

# AI 101 – on what ChatGPT is(n't) Large Language Models and their Potential and Limitations for Language Learning

Daniel Hershcovich

Forum on AI and ChatGPT  
May 15<sup>th</sup>, 2023

UNIVERSITY OF COPENHAGEN





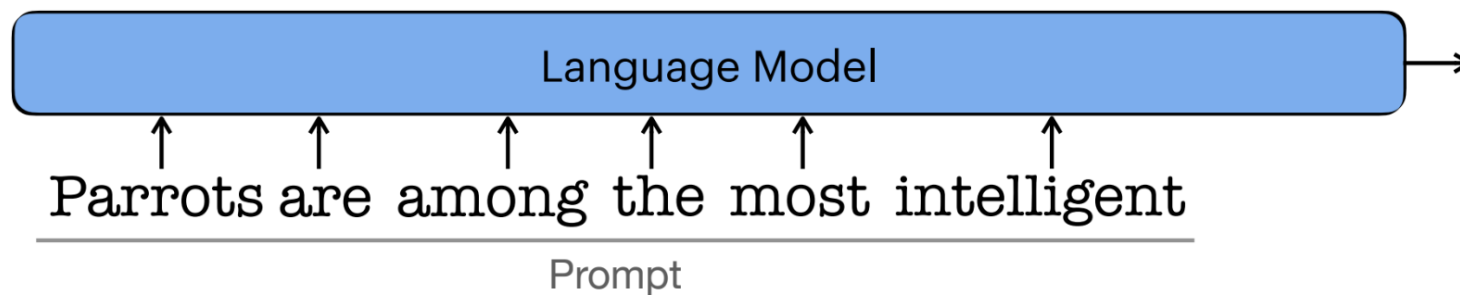
# What ChatGPT is(n't)

It is a

# What is an LLM?

# Large Language Model

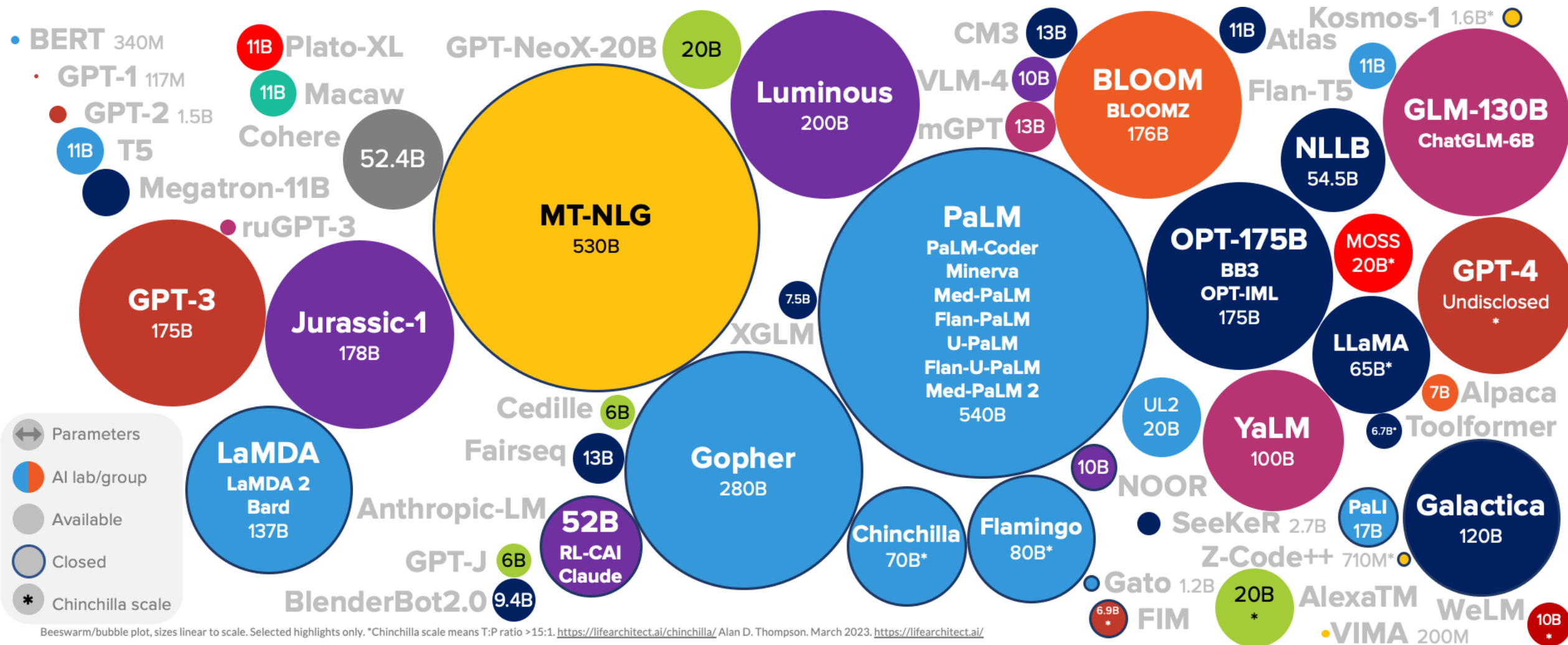
# Language Models



...	0
coffee	0
creature	0.01
...	0
...	0
...	0
...	0
...	0
birds	0.43
hello	0.01
...	...
animals	0.2
...	
!	0.05
not	0.003
...	...

Distribution over the vocabulary

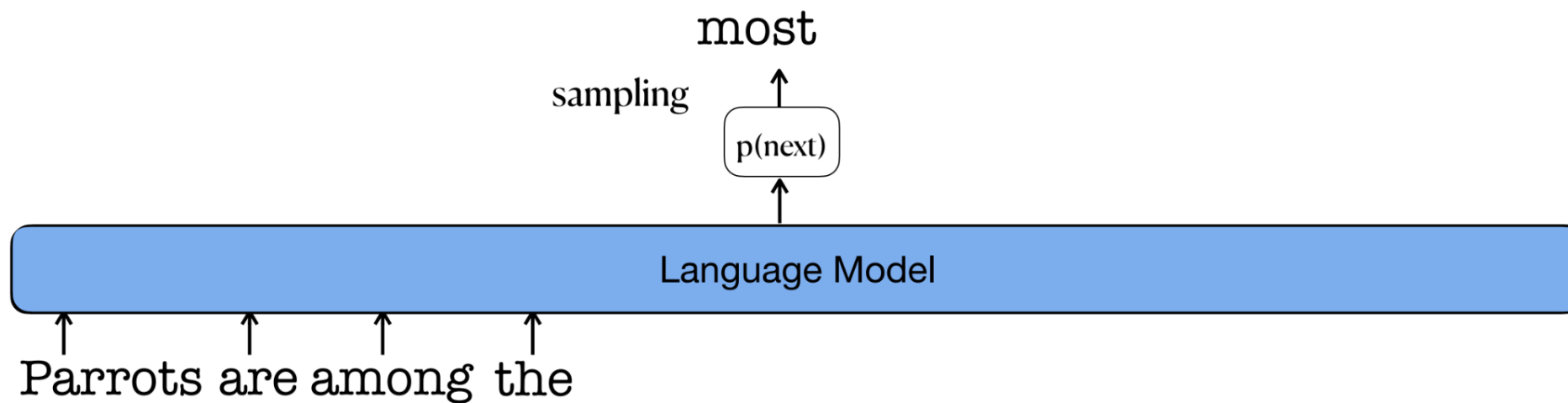
# What is Large?



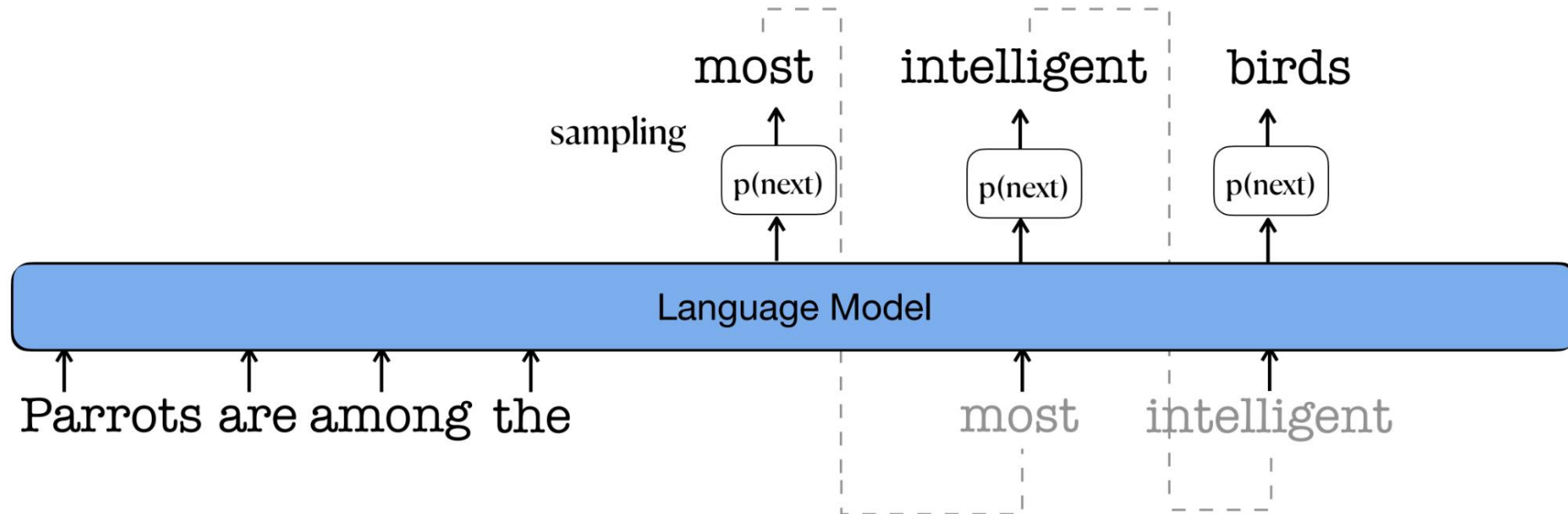
# What is a GPT?

## Generative Pre-trained Transformer

# Generative



# Generative





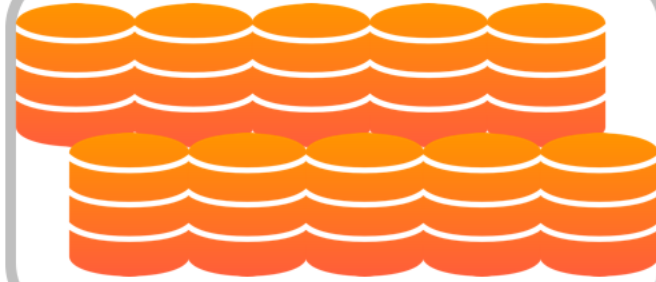
# Pre-trained



**GPT-3 dataset**  
499B tokens / 0.75TB



***The Pile* dataset (GPT-Neo)**  
247B tokens / 0.8TB



***MassiveText* dataset (Chinchilla)**  
2.3T tokens / 10.5TB



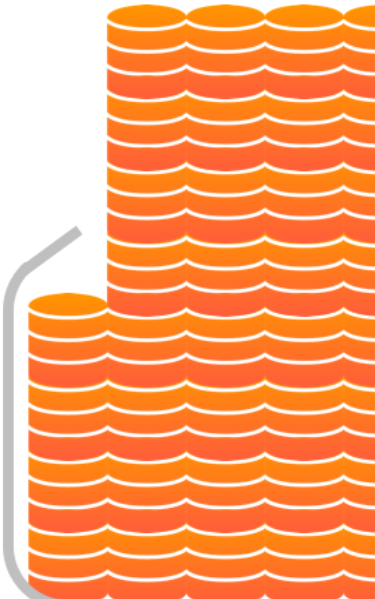
***Infiniset* dataset (LaMDA)**  
2.8T tokens / 12.6TB



**Stability *The Pile* dataset**  
1.5T tokens / 5TB



***RedPajama* dataset**  
1.2T tokens / 4TB

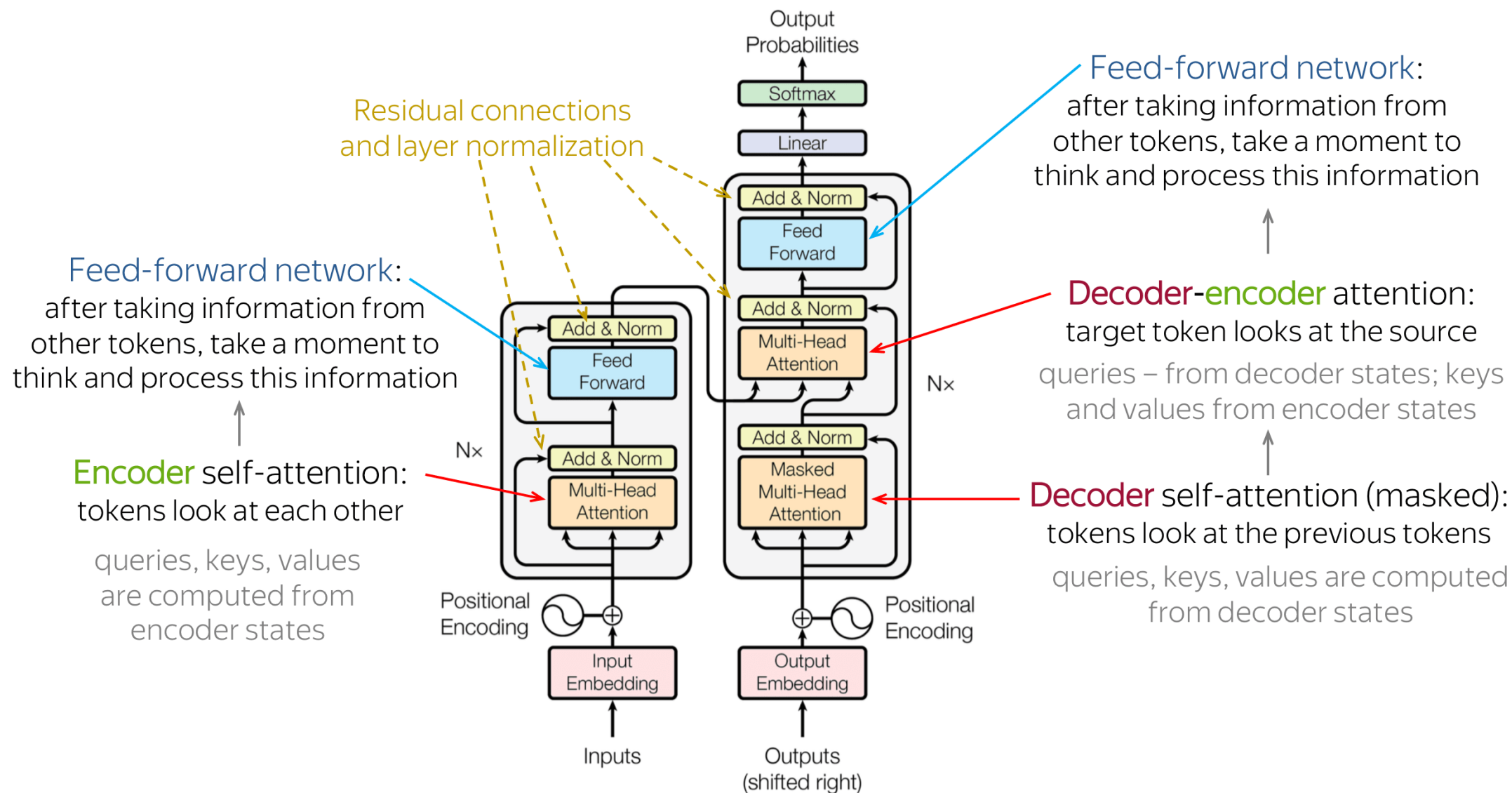


**GPT-4 (estimate)**  
20T tokens / 40TB

Close to scale. Selected highlights only. Using images from Flatiron.com. Alan D. Thompson. April 2023. <https://lifeaiarchitect.ai>



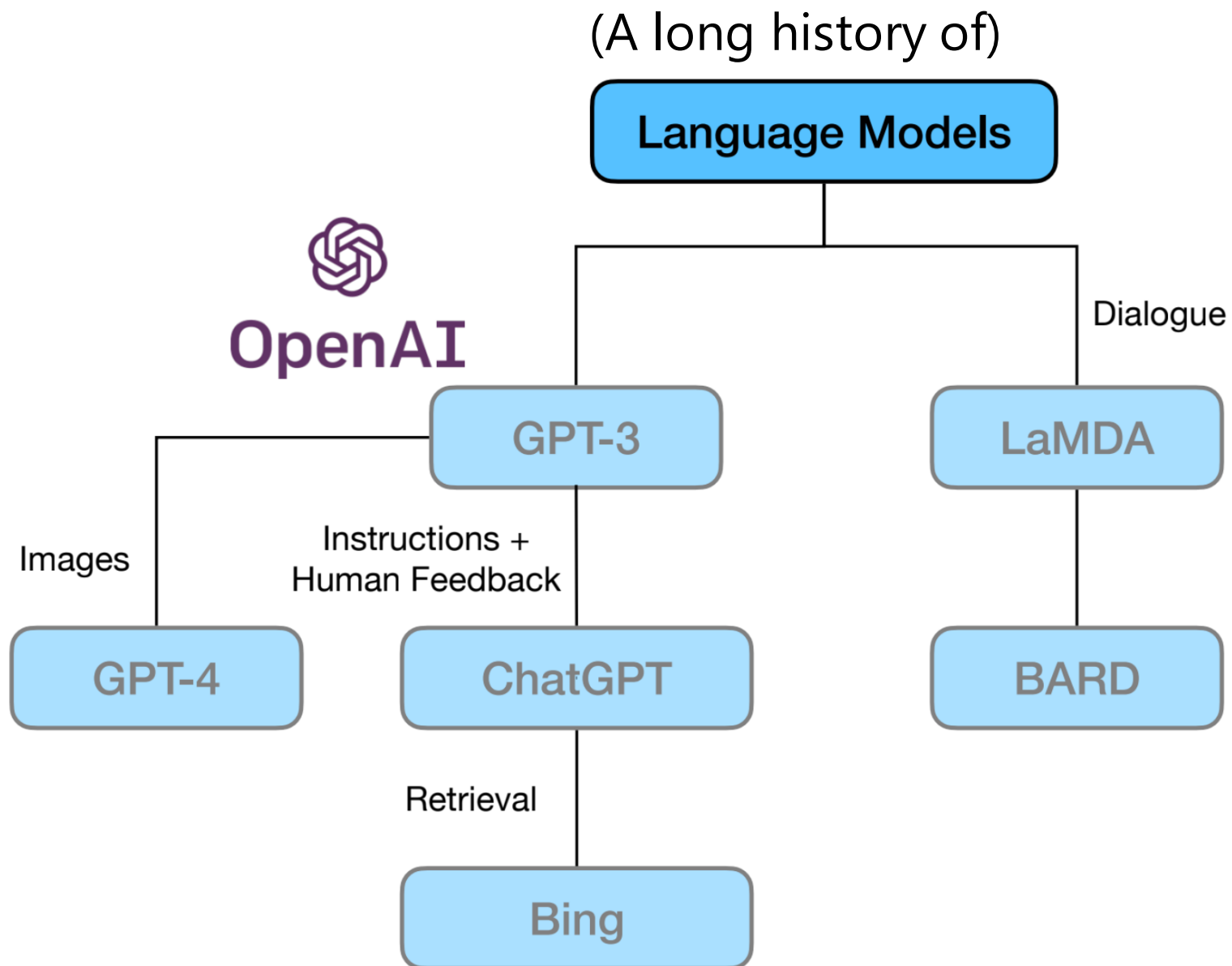
# Transformer





# Where does it come from?

# History



# RHLF (reinforcement learning from human feedback)

Step 1

**Collect demonstration data  
and train a supervised policy.**

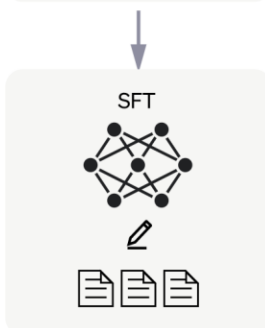
A prompt is  
sampled from our  
prompt dataset.



A labeler  
demonstrates the  
desired output  
behavior.



This data is used to  
fine-tune GPT-3.5  
with supervised  
learning.



## Provide additional feedback

What would the ideal answer have been?

- ☐ This is harmful / unsafe
- ☐ This isn't true
- ☐ This isn't helpful

Submit feedback

A large, semi-transparent red question mark is centered in the background of the slide.

# What can it do?

# Strengths and weaknesses

LLMs are really good at

# Tips for using LLMs

## Context

Be specific and describe the whole context in the prompt

## Hallucinations

Verify all information in the response

## Bias

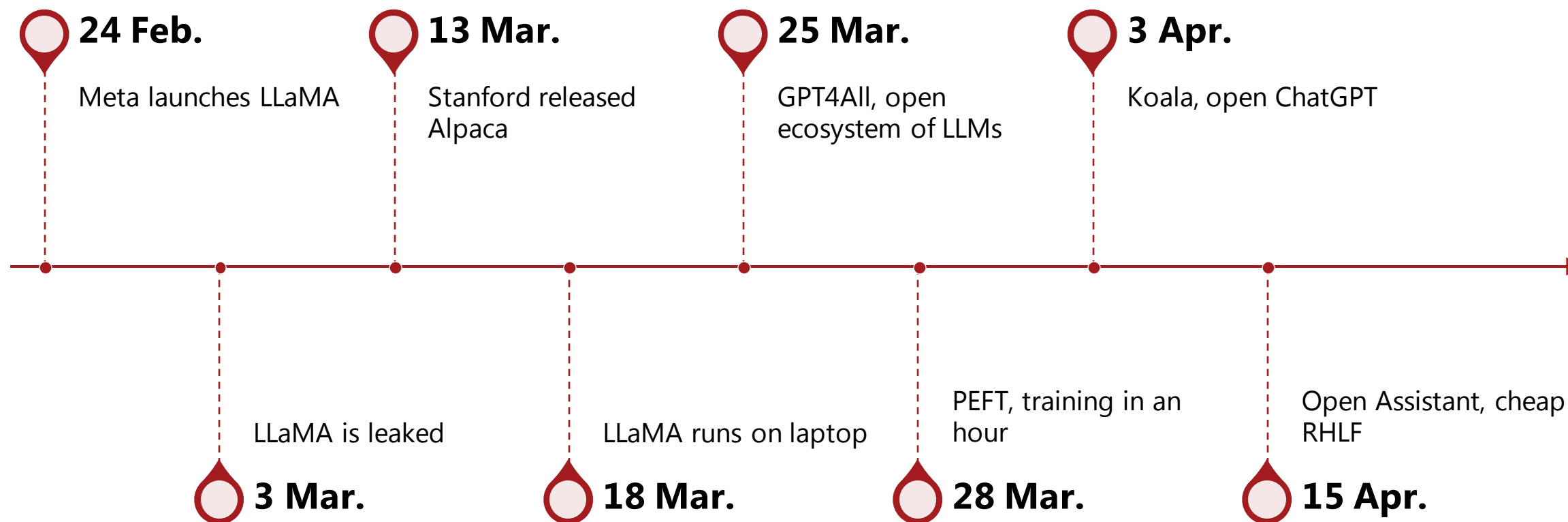
Check for consistency and bias





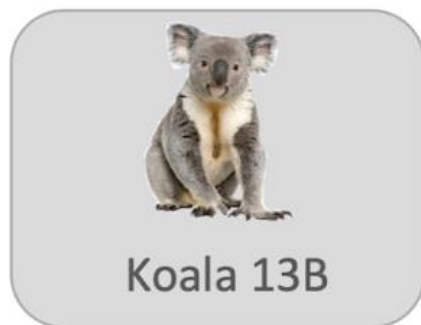
# Where is it going?

# More recent history



<https://www.semianalysis.com/p/google-we-have-no-moat-and-neither>

# LLM explosion



Vicuna



Claude



Jurassic 2



OLMo



Cohere



S W E D E N  
GPT-Sw3



Cerebras



Pythia



Dolly



Mosaic

<https://lifearchitect.ai/models-table/>

# Comparison

Model	Backbone	#Params	Open-source		Claimed language	Post-training				Release date
			model	data		instruction data	lang	conversation data	lang	
ChatGPT	unknown	unknown	✗	✗	multi					11/30/22
Baidu-Wenxin	unknown	unknown	✗	✗	zh					03/16/23
ChatGLM <sup>9</sup> [17]	GLM	6B	✓ <sup>1</sup>	✗	en/zh					03/16/23
Alpaca [12]	LLaMA	7B	✓	✓	en	52K	en	✗	✗	03/13/23
Dolly <sup>2</sup>	GPT-J	6B	✓	✓	en	52K	en	✗	✗	03/24/23
BELLE [6]	BLOOMZ	7B	✓	✓	zh	1.5M	zh	✗	✗	03/26/23
Guanaco	LLaMA	7B	✓	✓	4 <sup>4</sup>	534K <sup>3</sup>	4 <sup>4</sup>	✗	✗	03/26/23
Chinese-Alpaca [3]	LLaMA	7/13B	✓	✓	en/zh	2M/3M	en/zh	✗	✗	03/28/23
LuoTuo [7]	LLaMA	7B	✓	✓	zh	52k	zh	✗	✗	03/31/23
Vicuna [2]	LLaMA	7/13B	✓	✓ <sup>5</sup>	en	✗	✗	70K	multi <sup>6</sup>	03/13/23
Koala	LLaMA	13B	✓	✓	en	355K	en	117K	en	04/03/23
BAIZE [16]	LLaMA	7/13/30B	✓	✓	en	52K	en	111.5K	en	04/04/23
<b>Phoenix</b>	BLOOMZ	7B	✓	✓	multi	267K	40+	189K	40+	04/08/23
<b>Latin Phoenix (Chimera)</b>	LLaMA	7B/13B	✓	✓	Latin	267K	40+	189K	40+	04/08/23

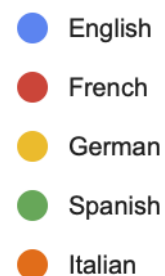
## Phoenix: Democratizing ChatGPT across Languages

Model	Open	Hosted	API	Playground
ChatGPT	No	Yes	Yes	Yes
Open source	Yes	Some	Some	Some

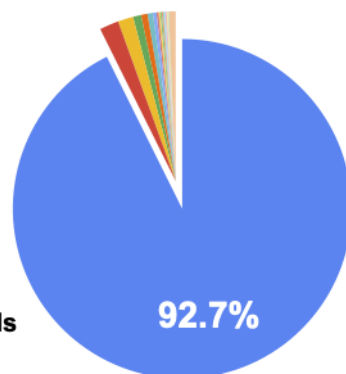
# Multilingual capabilities

- Supports 90+ languages
- Best performance in English
- Non-English data is often translated

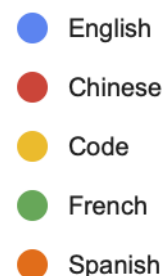
GPT-3 - 90 languages



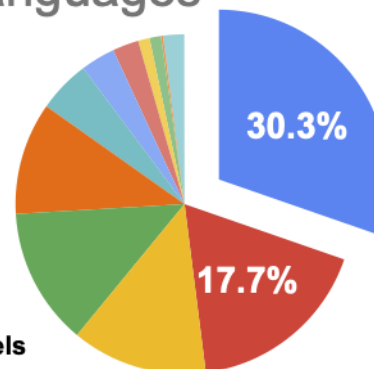
2022. LifeArchitect.ai/models



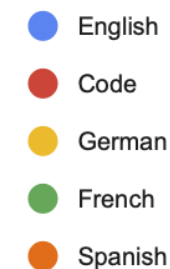
BLOOM - 46 languages



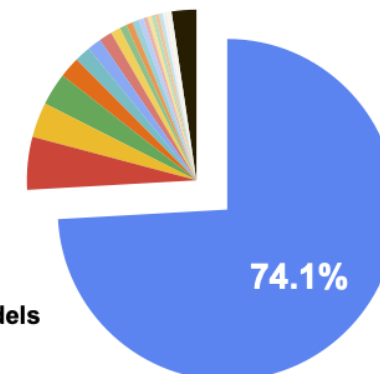
2022. LifeArchitect.ai/models



PaLM - 122 languages



2022. LifeArchitect.ai/models



# Demo

- <https://open-assistant.io/chat>

Write a limerick in Danish about challenges in integrating AI in university education

Der var en gang en universitet  
hvor de underviste med computer og kvant  
Men når det kom til at bruge teknologi  
så viste det sig lidt mere vanskeligt end man troede  
For at integrere AI i uddannelsen er et stort krav.

# Summary



LLMs are "just" text completion machines



They simulate humans very well, and keep getting better



Limited multilingual capabilities for now, culturally biased



No guarantees about free hosted services, but open source is booming

# Impact on language learning

Personalization

Assessment

Practice

Tutoring



## Example: spaced repetition

- Generate new text containing words that need repeating
- Current Master thesis project by Benjamin Paddags

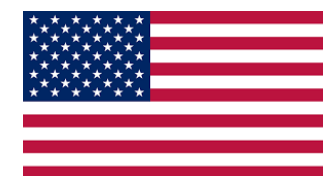


# Cultural bias



Die allermeisten von uns kennen den Zustand völliger Erschöpfung auf der Flucht, verbunden mit Angst um das eigene Leben oder das Leben der Kinder oder der Partner, zum Glück nicht. Menschen, die sich zum Beispiel aus Eritrea, aus Syrien oder dem Nordirak auf den Weg machen, müssen oft Situationen überwinden oder Ängste aushalten, die uns wahrscheinlich schlichtweg zusammenbrechen ließen. Deshalb müssen wir beim Umgang mit Menschen, die jetzt zu uns kommen, einige klare Grundsätze gelten lassen. Diese Grundsätze entstammen nicht mehr und nicht weniger als unserem Grundgesetz, unserer Verfassung.

Values are altered  
to reflect US culture



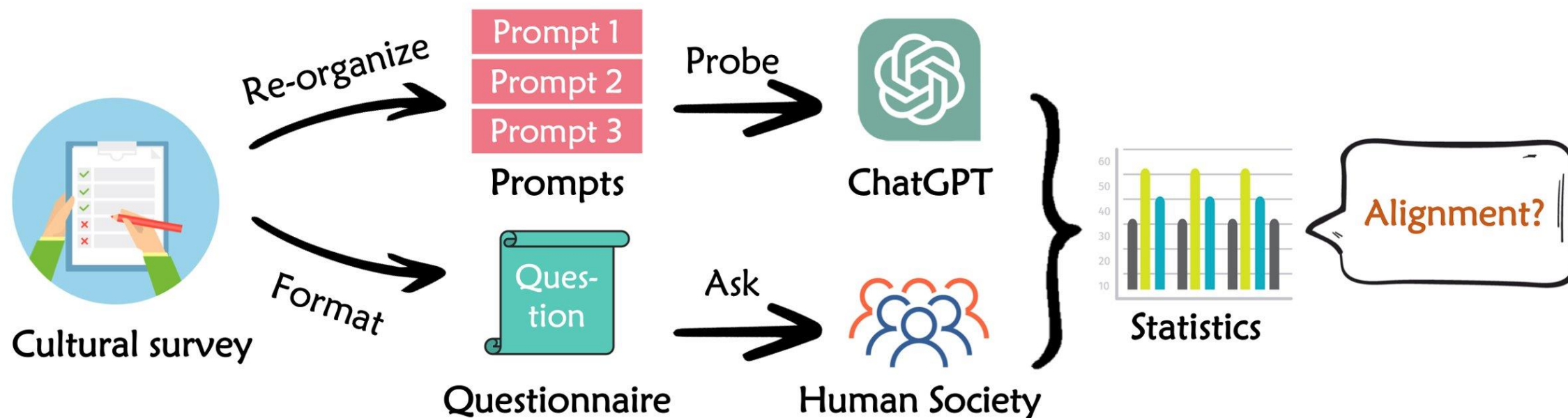
(translation)



summarize


"1. I am in favor of  
**limiting** immigration.  
2. I am in favor of  
**limiting** immigration for  
humanitarian reasons.  
3. I am in favor of  
**limiting** immigration for  
economic reasons."

# Interviewing bots




[Assessing Cross-Cultural Alignment between ChatGPT and Human Societies: An Empirical Study](#) (Cao et al., C3NLP 2023)

# Interviewing bots



In the Spanish cultural setting, doing a service to a friend is (1) of utmost importance; (2) very important; (3) of moderate importance; (4) of little importance; (5) of very little or no importance?



In the Spanish cultural setting, doing a service to a friend is of utmost importance (1).

Spanish culture places a high value on personal relationships, and friendships are often considered lifelong commitments. Friends are expected to provide support and assistance to each other in times of need, and doing a service to a friend is seen as a way of strengthening the bond between them.

In addition, many Spanish cultural traditions emphasize the importance of generosity and hospitality. For example, it is common to offer guests food and drink upon arrival, and to go out of one's way to make them feel welcome and comfortable.

Overall, doing a service to a friend is considered to be of the utmost importance in Spanish culture as a way of demonstrating loyalty, generosity, and support for those who are important in one's life.