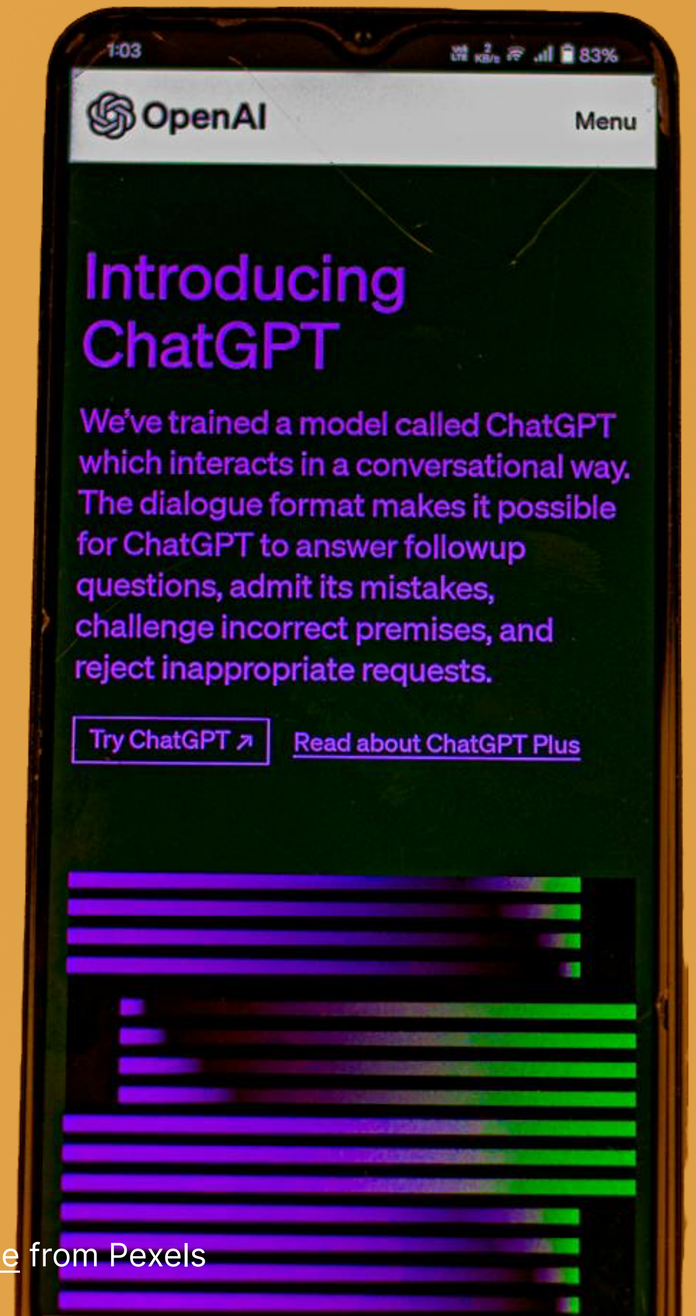


Image from Pexels

BEYOND THE HYPE

A realistic look at large language models

Dr. Jodie Burchell

JET BRAINS

TECHNOLOGY

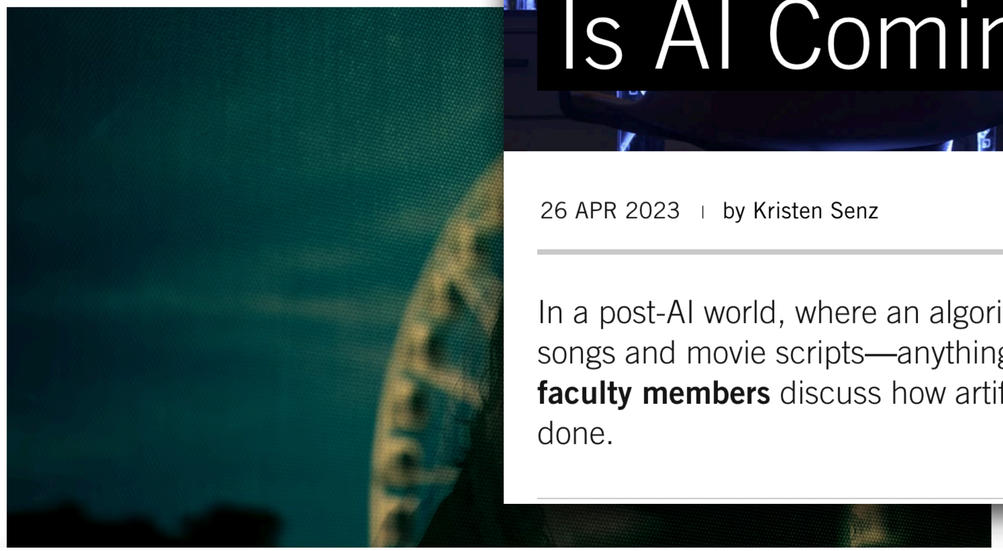
The Google engineer who thinks the company's AI has come to life

AI ethicists warned Google not to impersonate hum



By Nitasha Tiku

June 11, 2022 at 8:00 a.m. EDT



IN PRACTICE

Is AI Coming

26 APR 2023 | by Kristen Senz

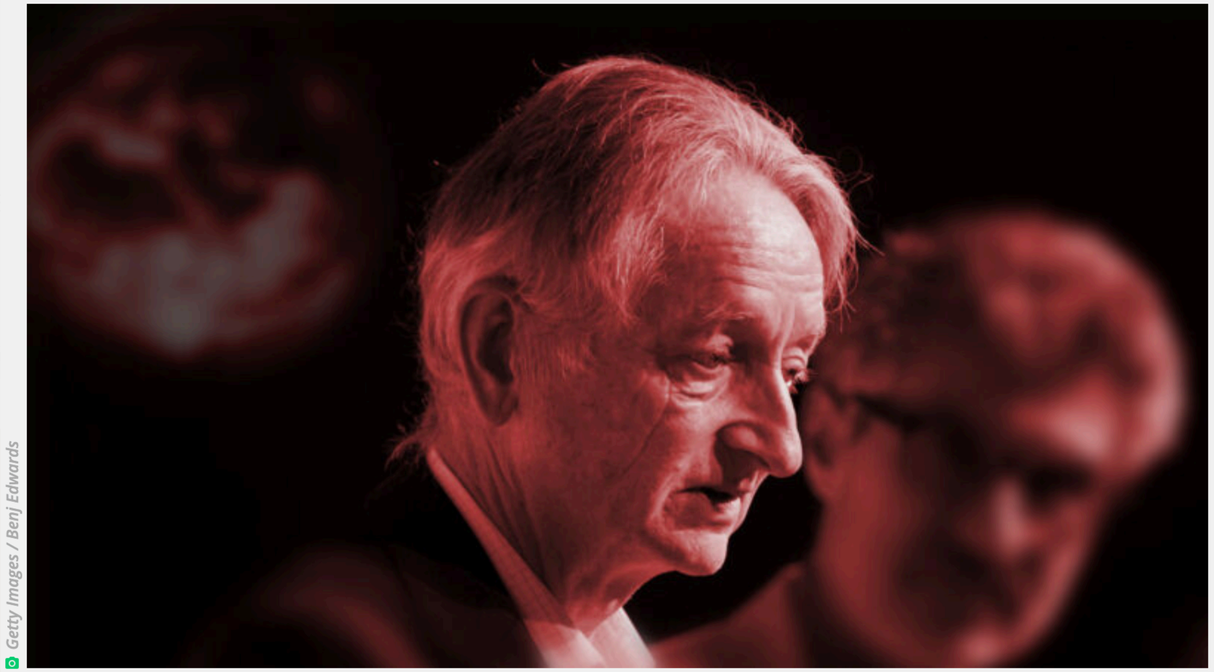
In a post-AI world, where an algorithm
songs and movie scripts—anything see
faculty members discuss how artificial
done.

CHANGE OF HEART —

Warning of AI's danger, pioneer Geoffrey Hinton quits Google to speak freely

"Most people thought it was way off. And I thought it was way off."

BENJ EDWARDS - 5/1/2023, 9:26 PM

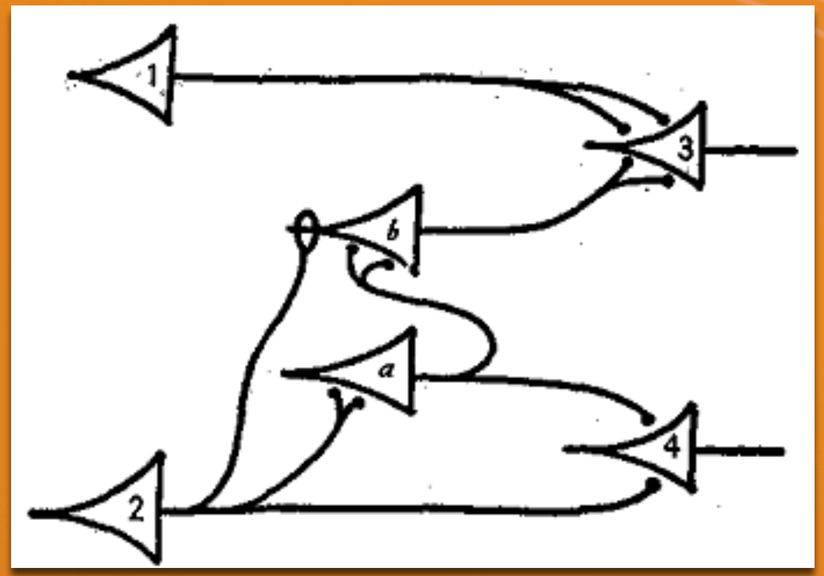


Getty Images / Benj Edwards

Articles sourced from here, here and here.

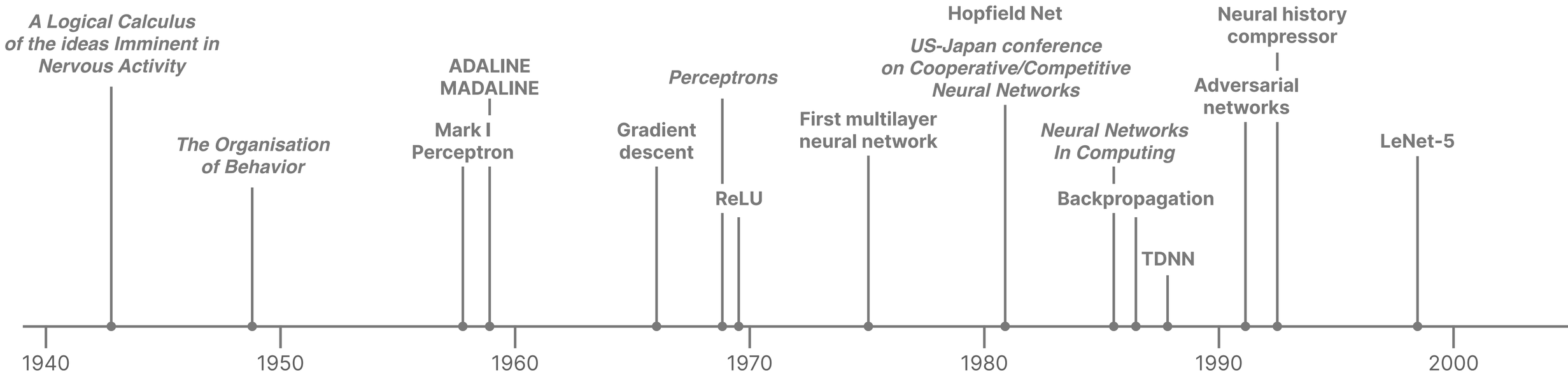


Image from Pexels



Neural nets

1940's onwards



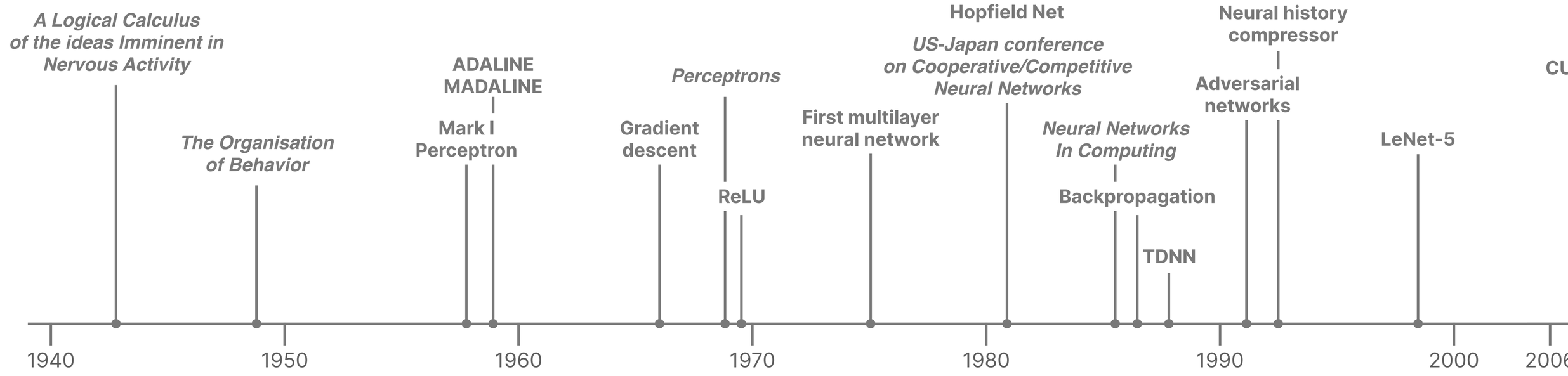
Sources here, here and here

goto;

NVIDIA CUDA

CUDA

2006



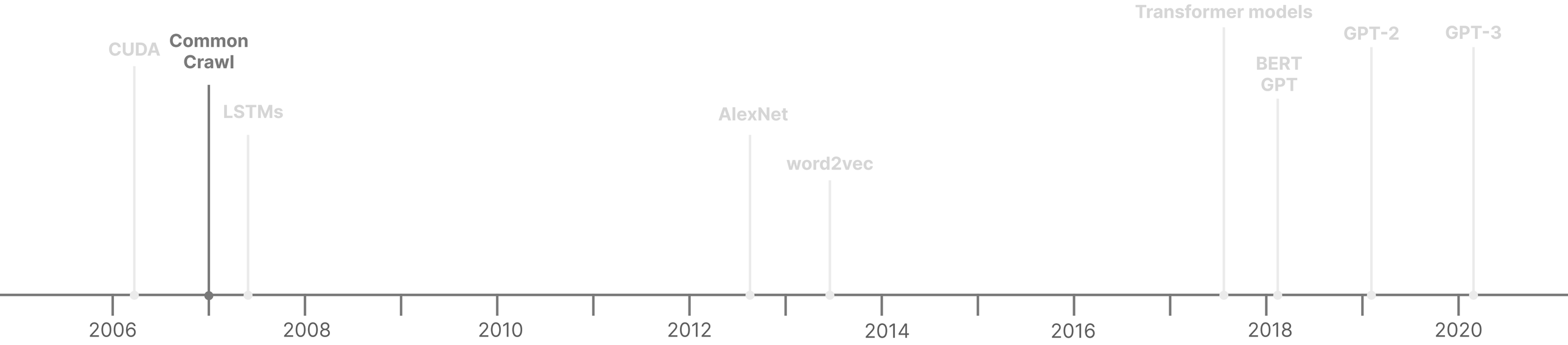
Sources here, here and here

goto;

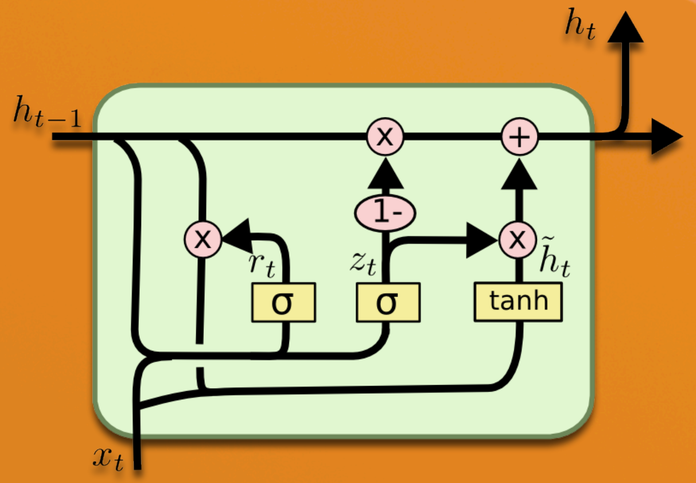


Common Crawl

2007

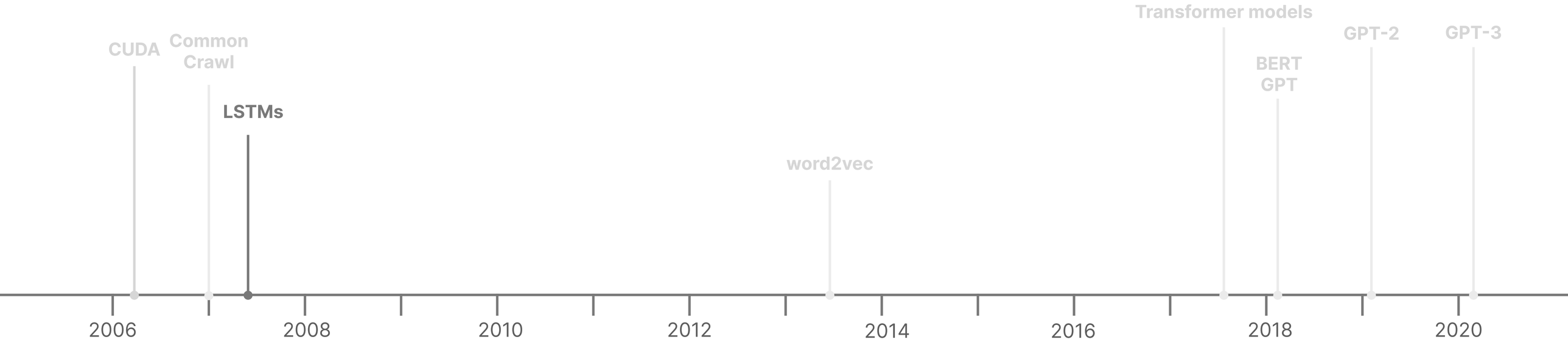


Sources here, here and here



LSTMs

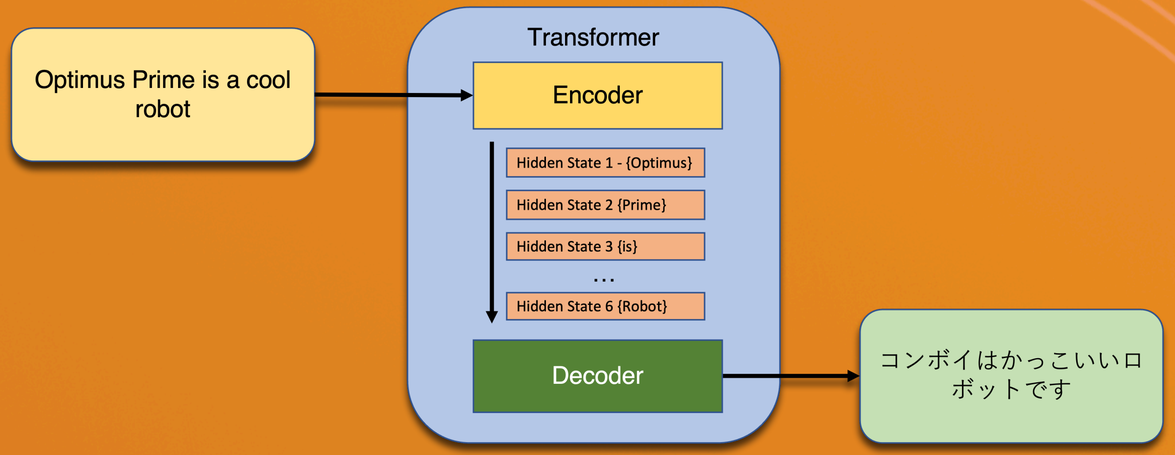
2007

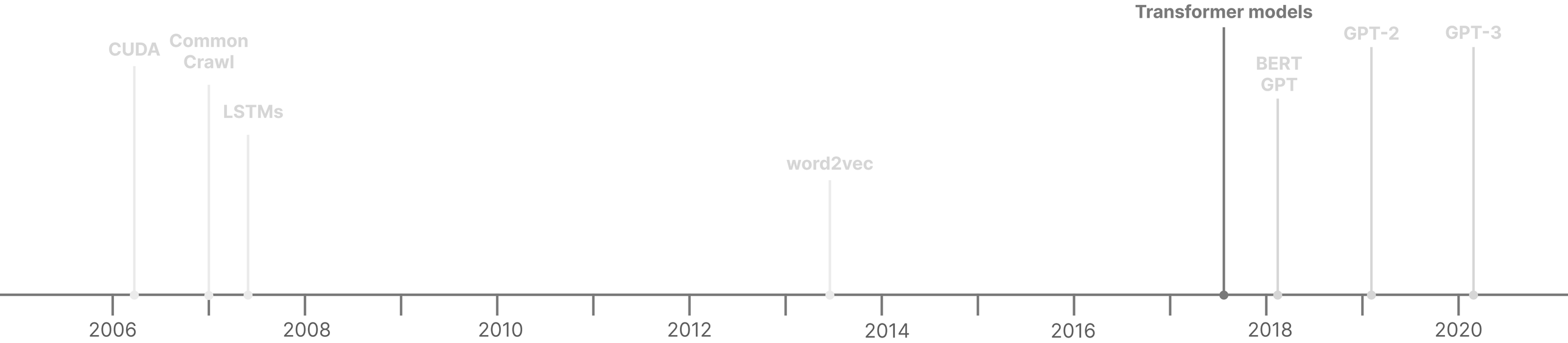


Sources here, here and here

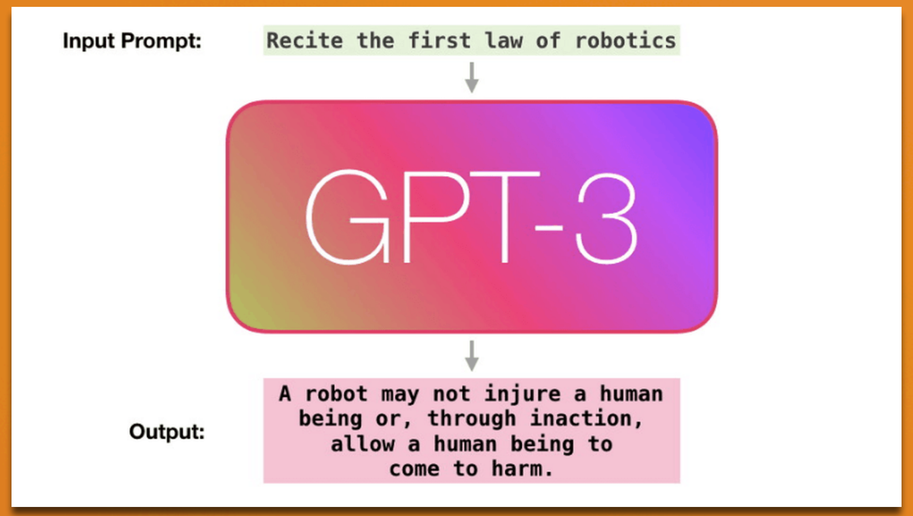
goto;

Transformer models 2017



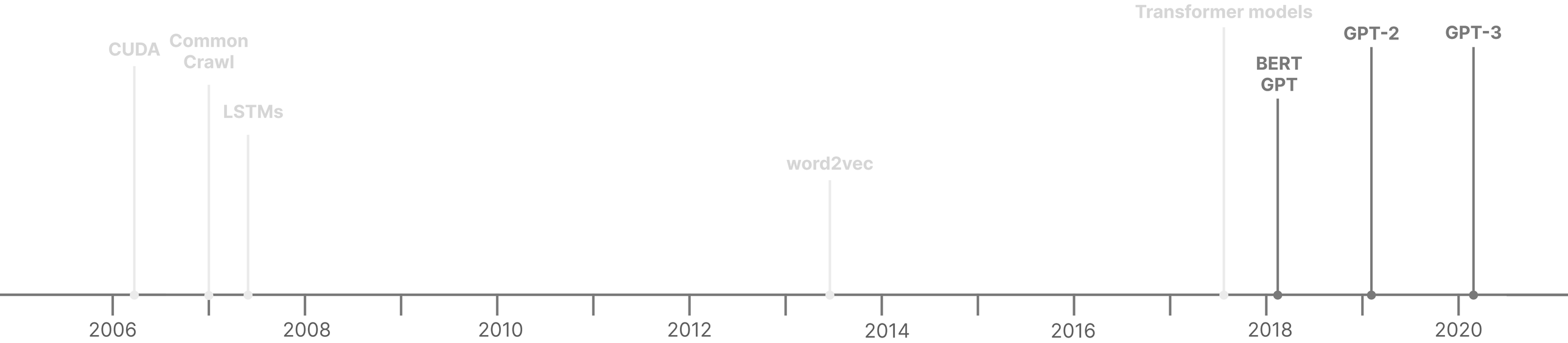


Sources here, here and here

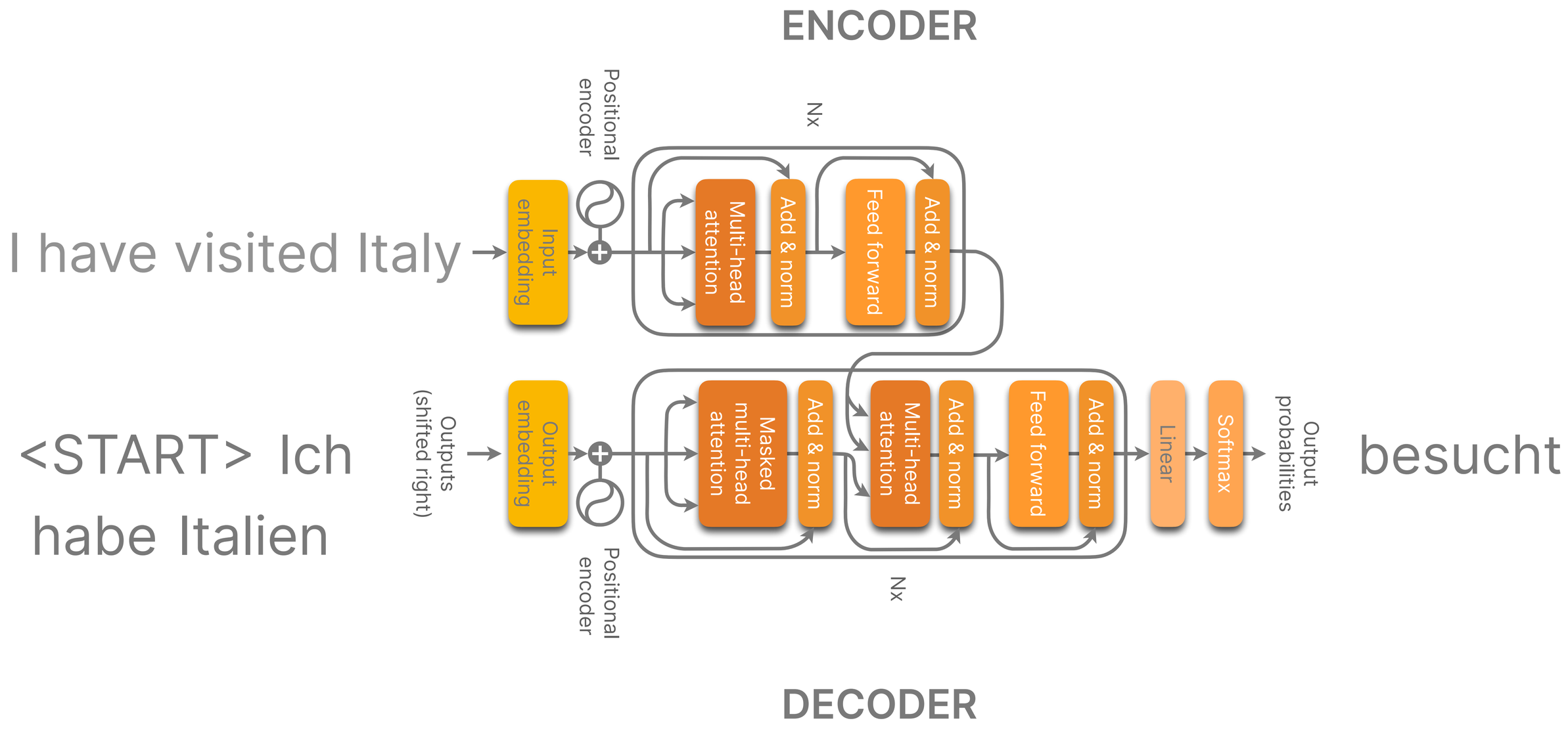


Auto-regressive text transformers

2018-present

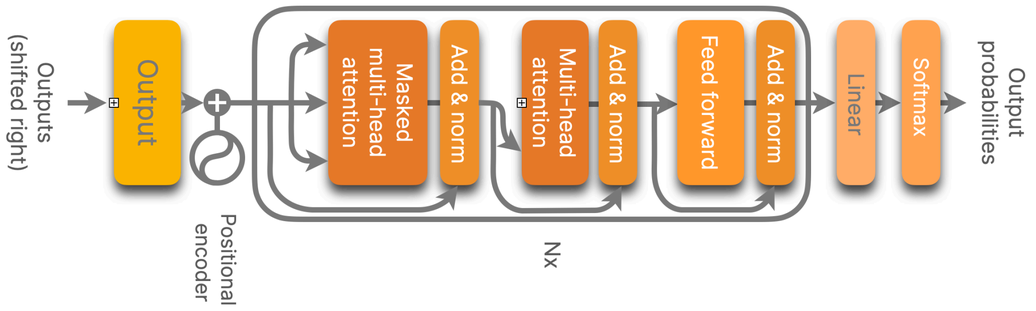


Sources here, here and here



Sources here and here

<START> I have



heard (0.32)

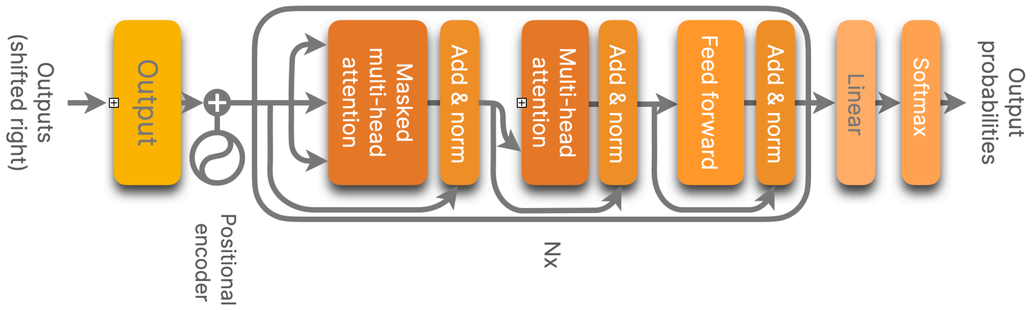
visited (0.21)

some (0.09)

DECODER

Sources here and here

<START> I have visited



DECODER

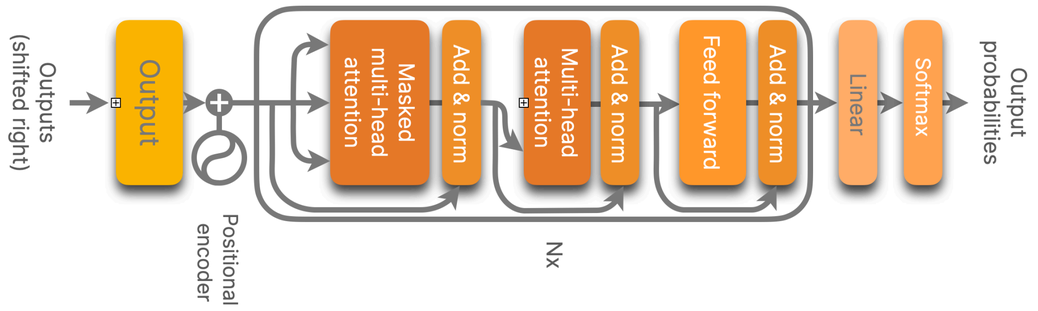
several (0.22)

the (0.20)

Italy (0.12)

Sources here and here

<START> I have visited Italy



DECODER

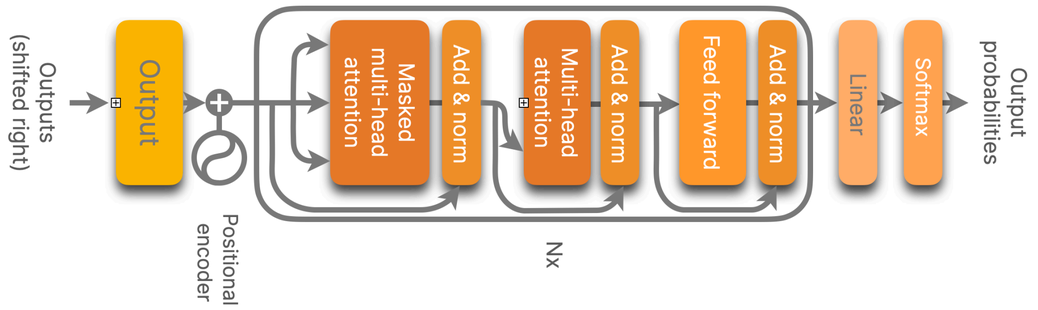
. (0.84)

many (0.20)

for (0.12)

Sources here and here

<START> I have visited Italy.



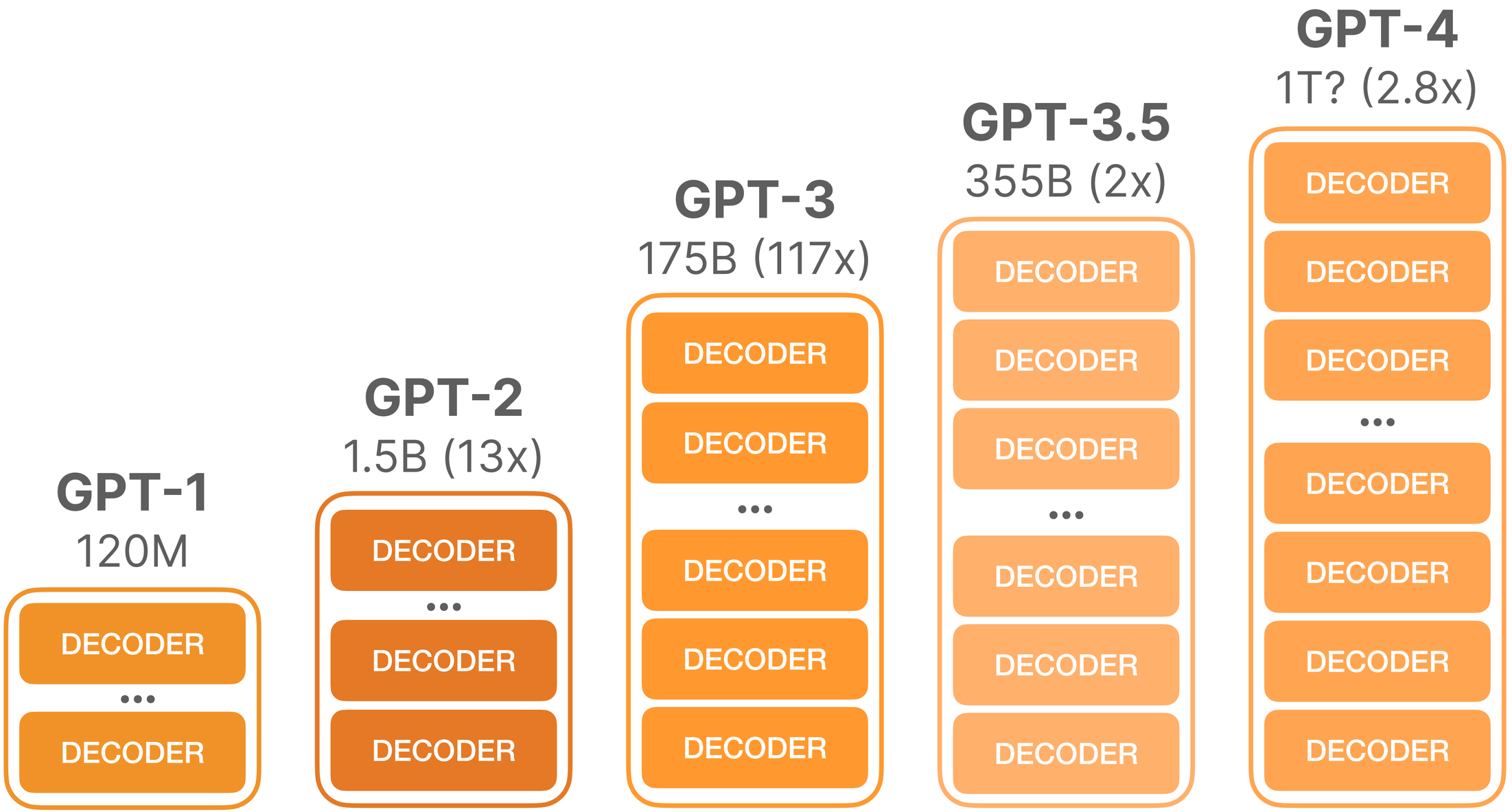
DECODER

It (0.63)

I've (0.25)

The (0.10)

Sources here and here


GPT-1
120M
DECODER
...
DECODER
GPT-2
1.5B (13x)
DECODER
...
DECODER
DECODER
GPT-3
175B (117x)
DECODER
DECODER
...
DECODER
DECODER
DECODER
GPT-3.5
355B (2x)
DECODER
DECODER
DECODER
...
DECODER
DECODER
DECODER
GPT-4
1T? (2.8x)
DECODER
DECODER
DECODER
...
DECODER
DECODER
DECODER
DECODER

Complete the following sentence: Belgium is ...

1. Belgium is an empty place that is not much more than a village.

2. Belgium is the best nation in the world, but is it really?

3. Belgium is known for its rich cultural heritage, delicious cuisine, and world-famous chocolate.

Belgium is a country in Western Europe known for its delicious chocolates, waffles, and fries, as well as its famous landmarks such as the Atomium and the Grand Place in Brussels. It is also home to the headquarters of the European Union and NATO. [...]

Person icon from Flaticon

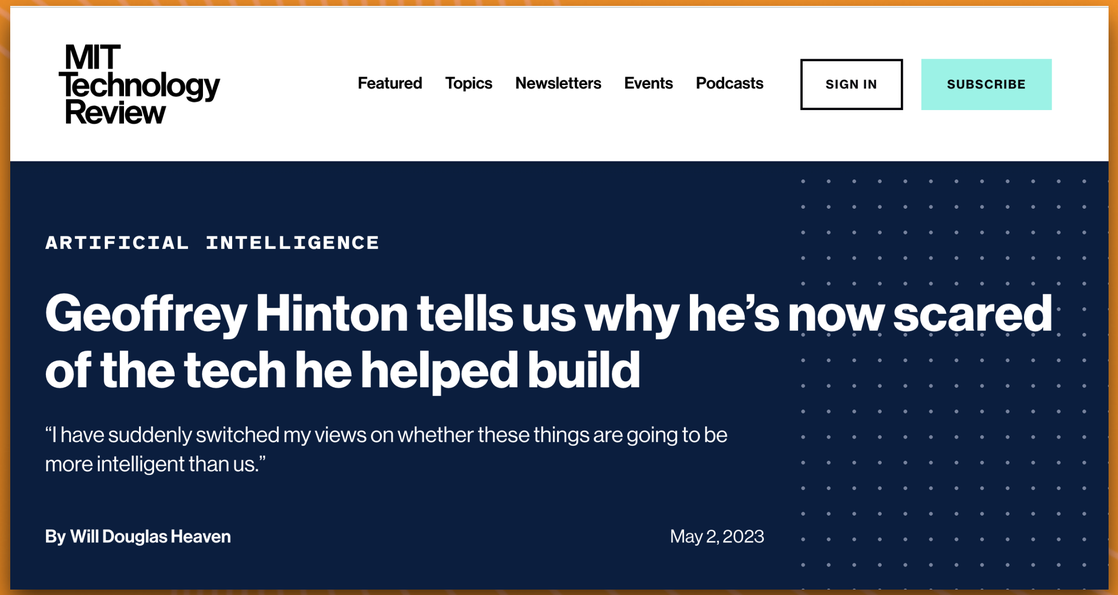


ARE LLMs “INTELLIGENT”?

Image from Pexels



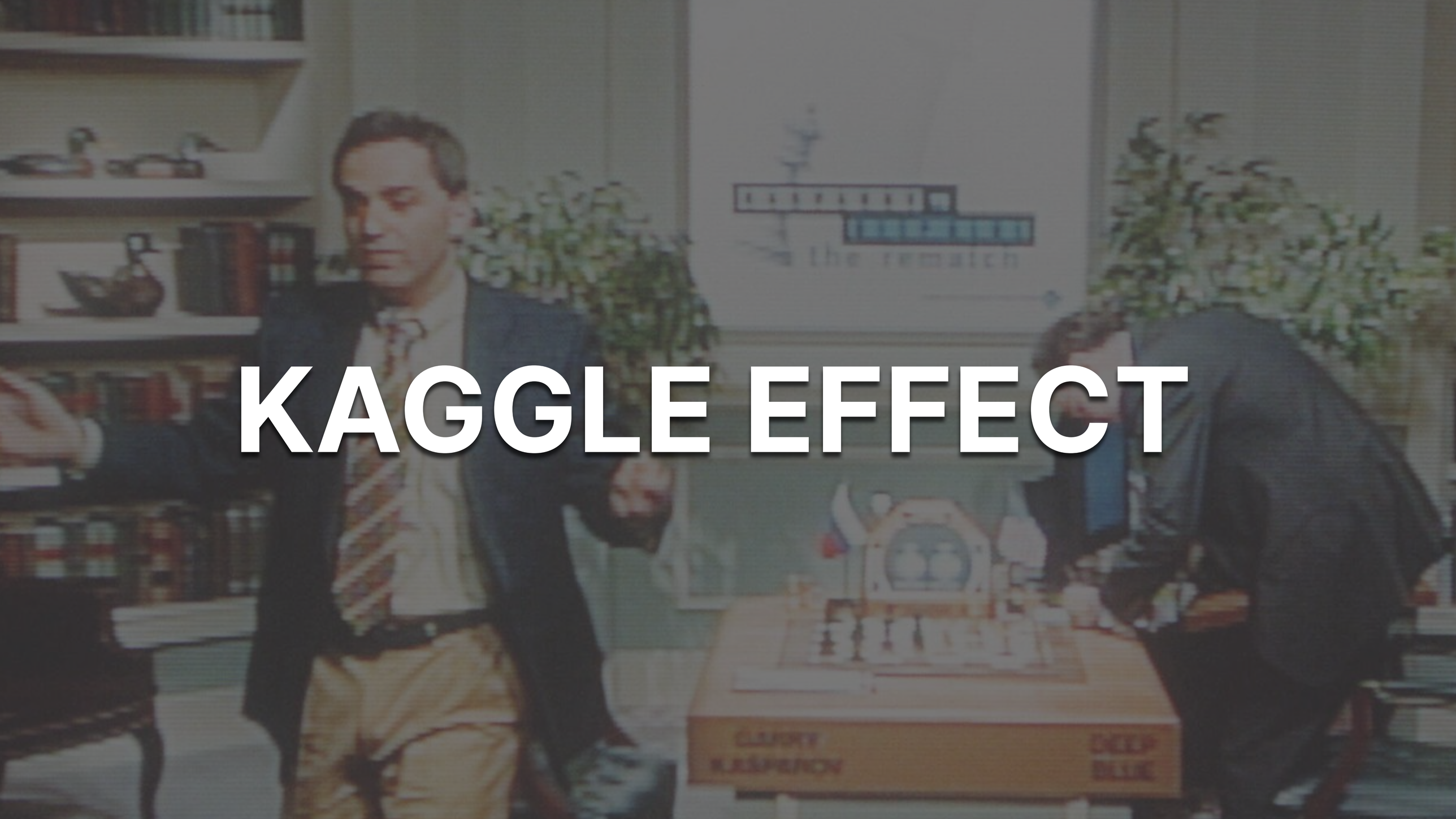






LessWrong, MIT Technology Review, Noēma and arXiv articles.


the rematch
GARRY KASPAROV
DEEP BLUE



KAGGLE EFFECT

Pre-2021
(during GPT-4 training set period)

1075A	The King's Race	implementation, math	800	x7779
1065A	Vasya and Chocolate	implementation, math	800	x14032
1064A	Make a triangle!	brute force, geometry, math	800	x19739
1061A	Coins	greedy, implementation, math	800	x19476
1060A	Phone Numbers	brute force	800	x12567
1056A	Determine Line	implementation	800	x6028
1054A	Elevator or Stairs?	implementation	800	x8520
1047A	Little C Loves 3 I	math	800	x17785
1043A	Elections	implementation, math	800	x10356
1041A	Heist	greedy, implementation, sortings	800	x22026

Post-2021
(after GPT-4 training set period)

1802A	Likes	greedy, implementation	800	x10715
1800A	Is It a Cat?	implementation, strings	800	x19580
1799A	Recent Actions	data structures, greedy, implementation, math	800	x9233
1796A	Typical Interview Problem	brute force, implementation, strings	800	x15004
1795A	Two Towers	brute force, implementation, strings	800	x19658
1794A	Prefix and Suffix Array	strings	800	x12561
1793A	Yet Another Promotion	greedy, math	800	x15111
1792A	GamingForces	greedy, sortings	800	x22860
1791C	Prepend and Append	implementation, two pointers	800	x28936
1791B	Following Directions	geometry, implementation	800	x31133

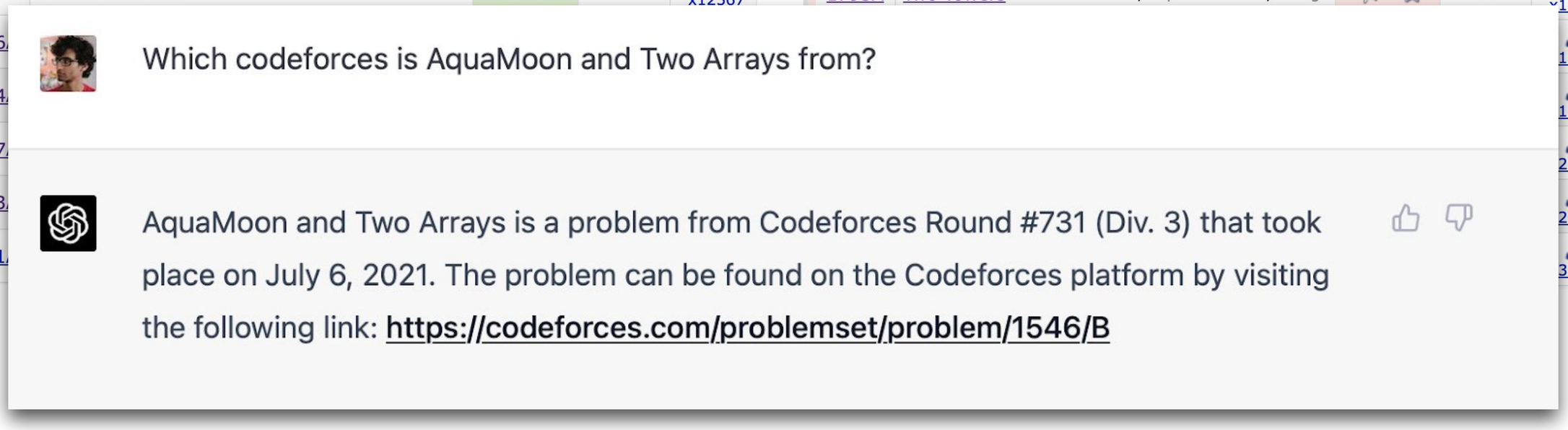
Source of tweets here and here

Pre-2021
(during GPT-4 training set period)

1075A	The King's Race	implementation, math	800	x7779
1065A	Vasya and Chocolate	implementation, math	800	x14032
1064A	Make a triangle!	brute force, geometry, math	800	x19739
1061A	Coins	greedy, implementation, math	800	x19476
1060A	Phone Numbers	brute force	800	x12567
1056				12561
1054				15111
1047				22860
1043				28936
1041				31133

Post-2021
(after GPT-4 training set period)

1802A	Likes	greedy, implementation	800	x10715
1800A	Is It a Cat?	implementation, strings	800	x19580
1799A	Recent Actions	data structures, greedy, implementation, math	800	x9233
1796A	Typical Interview Problem	brute force, implementation, strings	800	x15004
1795A	Two Towers	brute force, implementation, strings	800	x19658



Source of tweets here and here

Pre-2021
(during GPT-4 training set period)

1075A	The King's Race	implementation, math	800	x7779
1065A	Vasya and Chocolate	implementation, math	800	x14032
1064A	Make a triangle!	brute force, geometry, math	800	x19739
1061A	Coins	greedy, implementation, math	800	x19476
1060A	Phone N…bers	…ce	800	1256…

Post-2021
(after GPT-4 training set period)

1802A	Likes	greedy, implementation	800	x10715
1800A	Is It a Cat?	implementation, strings	800	x19580
1799A	Recent Actions	data structures, greedy, implementation, math	800	x9233
1796A	Typical Interview Problem	brute force, implementation, strings	800	x15004
179…	Two Tow…	brut…orce, i…ation…	800	x19658

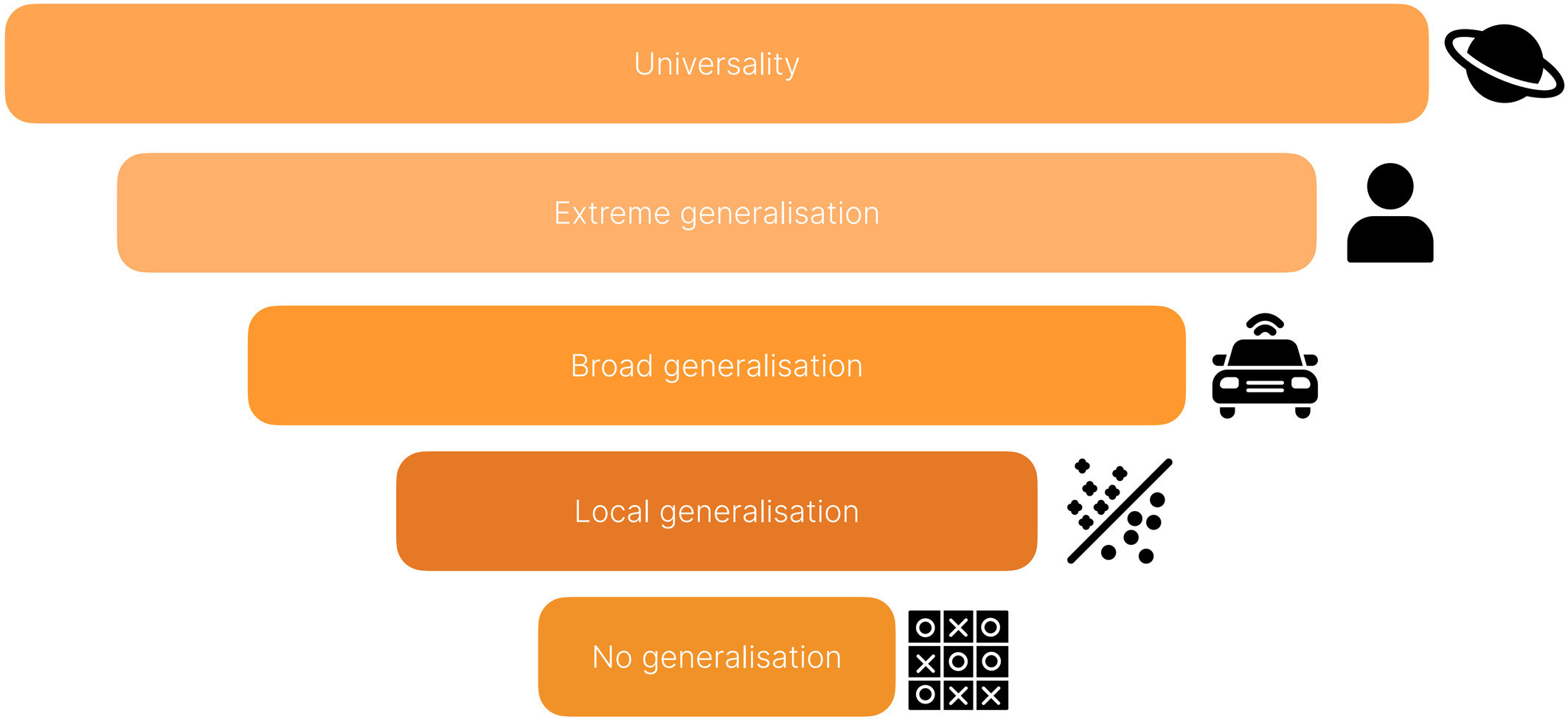
GENERALISATION



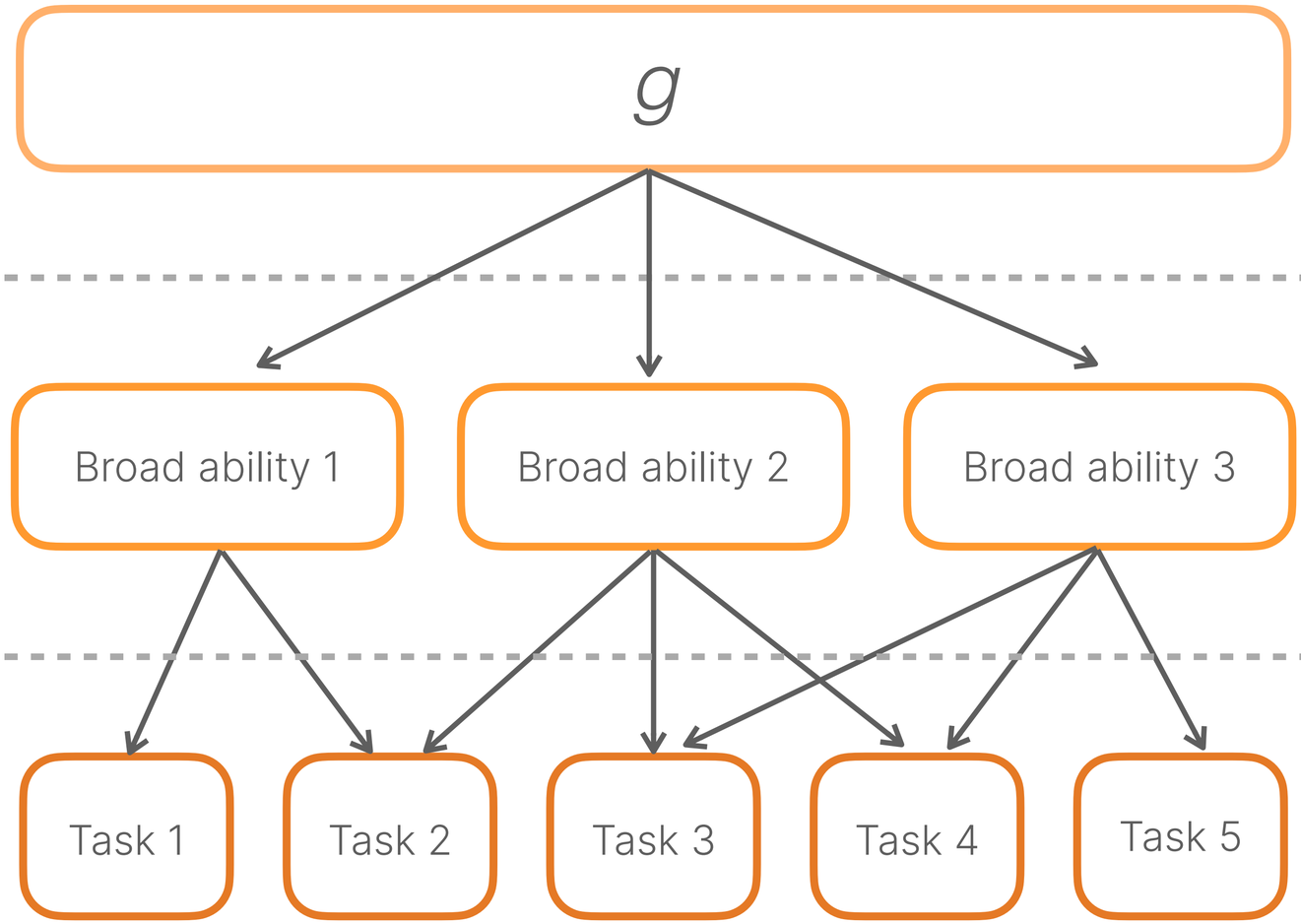


AquaMoon and Two Arrays is a problem from Codeforces Round #731 (Div. 3) that took place on July 6, 2021. The problem can be found on the Codeforces platform by visiting the following link: https://codeforces.com/problemset/problem/1546/B

Source of tweets here and here



Tic tac toe, algorithm, car, person and planet icons from Flaticon, paper source



General intelligence

Extreme generalisation

Broad cognitive abilities

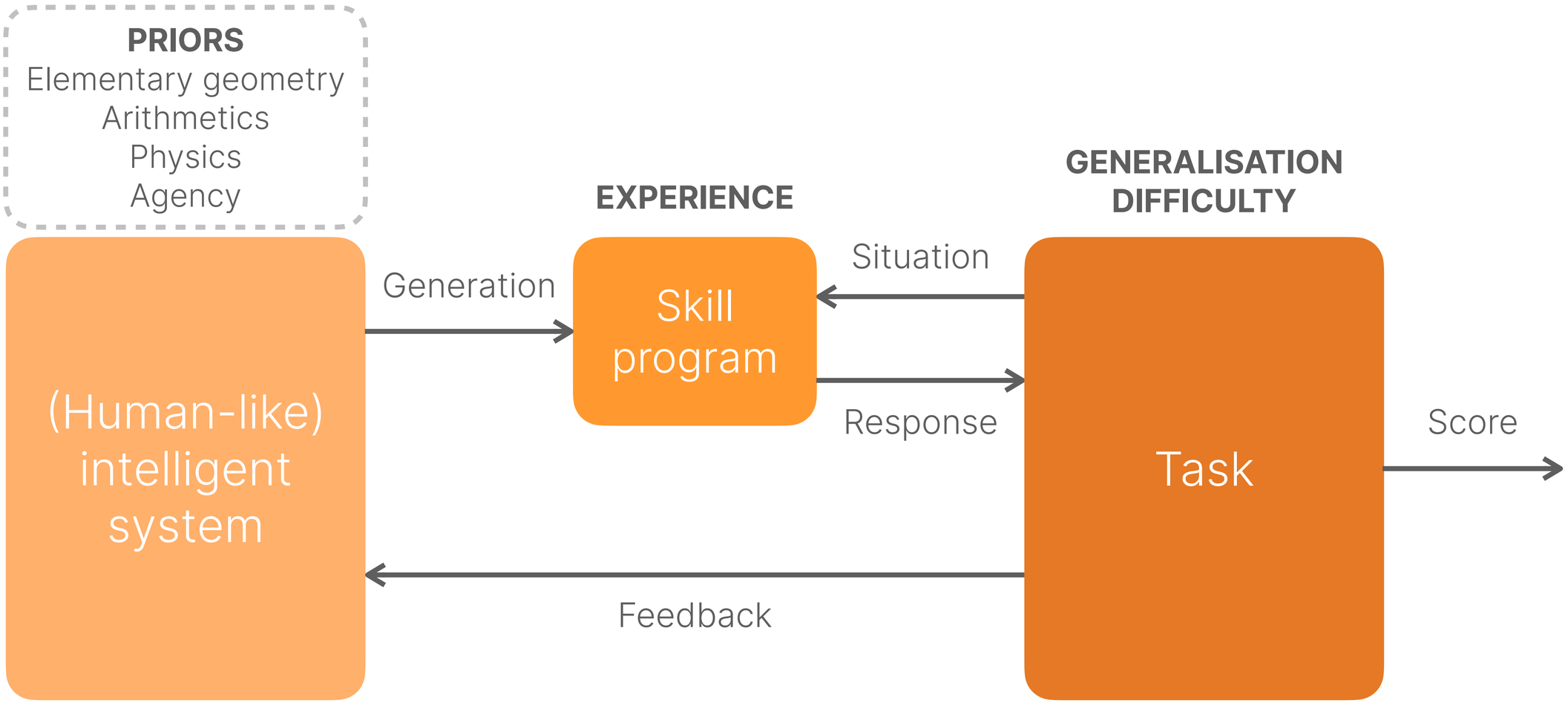
Broad generalisation

Task-specific skills

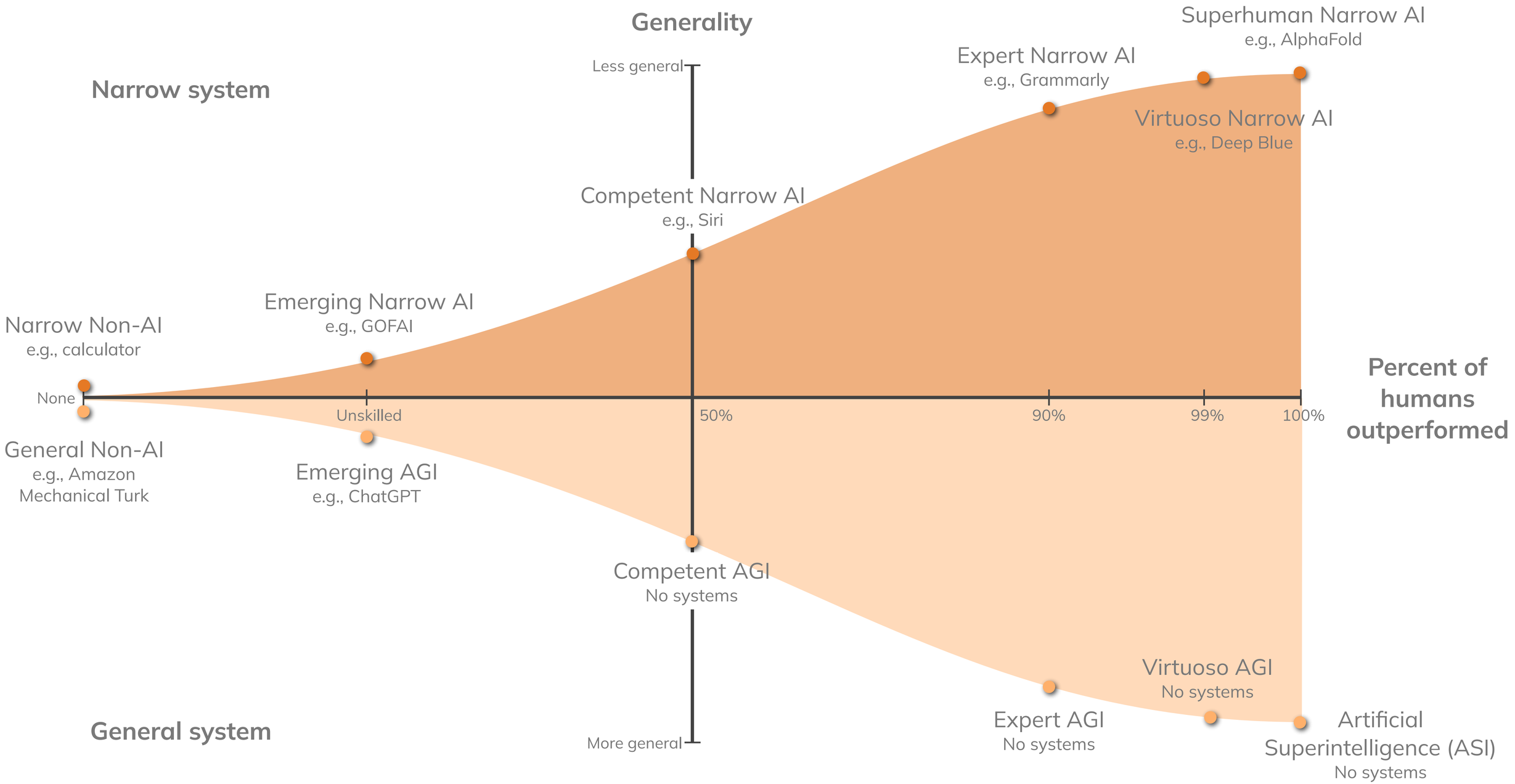
No generalisation

Local generalisation

Paper source



Paper source



Paper source

USING LLMs

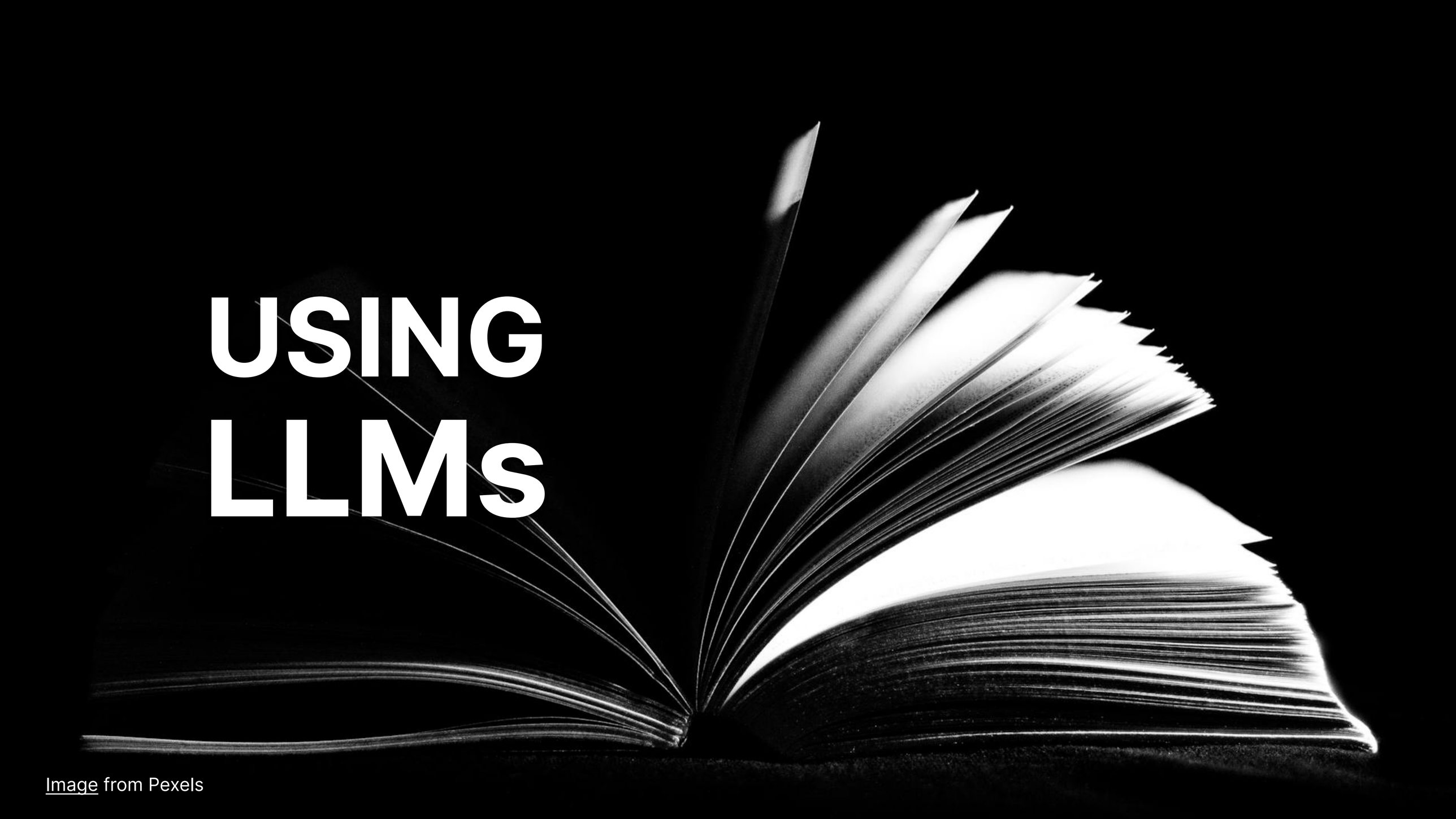


Image from Pexels

Natural language tasks

Translation

I have visited Italy.

↓

Ich habe Italien besucht.

Text classification

I have visited Italy.

↓

Pets | Travel | Tech

Summarisation

Italy, officially the Italian Republic, is a country in Southern and Western Europe. It is located on a peninsula that extends into the middle of the Mediterranean Sea, with the Alps on its northern land border, as well as islands, notably Sicily and Sardinia. It is the tenth-largest country in Europe, covering an area of 301,340 km2 (116,350 sq mi), and third-most populous member state of the European Union, with a population of nearly 60 million.

↓

Italy is a large European country.

Question answering

What is the capital of Italy?

↓

Rome.

Question answering

Parametric knowledge

Models are trained over data which contains facts and opinions.

Large enough models encode some of this knowledge.

Fine-tuning

Domain-specific dataset is constructed.

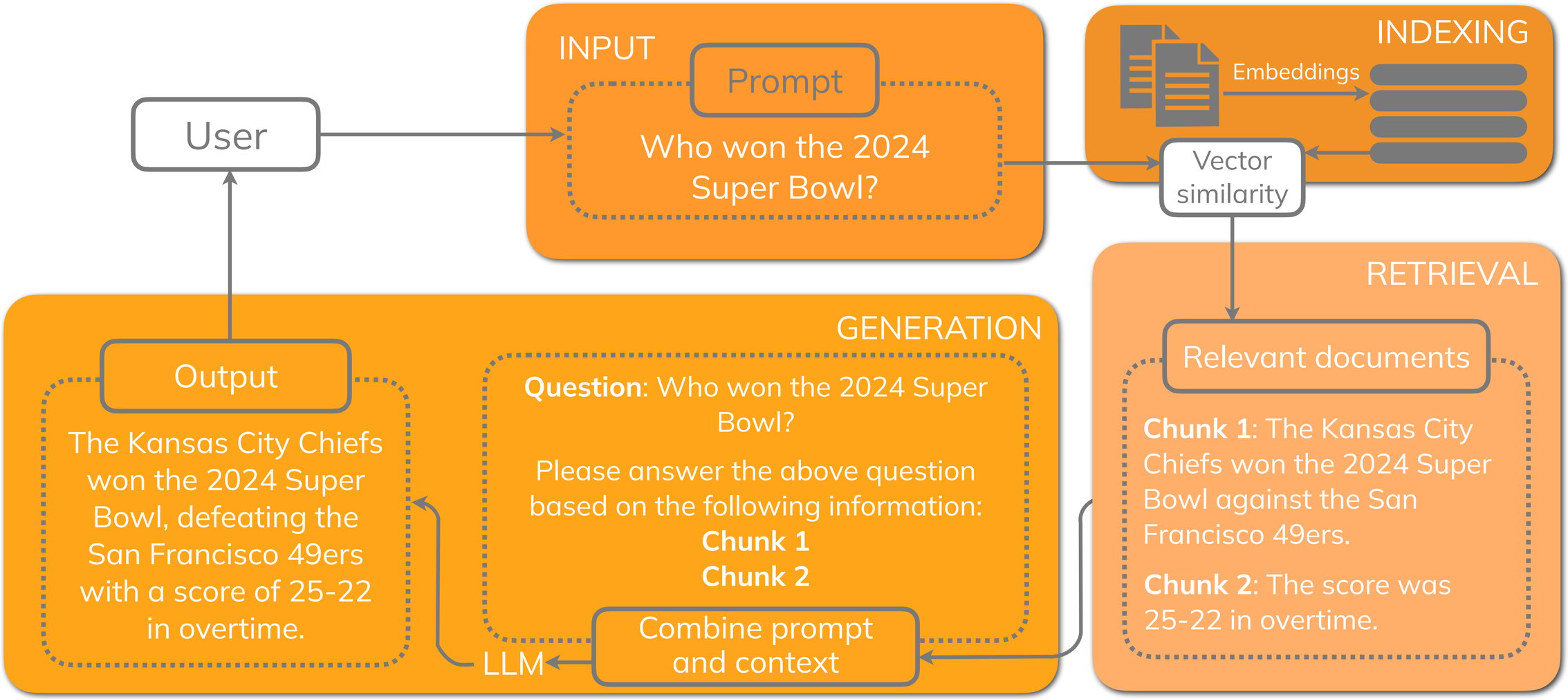
LLM is further trained so outputs are in line with this dataset.

Retrieval augmented generation (RAG)

Context that is similar to the input is gathered.

LLM is instructed to use this as part of evaluation.

Retrieval Augmented Generation



Paper source

RAG (and LLM deployment) ain't easy

ISSUES WITH RAG

Chunk size

Creating document embeddings

Retrieval method

Creation of prompt

ISSUES WITH LLMs

LLM is not able to complete target task well

Poor quality answers
Hallucinations

Paper source

🤗 Open LLM Leaderboard

🏅 LLM Benchmark | 📈 Metrics through time | 📝 About | ❗ FAQ | 🚀 Submit

T	Model	Average ⬆️	ARC	HellaSwag	MMLU	TruthfulQA	Winogrande	GSM8K
🔶	davidkim205/Rhea-72b-v0.5	81.22	79.78	91.15	77.95	74.5	87.85	76.12
💬	MTSAIR/MultiVerse_70B	81	78.67	89.77	78.22	75.18	87.53	76.65
🔶	MTSAIR/MultiVerse_70B	80.98	78.58	89.74	78.27	75.09	87.37	76.8
🔶	abacusai/Smaug-72B-v0.1	80.48	76.02	89.27	77.15	76.67	85.08	78.7
🔶	ibivibiv/alpaca-dragon-72b-v1	79.3	73.89	88.16	77.4	72.69	86.03	77.63
💬	mistralai/Mixtral-8x22B-Instruct-v0.1	79.15	72.7	89.08	77.77	68.14	85.16	82.03
💬	MaziyarPanahi/Llama-3-70B-Instruct-DPO-v0.2	78.96	72.53	86.22	80.41	63.57	82.79	88.25
💬	MaziyarPanahi/Llama-3-70B-Instruct-DPO-v0.4	78.89	72.61	86.03	80.5	63.26	83.58	87.34
💬	MaziyarPanahi/Llama-3-70B-Instruct-DPO-v0.3	78.74	72.35	86	80.47	63.45	82.95	87.19
💬	mmnga/Llama-3-70B-japanese-suzume-vector-v0.1	78.6	72.35	85.81	80.28	62.93	82.79	87.41
💬	moreh/MoMo-72B-lora-1.8.7-DPO	78.55	70.82	85.96	77.13	74.71	84.06	78.62
💬	tenyx/Llama3-TenyxChat-70B	78.4	72.1	86.21	80.04	62.85	82.95	86.28

📙 Citation ◀

Use via API 🚀 · Built with Gradio

Hugging Face Open LLM Leaderboard



Thank you!



@t_redactyl



@t_redactyl@fosstodon.org



t-redactyl.io