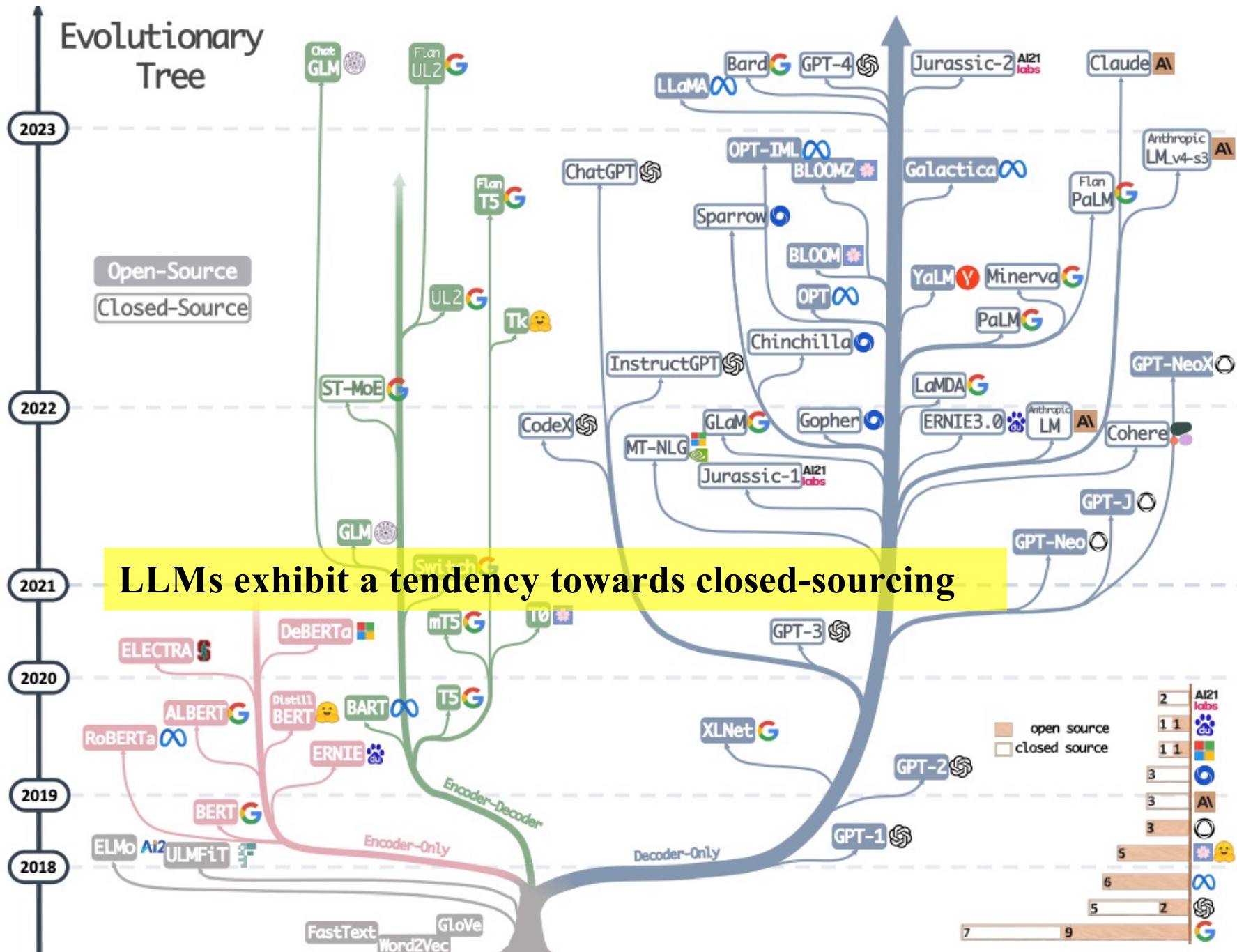


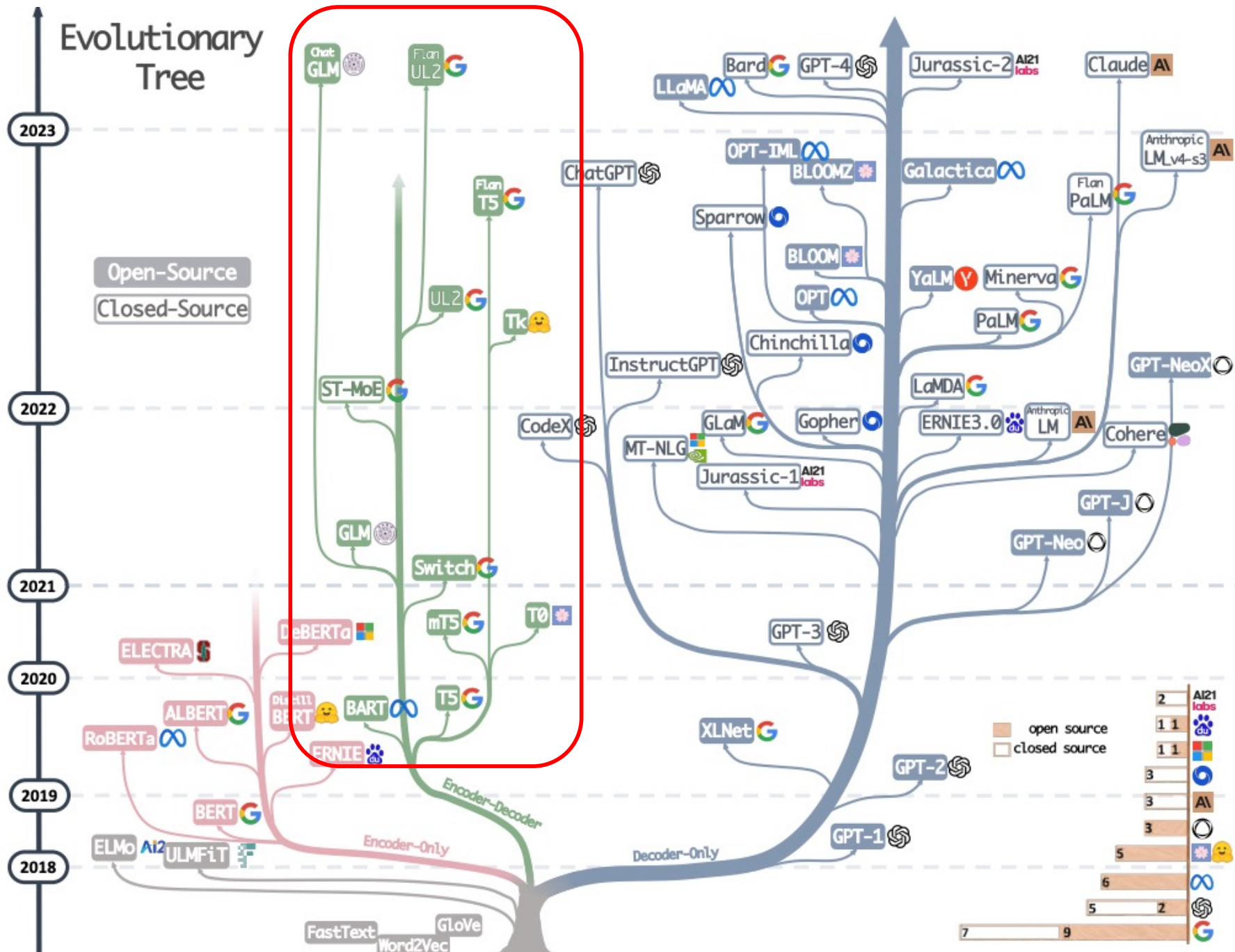






Meta contributes significantly to open-source LLMs and promotes research of LLMs.



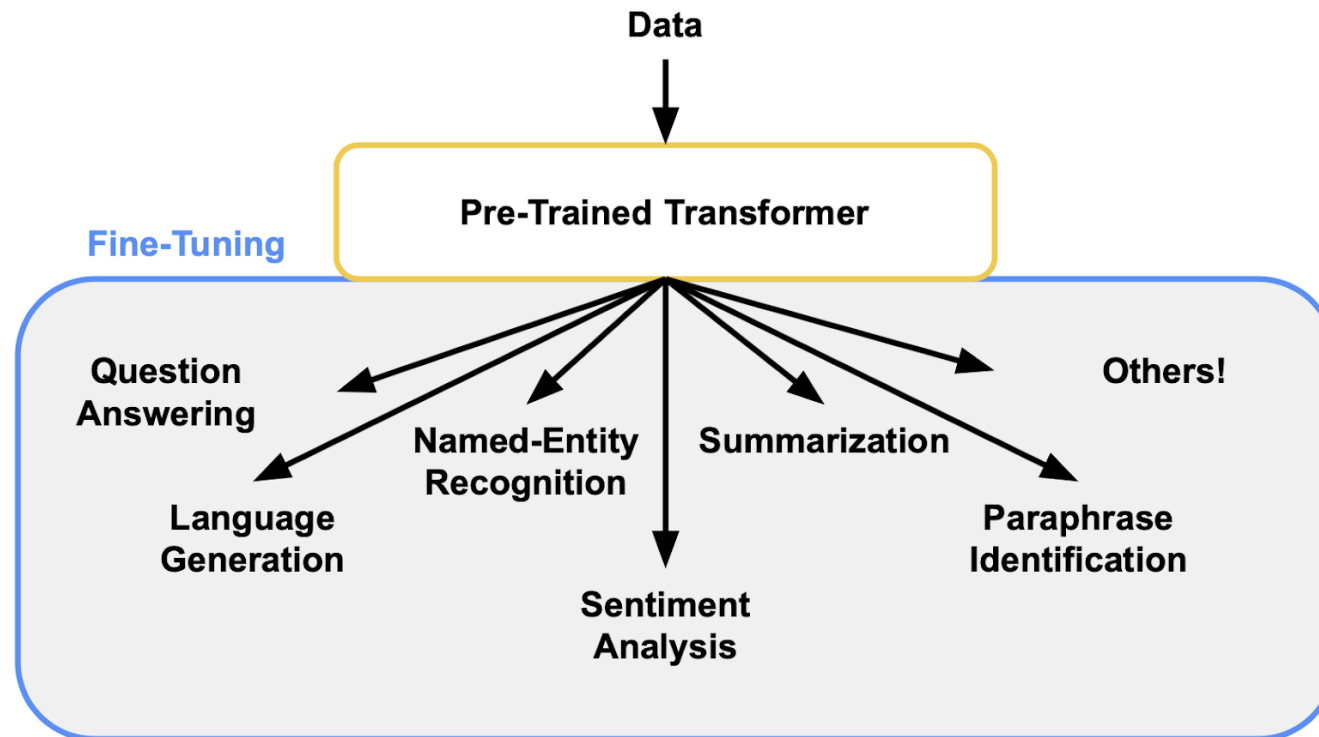


# How to use LLMs

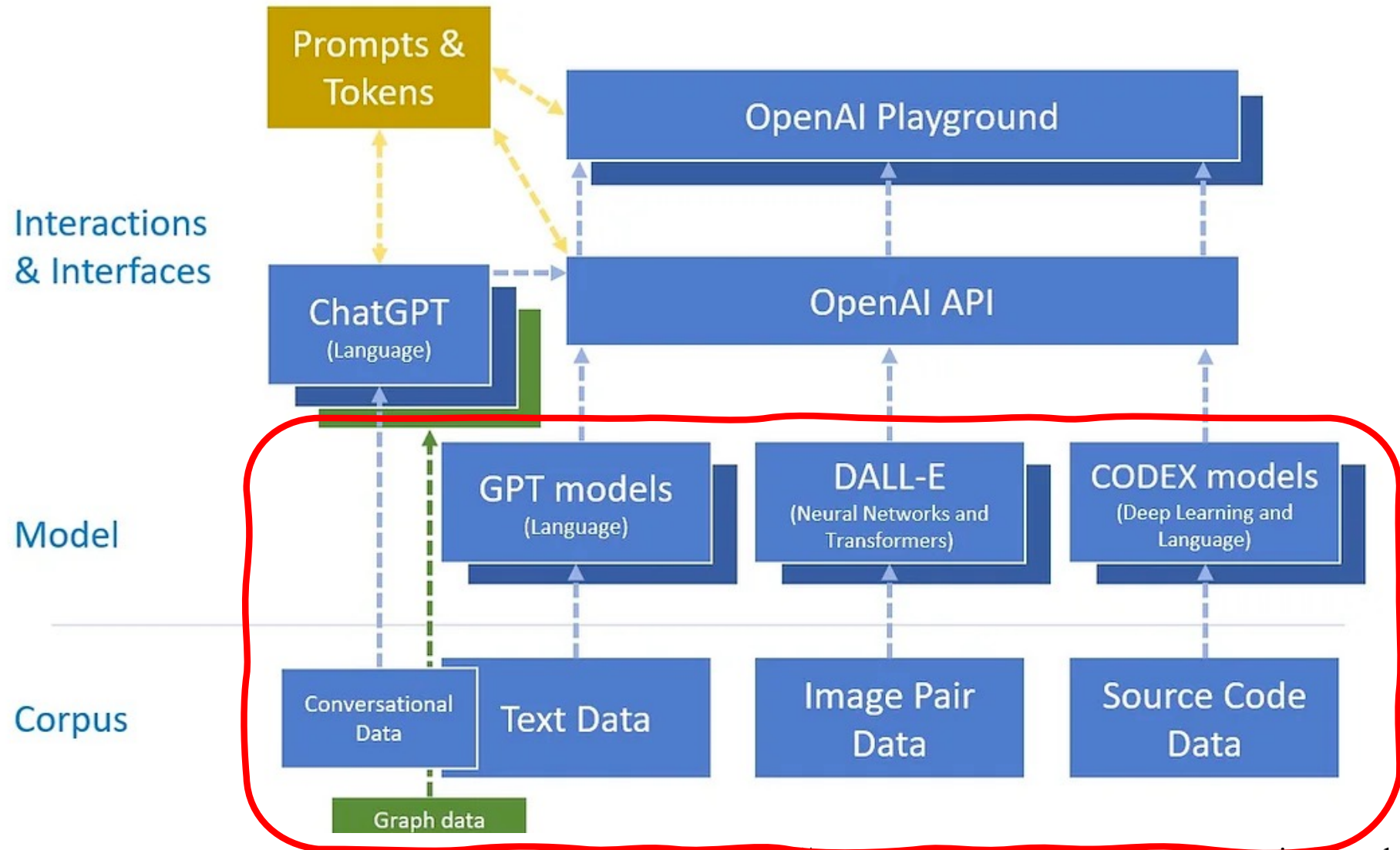
*Fine-tuned models VS In context learning*

# Fine-tuned models

- Fine-tuned models are typically smaller language models which are also pretrained and then further tuned on a smaller, task-specific dataset to optimize their performance on that task

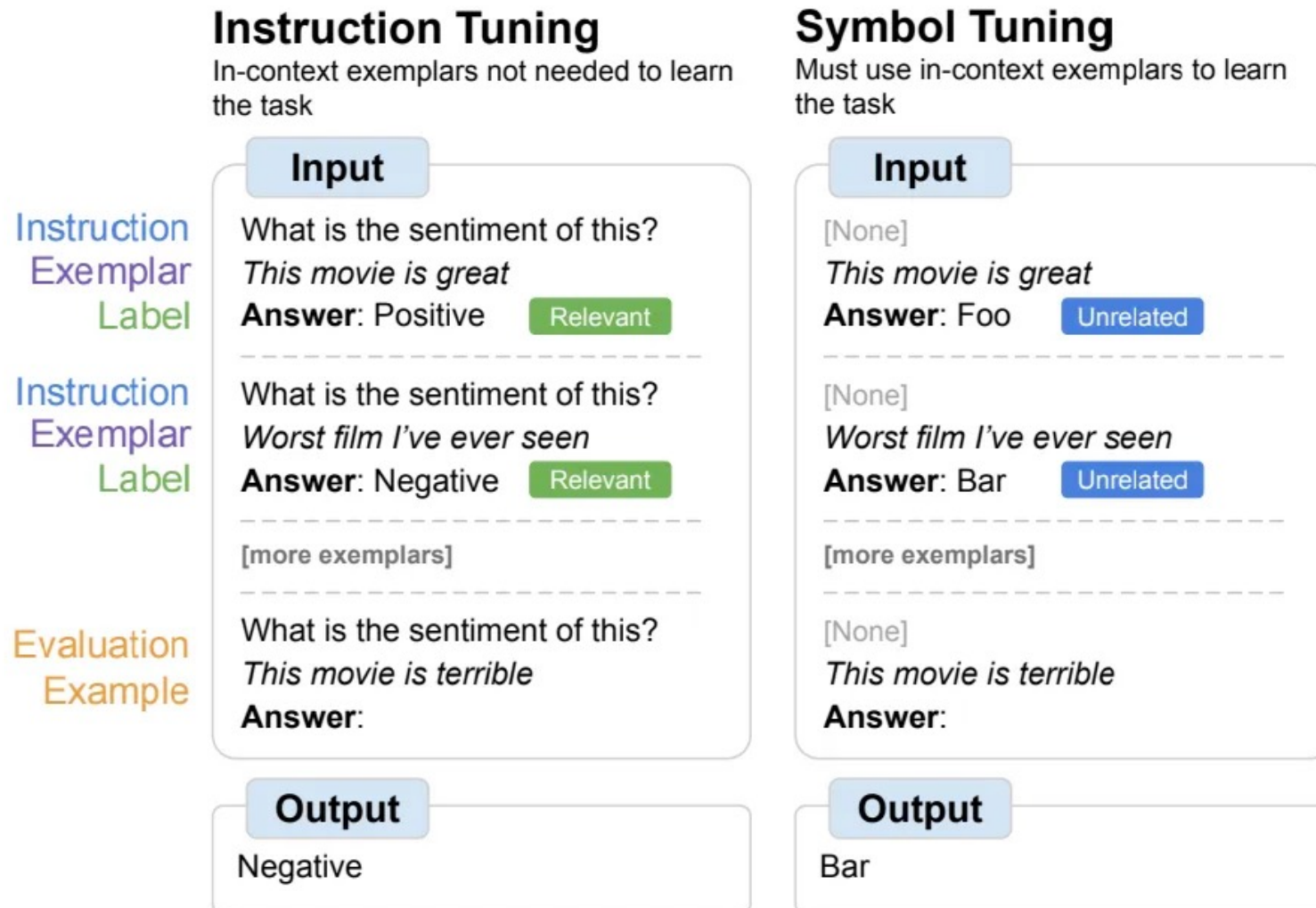


# Fine-tuned models with GPT



# In context learning

*Using LLMs directly **without tuning** on data for specific tasks*

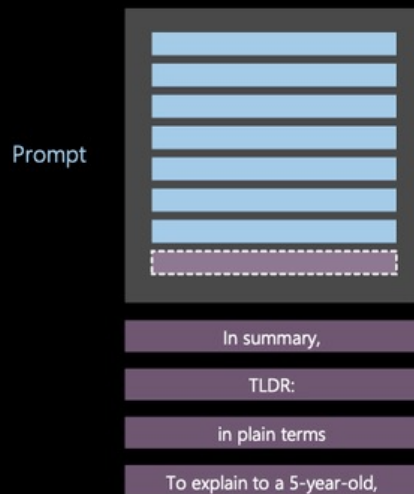




# Prompting & Fine Tuning

## 1) Prompt Engineering

### 1.1) Tuning the task description



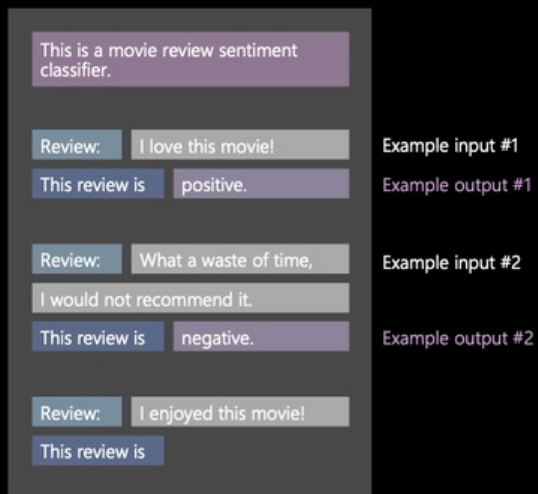
Task description

Example #1

Example #2

Current output

### 1.2) Adding examples in the prompt



This is a movie review sentiment classifier.

Review: I love this movie!

This review is positive.

Review: What a waste of time,

I would not recommend it.

This review is negative.

Review: I enjoyed this movie!

This review is

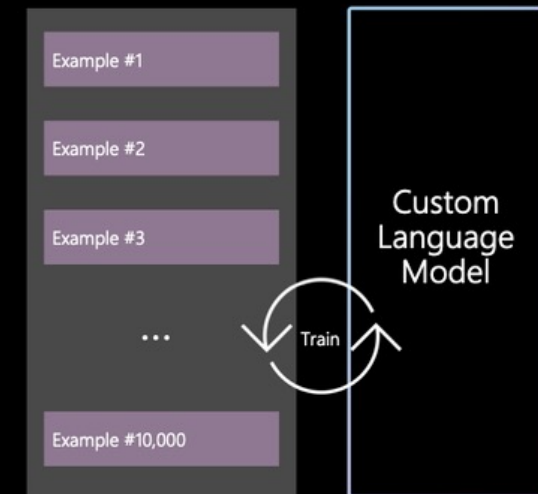
Example input #1

Example output #1

Example input #2

Example output #2

## 2) Finetune on many examples



Example #1

Example #2

Example #3

...

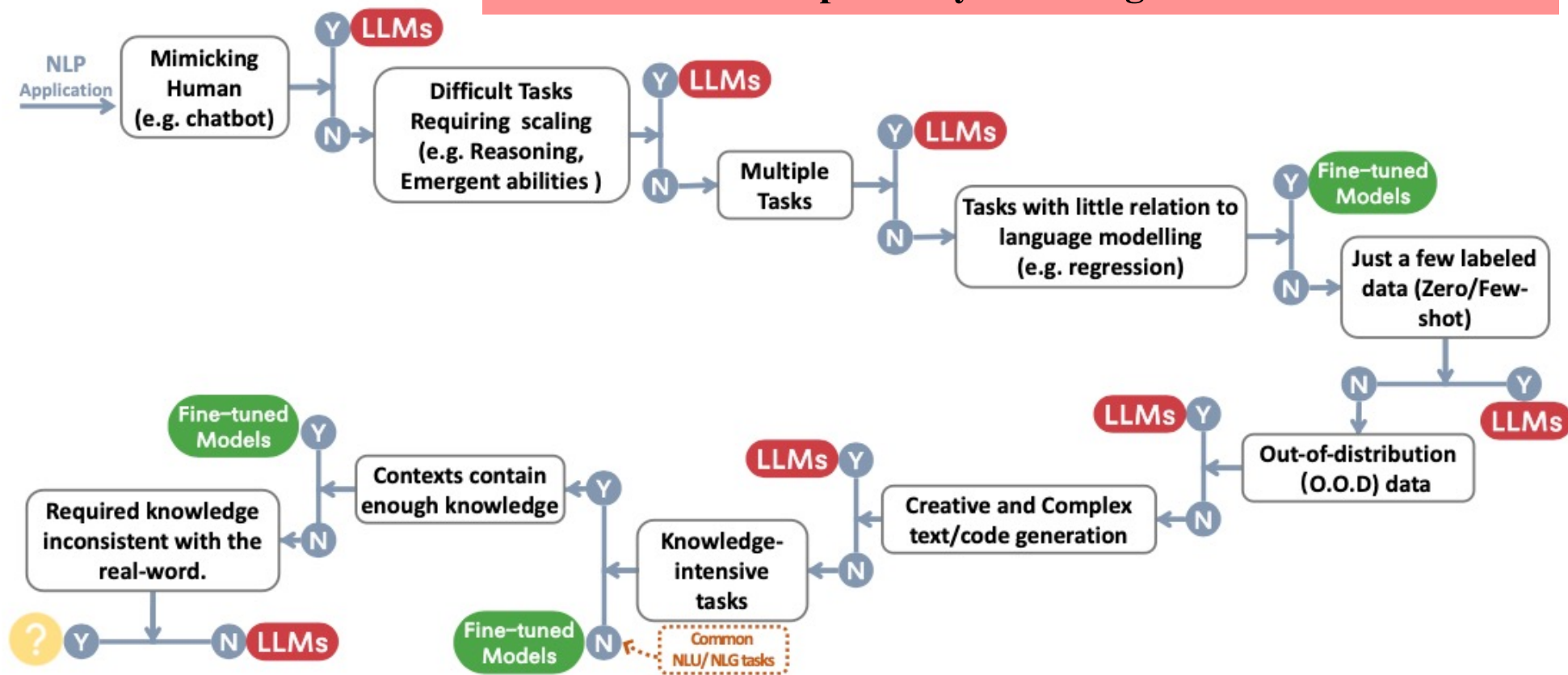
Example #10,000

Custom Language Model

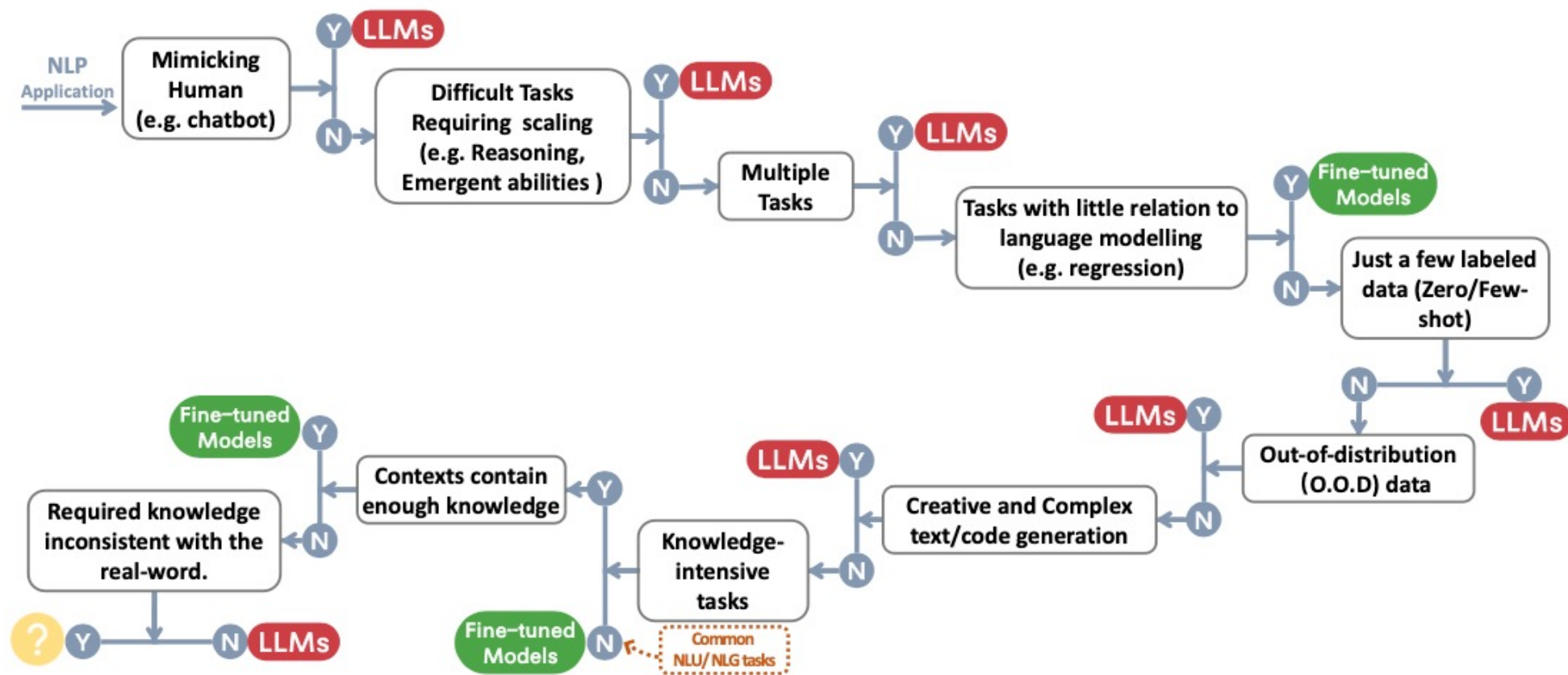
Train

# Choosing LLMs or fine-tuned models

1. Due to their strong generation ability and creativity, LLMs show superiority at most generation tasks.

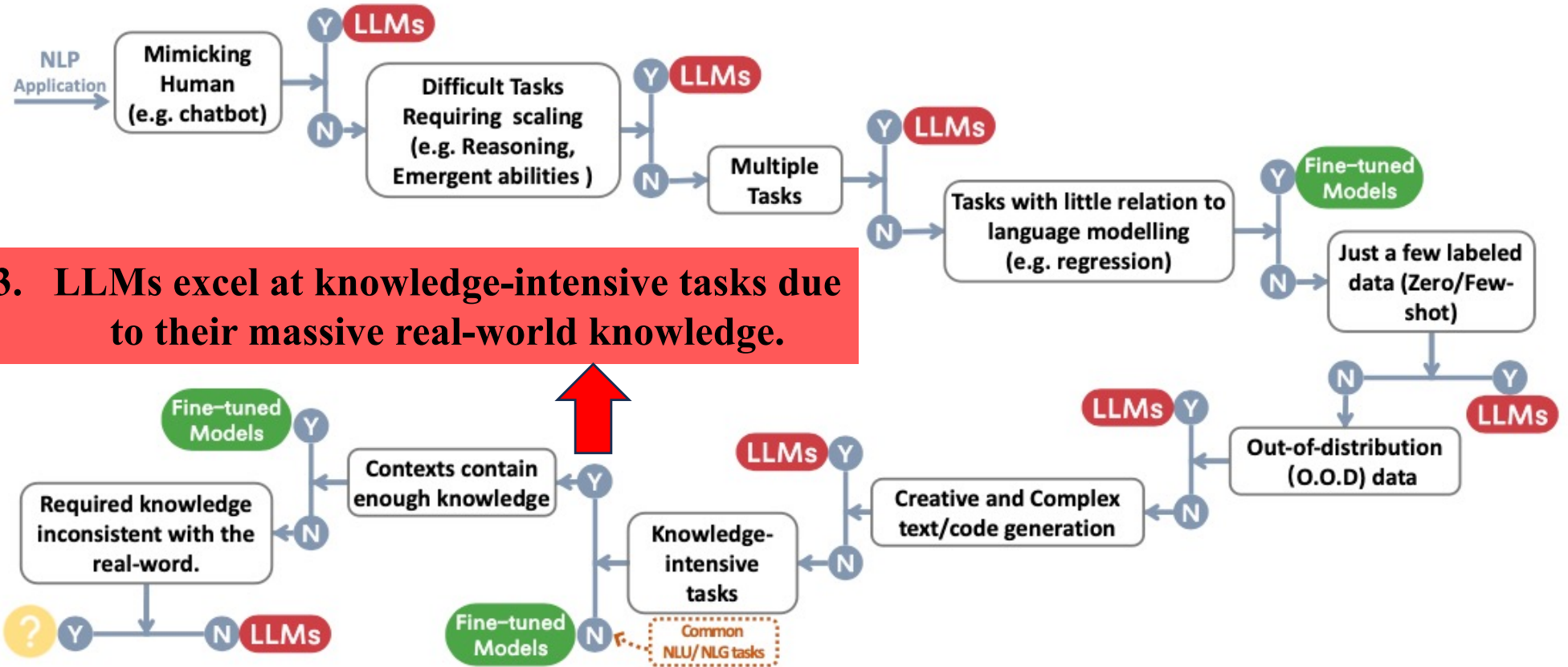


# Choosing LLMs or fine-tuned models



**2. For most traditional NLU tasks, a fine-tuned model is a better choice**

# Choosing LLMs or fine-tuned models



**3. LLMs excel at knowledge-intensive tasks due to their massive real-world knowledge.**

# Abilities Regarding Scaling

- With the exponential increase of model scales, LLMs become especially capable of reasoning like arithmetic reasoning and commonsense reasoning.
- Emergent abilities become serendipity (意外新发现) for uses that arise as LLMs scale up, such as ability in word manipulation and logical ability.
  - emergent abilities of LLMs are abilities that are not present in smaller-scale models but are present in large-scale models.
- In many cases, performance does not steadily improve with scaling due to the limited understanding of how large language models' abilities change as they scale up.

# Reasoning test with 通义千问

小爱同学?  
refuse to  
answer the  
same question

我有15个苹果，吃掉3个，还剩几个苹果

$15 - 3 = 12$

我有n个苹果，吃掉n+1个，还剩几个苹果?

$n - (n+1) = -1$

-1个苹果，这个结果合理吗?

-1个苹果在现实中是无法存在的，所以这个结果在实际情况下不合理。

重新回答这个问题：“我有n个苹果，吃掉n+1个，还剩几个苹果?”

$n - (n+1) = -1$ ，这个结果在实际情况下不合理，因为在现实中不可能有负数个苹果。

- DS323: AI in Design
- 

# Generative Adversarial Networks

# Generative Adversarial Networks

- GANs can be trained on the images of
  - humans to generate realistic faces.
  - cartoon characters for generating faces of anime characters as well as Pokemon characters.



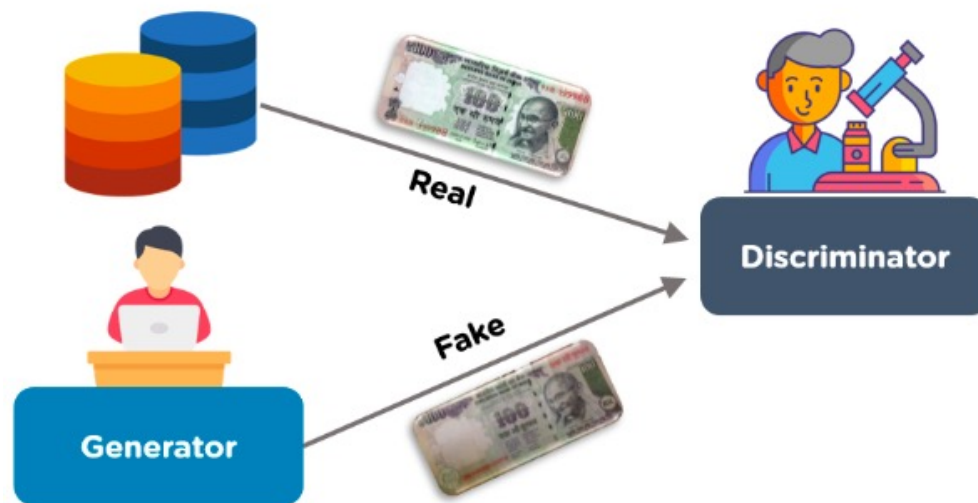


# Text to Image

Text description	This bird is red and brown in color, with a stubby beak	The bird is short and stubby with yellow on its body	A bird with a medium orange bill white body gray wings and webbed feet	This small black bird has a short, slightly curved bill and long legs	A small bird with varying shades of brown with white under the eyes	A small yellow bird with a black crown and a short black pointed beak
64x64 GAN-INT-CLS						
128x128 GAWWN						
256x256 StackGAN-v1						
256x256 StackGAN-v2						

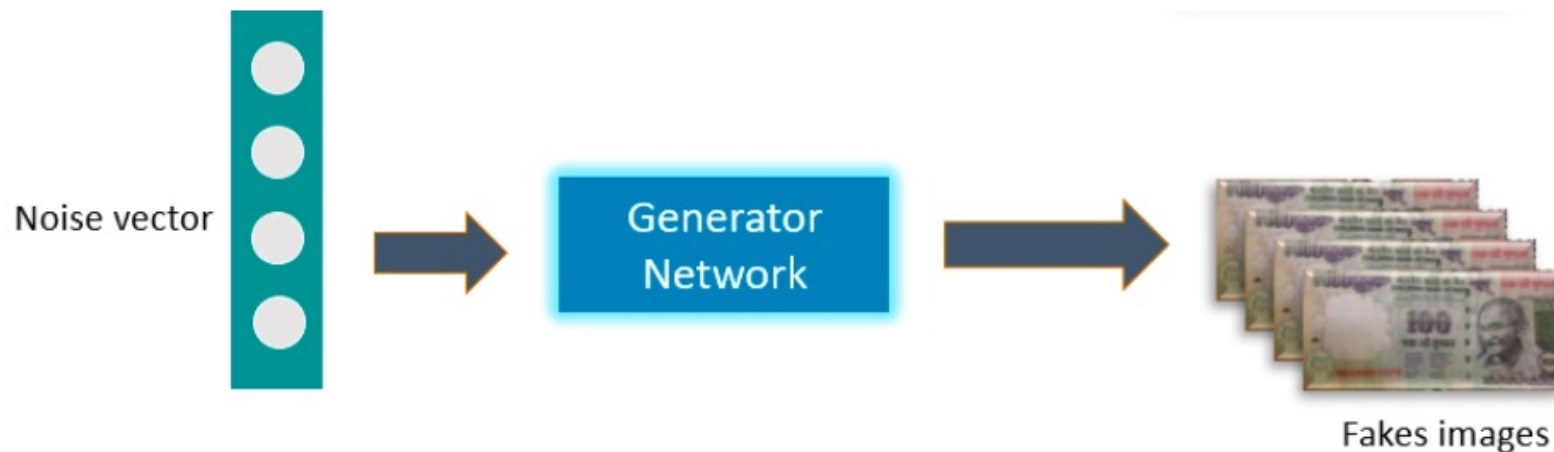
# What are Generative Adversarial Networks?

- Generative Adversarial Networks (GANs) were introduced in 2014 by Ian J. Goodfellow.
- GANs perform unsupervised learning tasks in machine learning.
- It consists of 2 models that automatically discover and learn the patterns in input data.



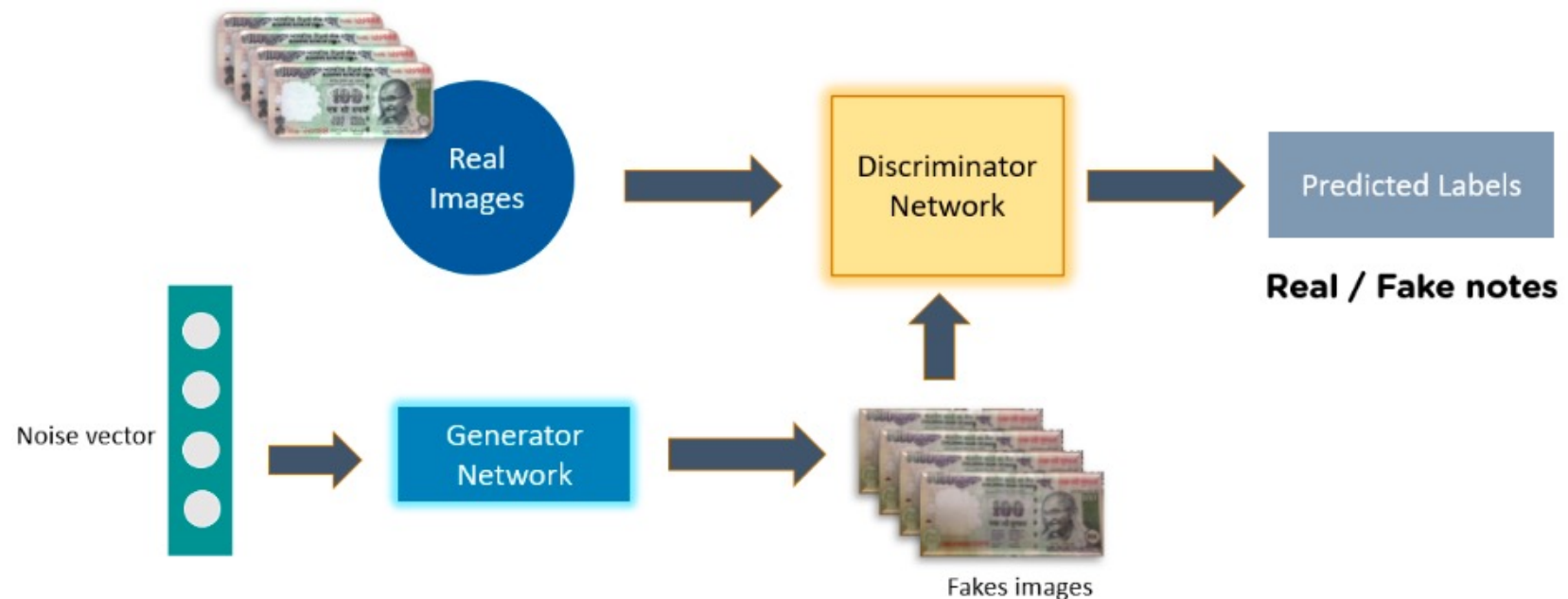
# What are Generative Adversarial Networks?

- A Generator in GANs is a neural network that creates fake data to be trained on the discriminator. It learns to generate plausible data.



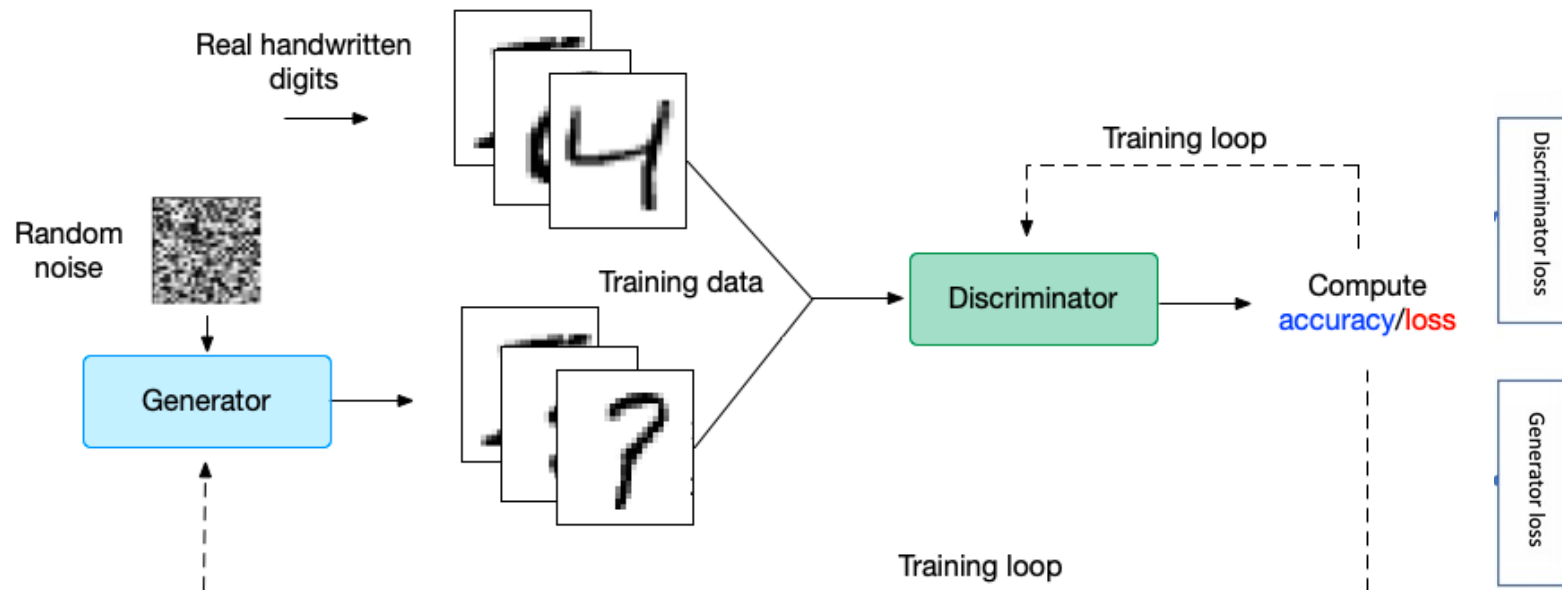
# What are Generative Adversarial Networks?

- The Discriminator is a neural network that identifies real data from the fake data created by the Generator. The discriminator's training data comes from different two sources



# Steps for Training GAN

1. Define the problem
2. Choose the architecture of GAN
3. Train discriminator on real data
4. Generate fake inputs for the generator
5. Train discriminator on fake data
6. Train generator with the output of the discriminator



# Exercise

- GPT token encoder and decoder:  
<https://observablehq.com/@simonw/gpt-tokenizer>
- 阿里云通义千问模型: <https://qianwen.aliyun.com/>
- GAN Lab: <https://poloclub.github.io/ganlab/>

## Resources

- <https://github.com/Hannibal046/Awesome-LLM>



# DS323: AI in Design (AIID)

<https://ds323.ancorasir.com/>

Autumn 2023

**Thank you~**

Wan Fang  
Southern University of Science and Technology