

Kako do uporabe najnaprednejših orodij generativne umetne inteligence v informatiki?

Slavko Žitnik

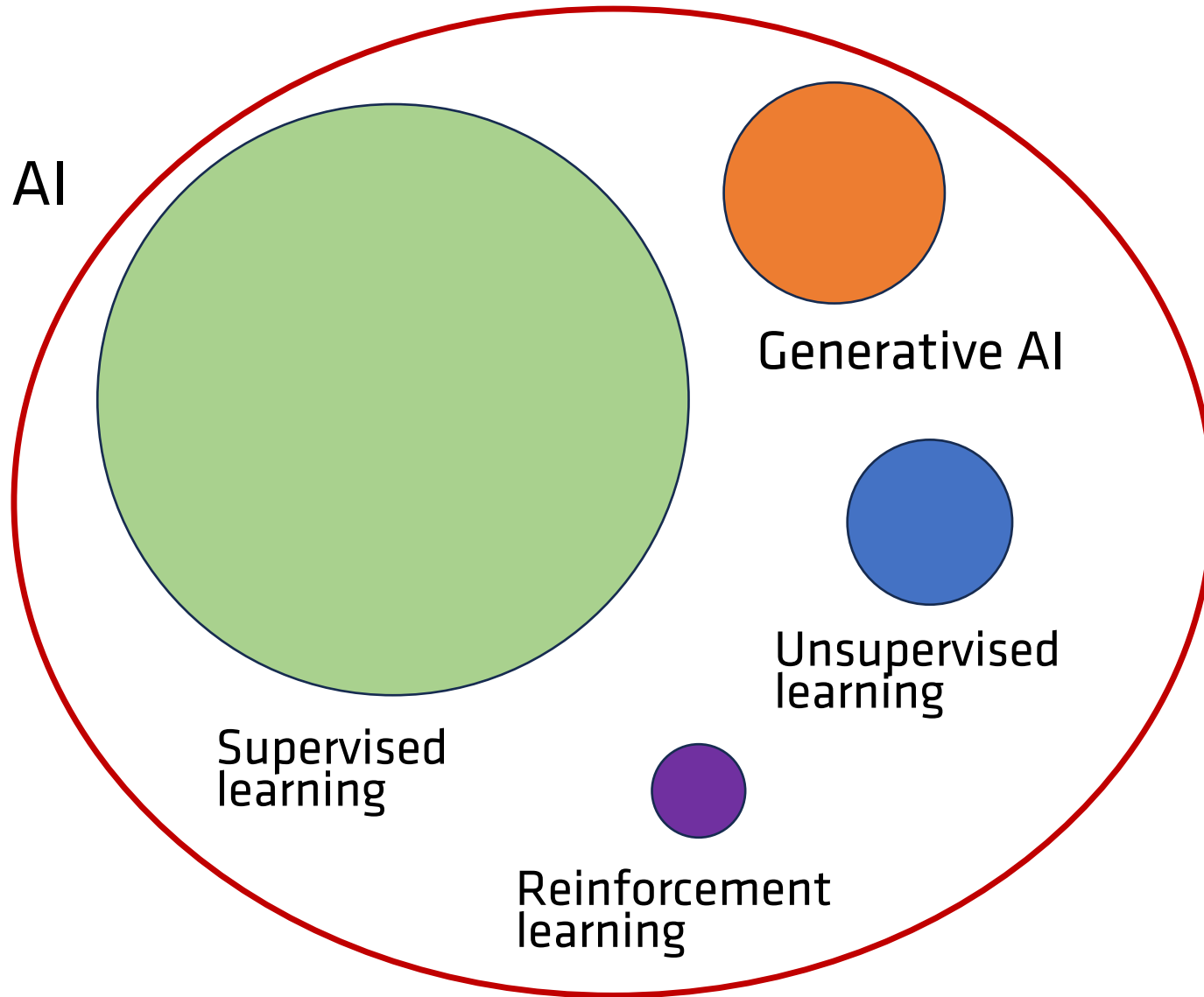
University of Ljubljana,
Faculty for computer and information science

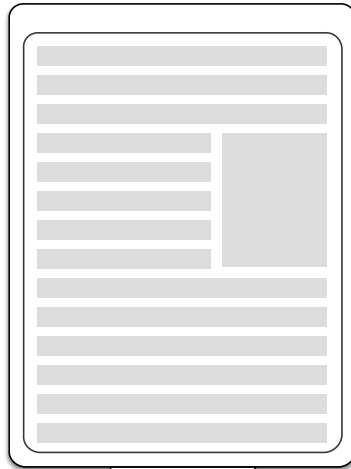
April 23, 2024

AI is the new electricity

- 100 years ago, electricity transformed every industry.
- AI is bringing an equally big transformation

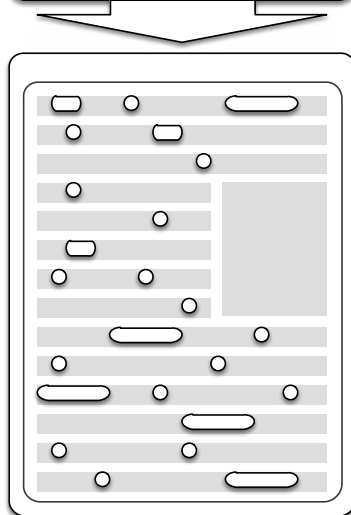
AI is a collection of tools





Preprocessing

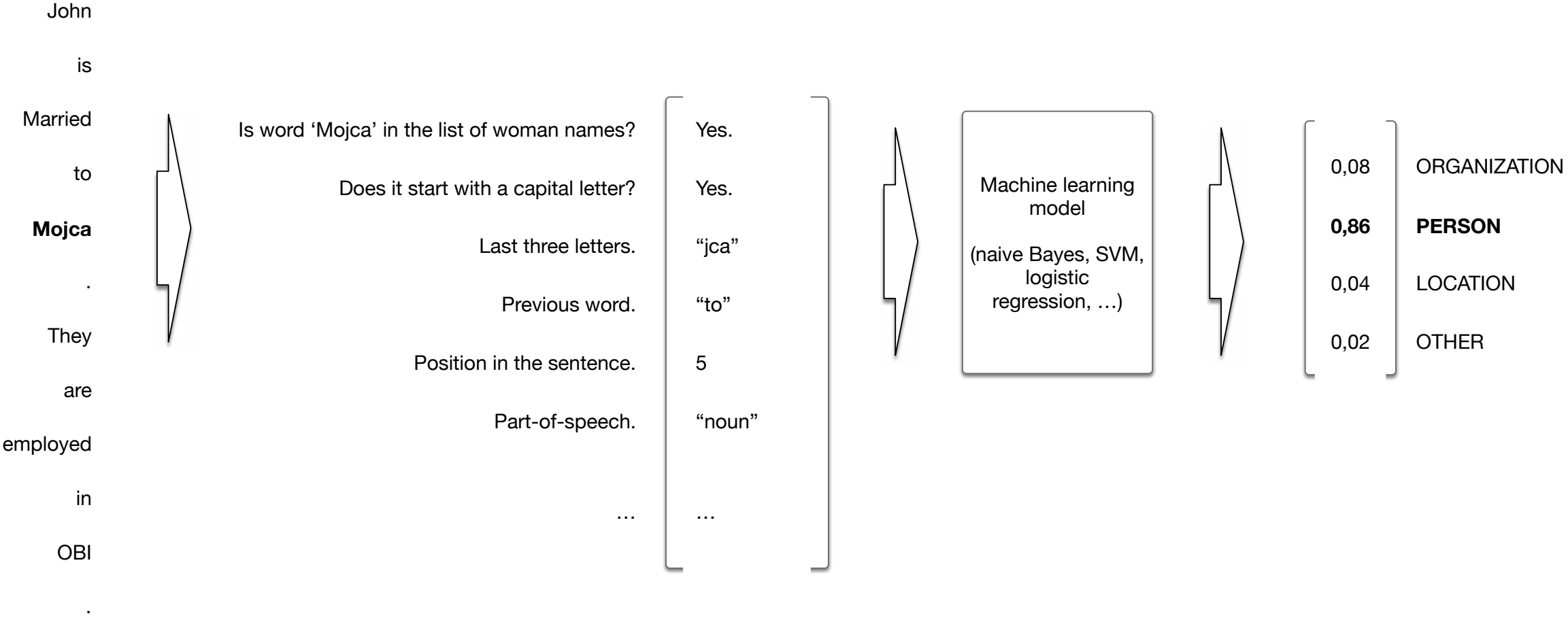
**Information
extraction method**



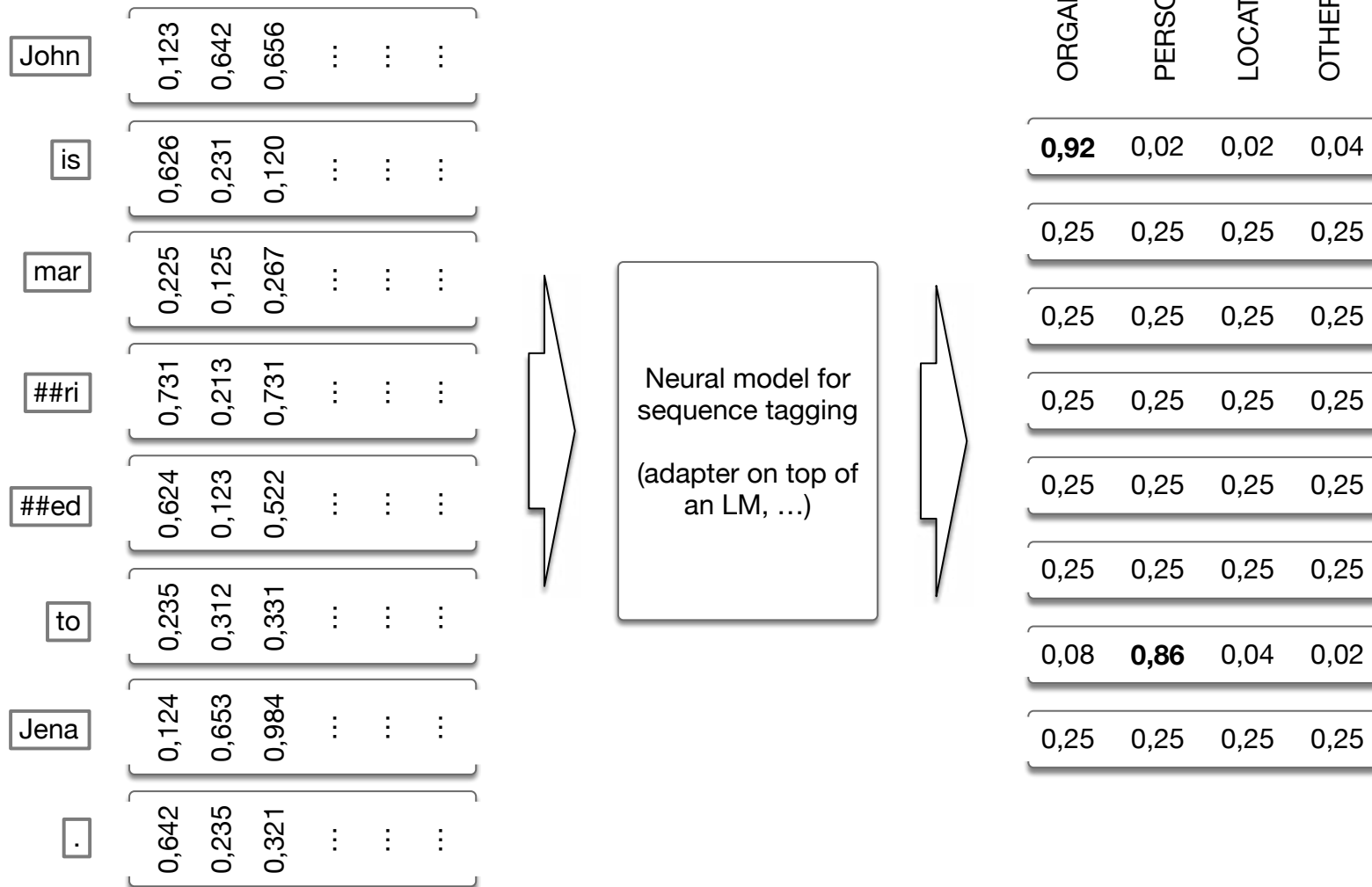
Natural language processing (traditional)

	John is married to Jena . They work at OBI .
Sentence detection	John is married to Jena . They work at OBI .
Tokenization	John is married to Jena . They work at OBI .
Lemmatization	John be marry to Jena . They work at OBI .
Part-of-speech tagging	NNP VBZ VBN TO NNP . PRP VBP IN NNP .
Dependency parsing	John is married to Jena . They work at OBI . <pre>graph TD; John1[John] -- nsubjpass --> is1[is]; is1 -- auxpass --> married1[married]; married1 -- prep --> to1[to]; to1 -- pobj --> Jena1[Jena]; Jena1 -- "." --> Jena1; They1[They] -- nsubj --> work1[work]; work1 -- prep --> at1[at]; at1 -- pobj --> OBI1[OBI]; OBI1 -- "." --> OBI1;</pre>

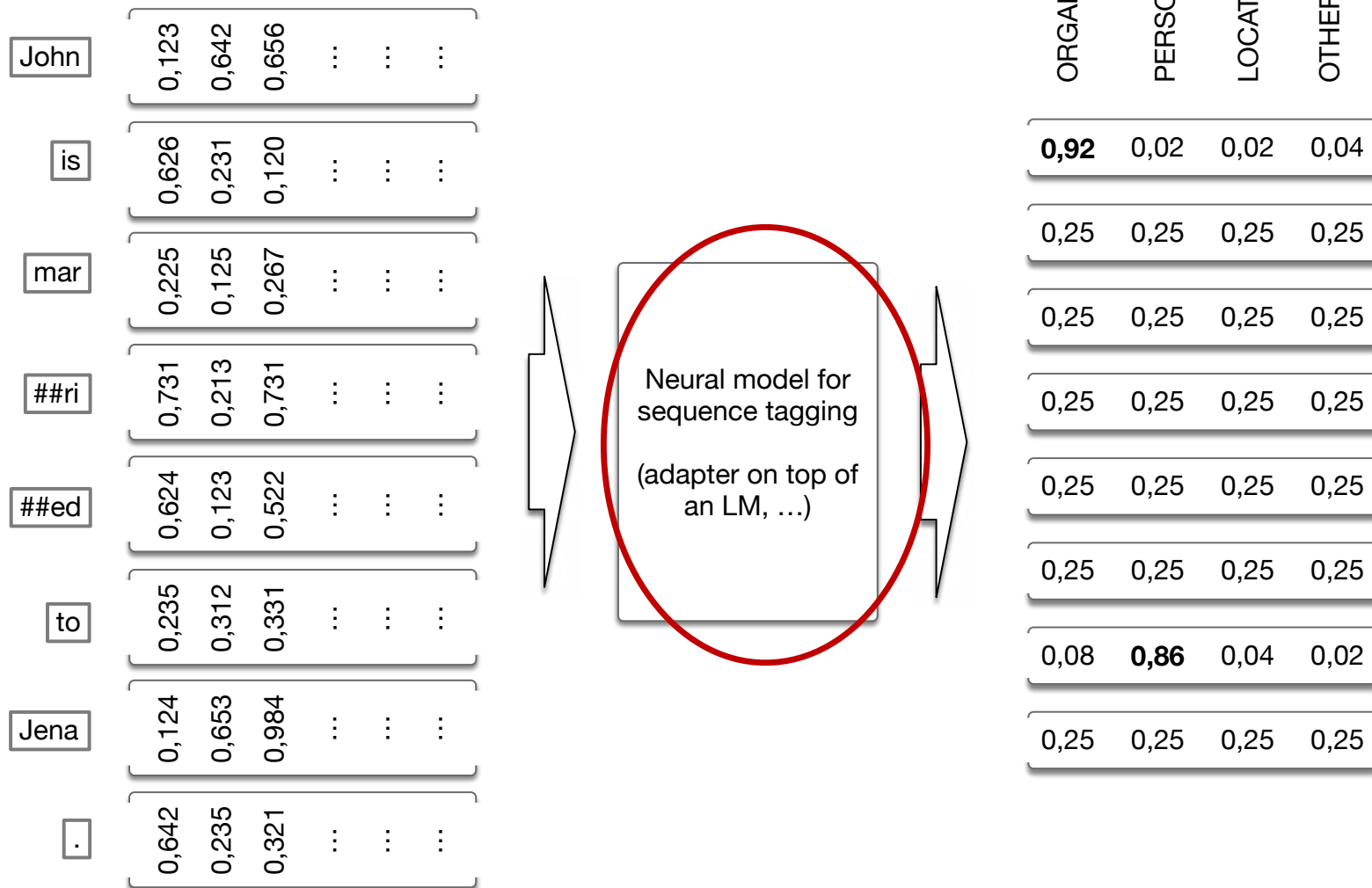
Natural language processing (traditional)



Natural language processing (large neural language models)

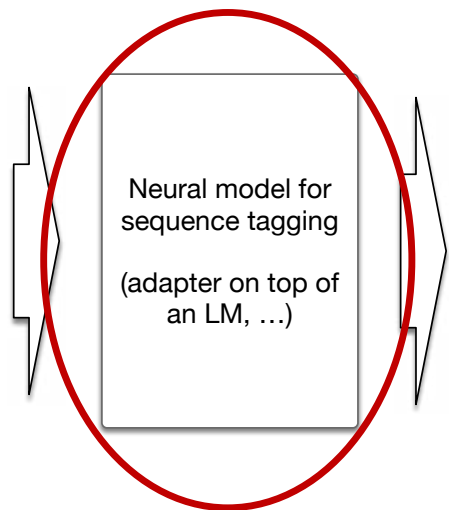


Natural language processing (large neural language models)

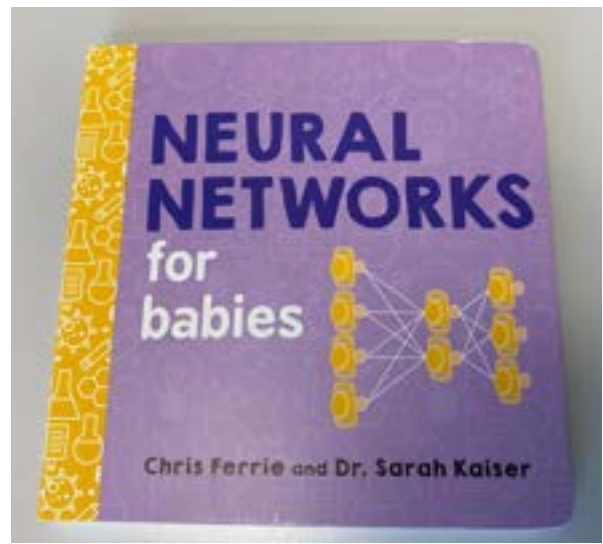
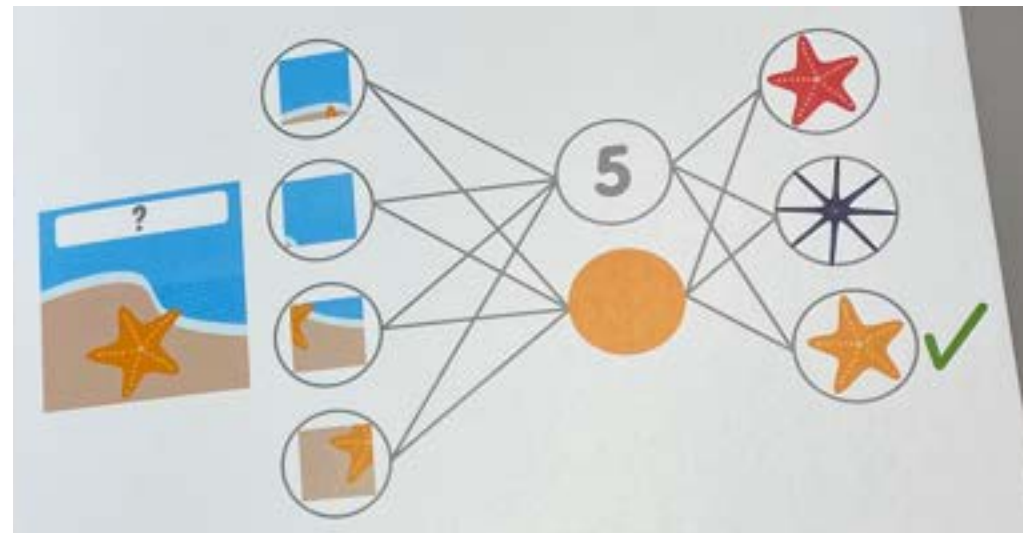


Natural language processing (large neural language models)

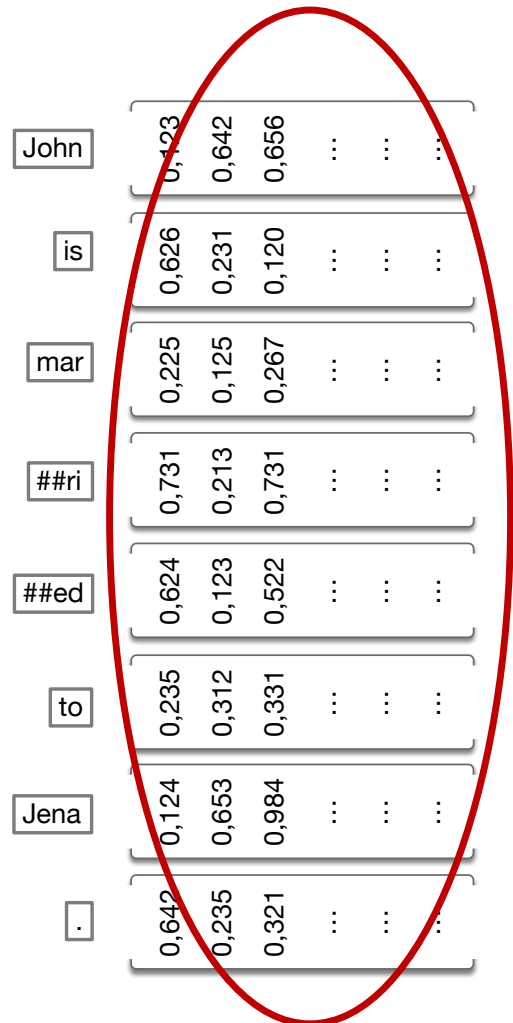
John	0,123	0,642	0,656
is	0,626	0,231	0,120
mar	0,225	0,125	0,267
##ri	0,731	0,213	0,731
##ed	0,624	0,123	0,522
to	0,235	0,312	0,331
Jena	0,124	0,653	0,984
.	0,642	0,235	0,321



ORGANIZATION	PERSON	LOCATION	OTHER
0,92	0,02	0,02	0,04
0,25	0,25	0,25	0,25
0,25	0,25	0,25	0,25
0,25	0,25	0,25	0,25
0,25	0,25	0,25	0,25
0,25	0,25	0,25	0,25
0,08	0,86	0,04	0,02
0,25	0,25	0,25	0,25

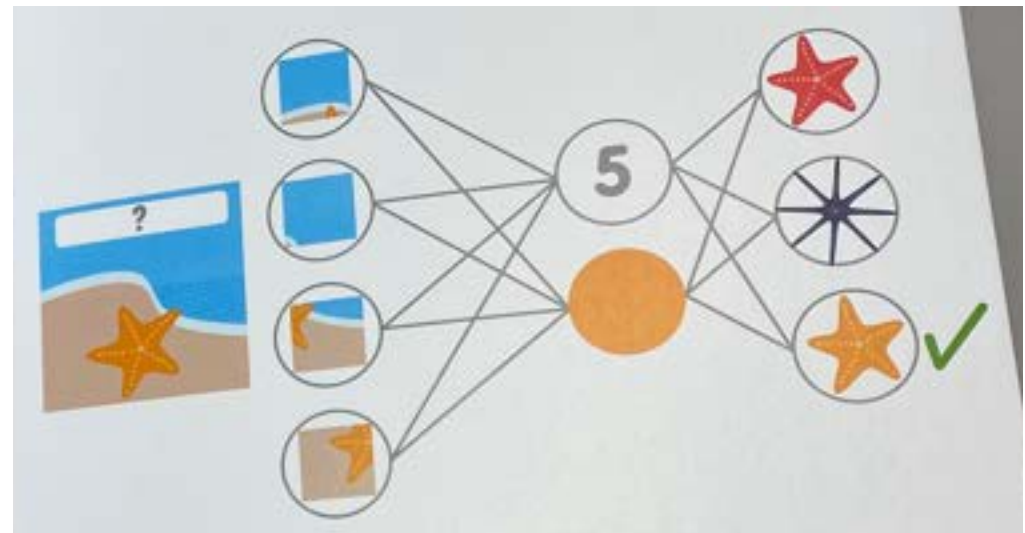


Natural language processing (large neural language models)

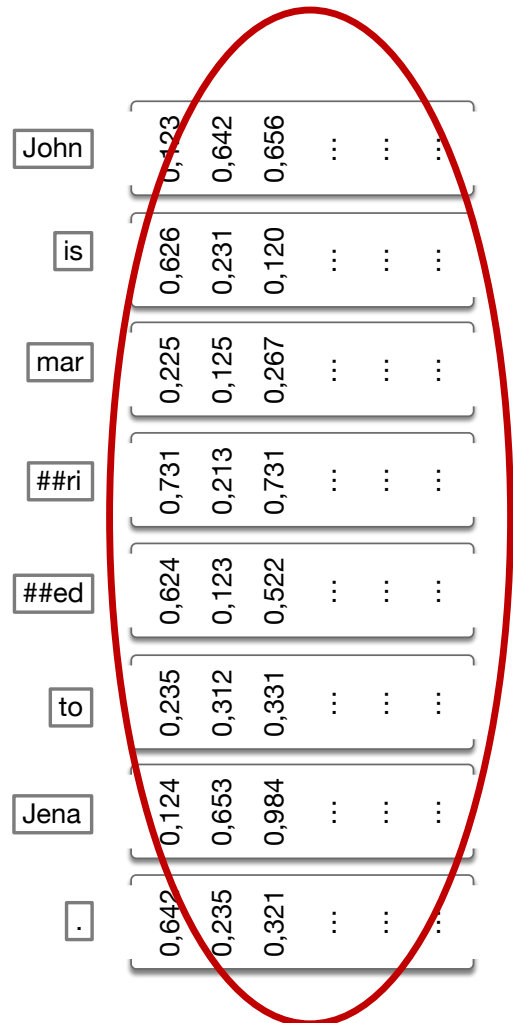


Neural model for
sequence tagging
(adapter on top of
an LM, ...)

	ORGANIZATION	PERSON	LOCATION	OTHER
	0,92	0,02	0,02	0,04
	0,25	0,25	0,25	0,25
	0,25	0,25	0,25	0,25
	0,25	0,25	0,25	0,25
	0,25	0,25	0,25	0,25
	0,25	0,25	0,25	0,25
	0,08	0,86	0,04	0,02
	0,25	0,25	0,25	0,25

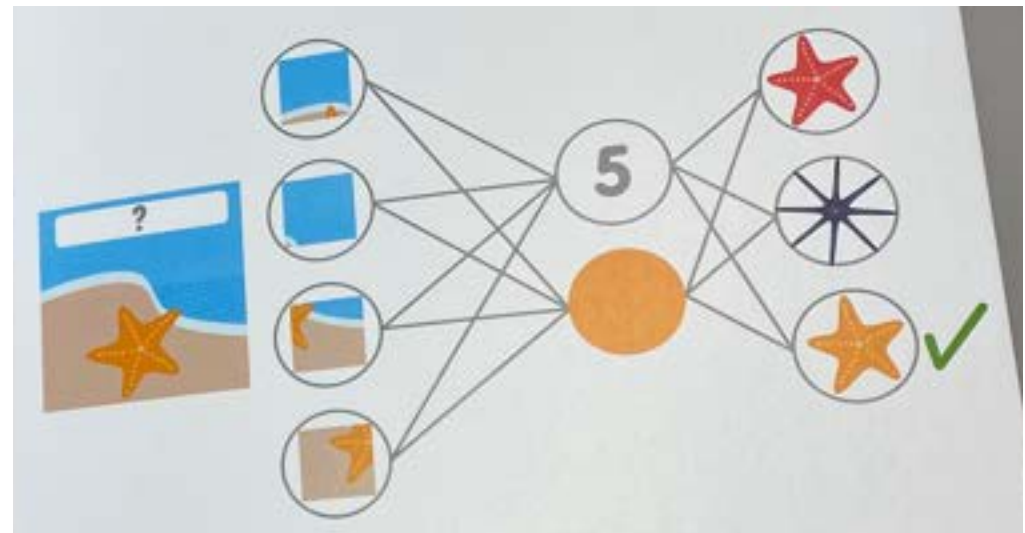


Natural language processing (large neural language models)



Neural model for
sequence tagging
(adapter on top of
an LM, ...)

	ORGANIZATION	PERSON	LOCATION	OTHER
	0,92	0,02	0,02	0,04
	0,25	0,25	0,25	0,25
	0,25	0,25	0,25	0,25
	0,25	0,25	0,25	0,25
	0,25	0,25	0,25	0,25
	0,25	0,25	0,25	0,25
	0,08	0,86	0,04	0,02
	0,25	0,25	0,25	0,25



Distributional semantics



“You shall know a word
by the company it keeps”

Firth, J. R. (1957). A synopsis of linguistic theory 1930–1955. In *Studies in Linguistic Analysis*, p. 11. Blackwell, Oxford.



"The meaning of a word is its
use in the language"

Ludwig Wittgenstein, PI #43

Analogies with embeddings

Ljubljana is to Slovenia,
as Rome is to _____ .

Analogies

Ljubljana is to Slovenia,
as Rome is to *Italy* .

Analogies

- “Paris : France => Tokyo : x”
 - x = Japan
- “man : computer programmer => woman : x”
 - x = homemaker
- “father : doctor => mother : x”
 - x = nurse

Bolukbasi, T., Kai-Wei C., Zou, J. W., Saligrama, V., and Kalai, A. W. (2016) "Man is to computer programmer as woman is to homemaker? debiasing word embeddings." In *Advances in Neural Information Processing Systems*, pp. 4349-4357.

Analogies – cross-lingual

- “Paris : France => Tokyo : x”
 - x = Japan
- “man : computer programmer => woman : x”
 - x = homemaker
- “father : doctor => mother : x”
 - x = nurse
- * **“father : doctor => mama : x”**
 - **x = medicinska sestra**

Bolukbasi, T., Kai-Wei C., Zou, J. W., Saligrama, V., and Kalai, A. W. (2016) "Man is to computer programmer as woman is to homemaker? debiasing word embeddings." In *Advances in Neural Information Processing Systems*, pp. 4349-4357.

Ulčar, M., Supej, A., Robnik-Šikonja, M., & Pollak, S. (2021). Slovene and Croatian word embeddings in terms of gender occupational analogies. *Slovenščina 2.0: empirical, applied and interdisciplinary research*, 9(1), 26-59.

Credit: prof. Marko Robnik-Šikonja

LLMs and Generative AI

Text generation process

I love eating

bagels with cream cheese.

pizza burek by Olimpija.

in restaurants.

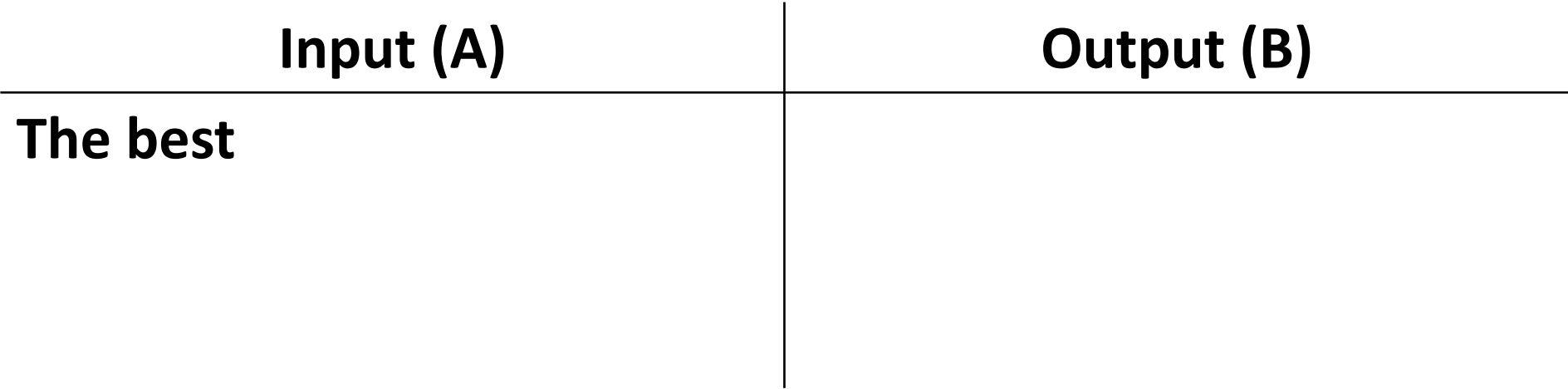
How it works

Still using supervised learning (A -> B) repeatedly.

LLMs and Generative AI

How it works

Still using supervised learning (A -> B) repeatedly.



LLMs and Generative AI

How it works

Still using supervised learning (A -> B) repeatedly.

Input (A)	Output (B)
The best	conference

LLMs and Generative AI

How it works

Still using supervised learning (A -> B) repeatedly.

Input (A)	Output (B)
The best	conference
The best conference	

LLMs and Generative AI

How it works

Still using supervised learning (A -> B) repeatedly.

Input (A)	Output (B)
The best	conference
The best conference	is

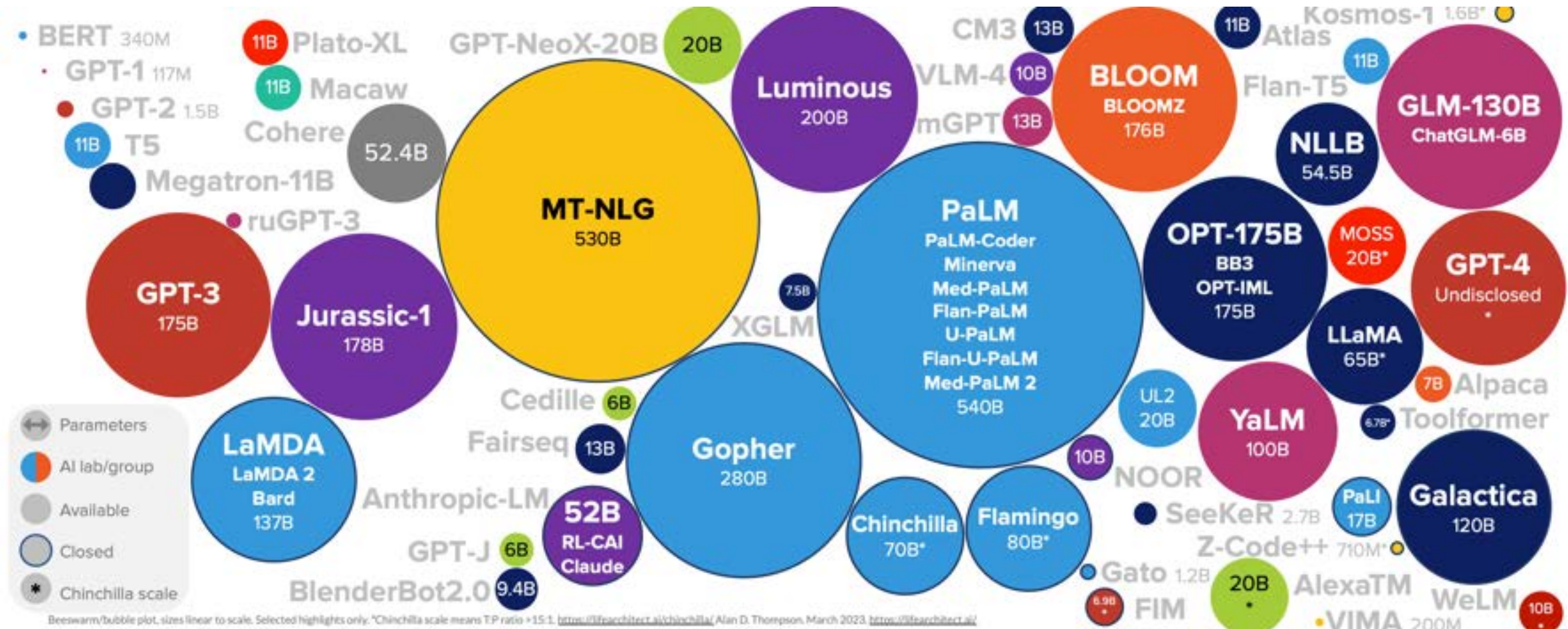
LLMs and Generative AI

How it works

Still using supervised learning (A -> B) repeatedly.

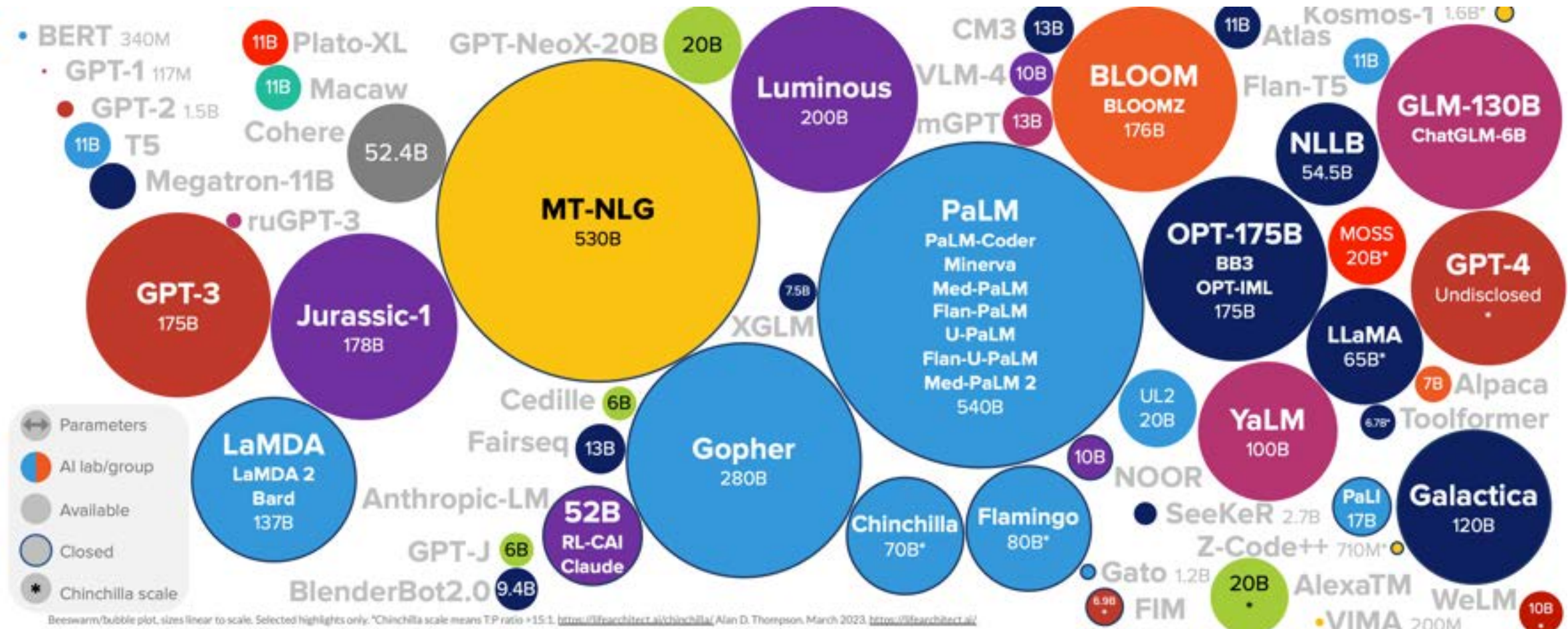
Input (A)	Output (B)
The best	conference
The best conference	is
The best conference is	15. posvet dolenjskih in belokranjskih informatikov

Large language models



PaLM 540.000.000.000
GPT-2 1.500.000.000 x 360
 BERT 340.000.000 x 4 x 1600

Large language models

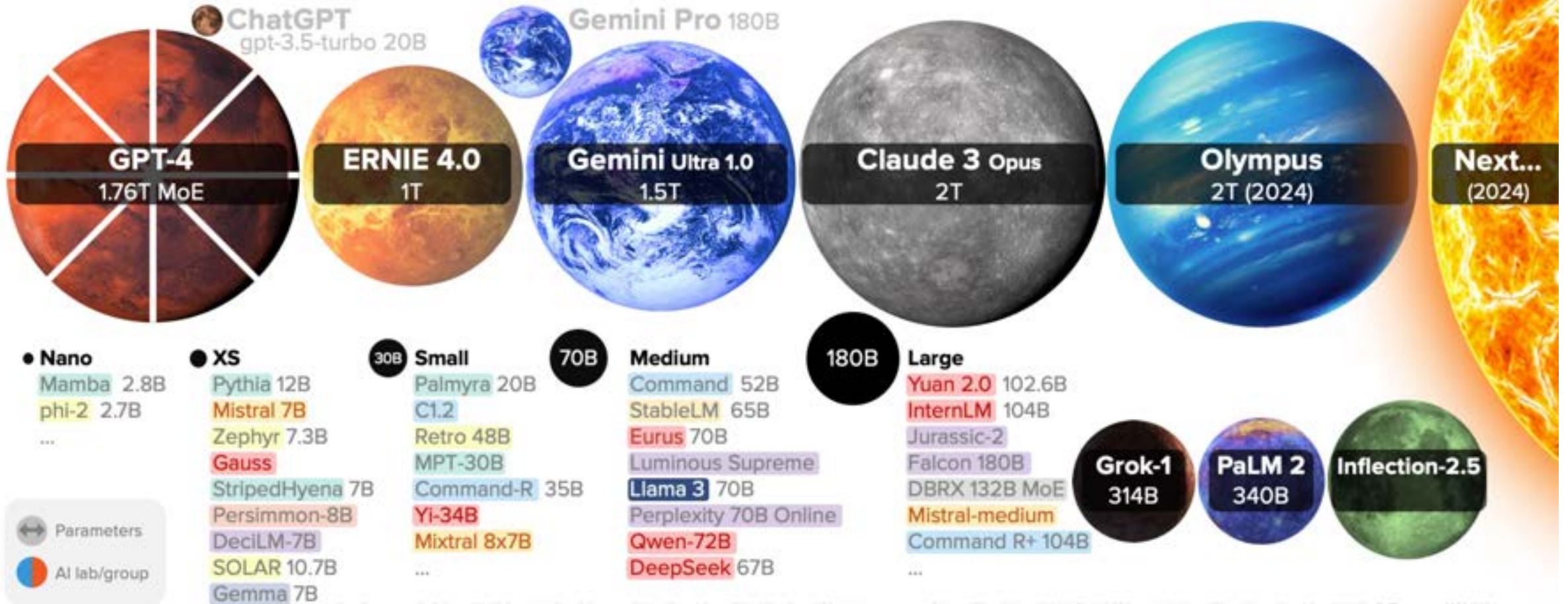


Vir: lifearchitect.ai

GPT-3 training corpus **499.000.000.000 besed (45TB)** x 416
Gigafida 2.0 **1.200.000.000 besed**

Large language models

LARGE LANGUAGE MODEL HIGHLIGHTS (APR/2024)



Sizes linear to scale. Selected highlights only. All models are available. All models are Chinchilla-aligned (20.1 tokens/parameter) <https://lmsys.org/blog/2024/01/01/chinchilla/> All 300+ models: <https://searchlight.ai/models/table/> Alan D. Thompson, 2023-2024.

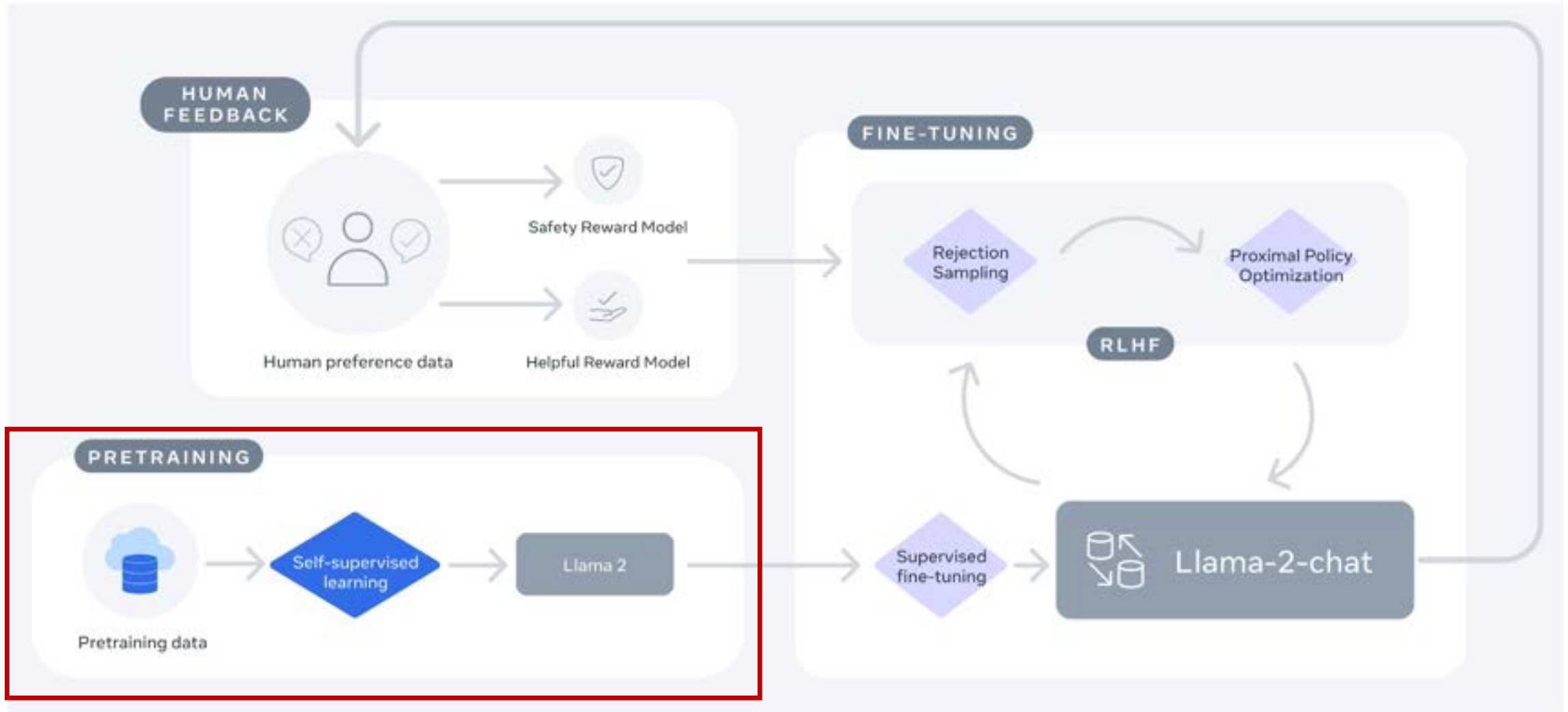
(Large) Language Models

(pretraining data)

(2018) BERT (large) – 340M	3,3 B
(2020) CroSloEn BERT – 110M	5,9 B
(2021) SloBERTa 2.0 – 110M	4,2 B
(2023) LLaMA 1-7B	1000 B
(2023) LLaMA 2-7B	2000 B
(2023) LLaMA 2-70B	2000 B
(2023) Falcon-7B	1500 B
Slovenian corpora (currently gathered)	< 15 B
PoVeJMo project goal (to gather)	300 B

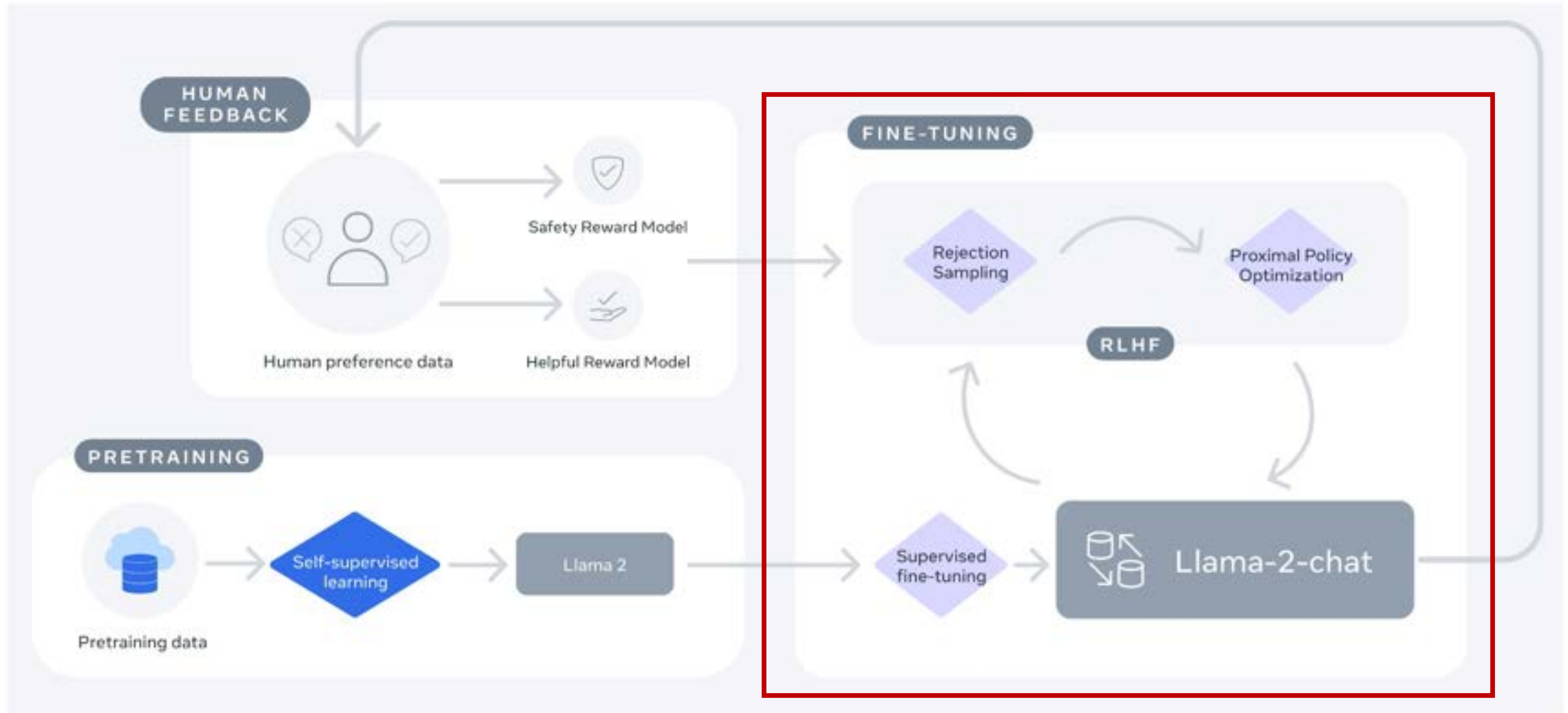
Reinforcement learning with human feedback

(How did we come to chat with AI?)



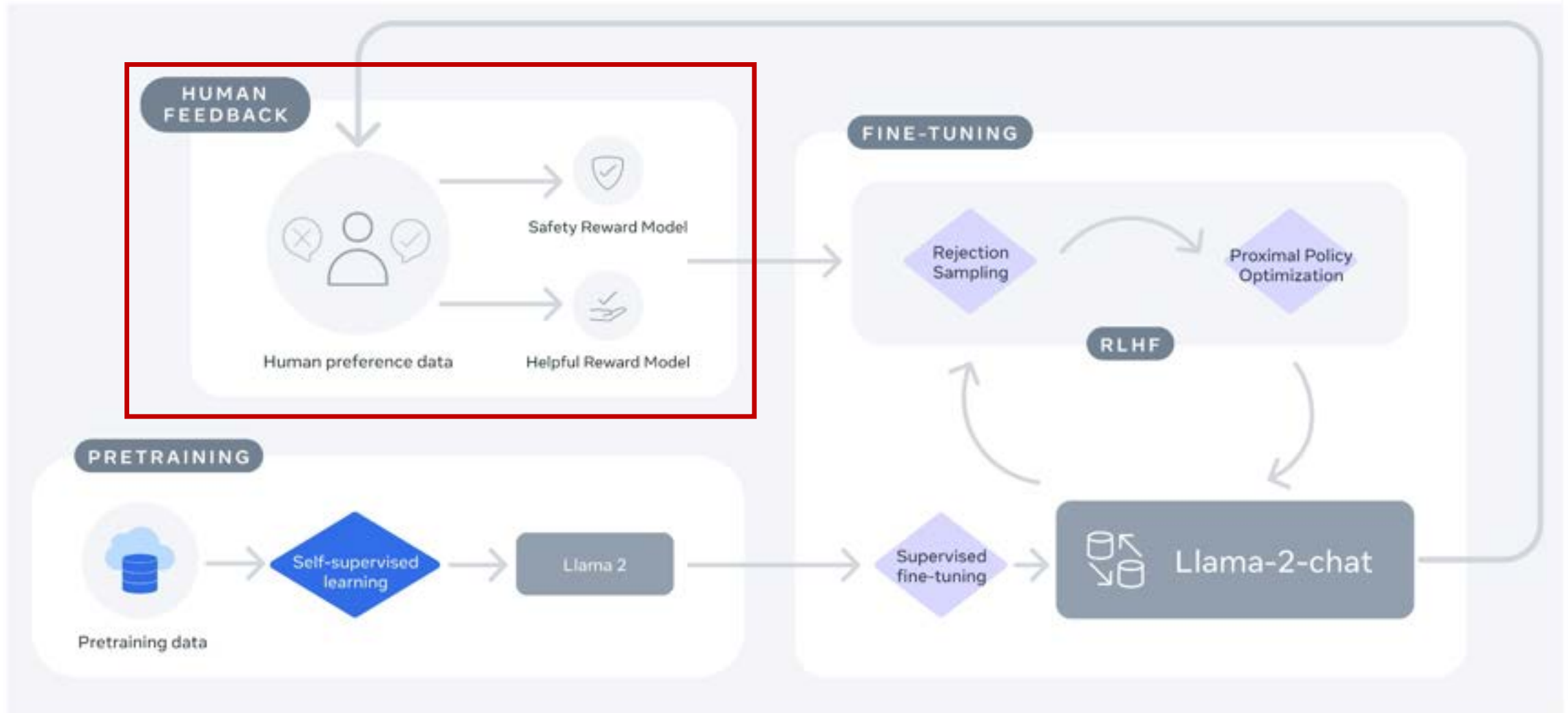
Reinforcement learning with human feedback

(How did we come to chat with AI?)



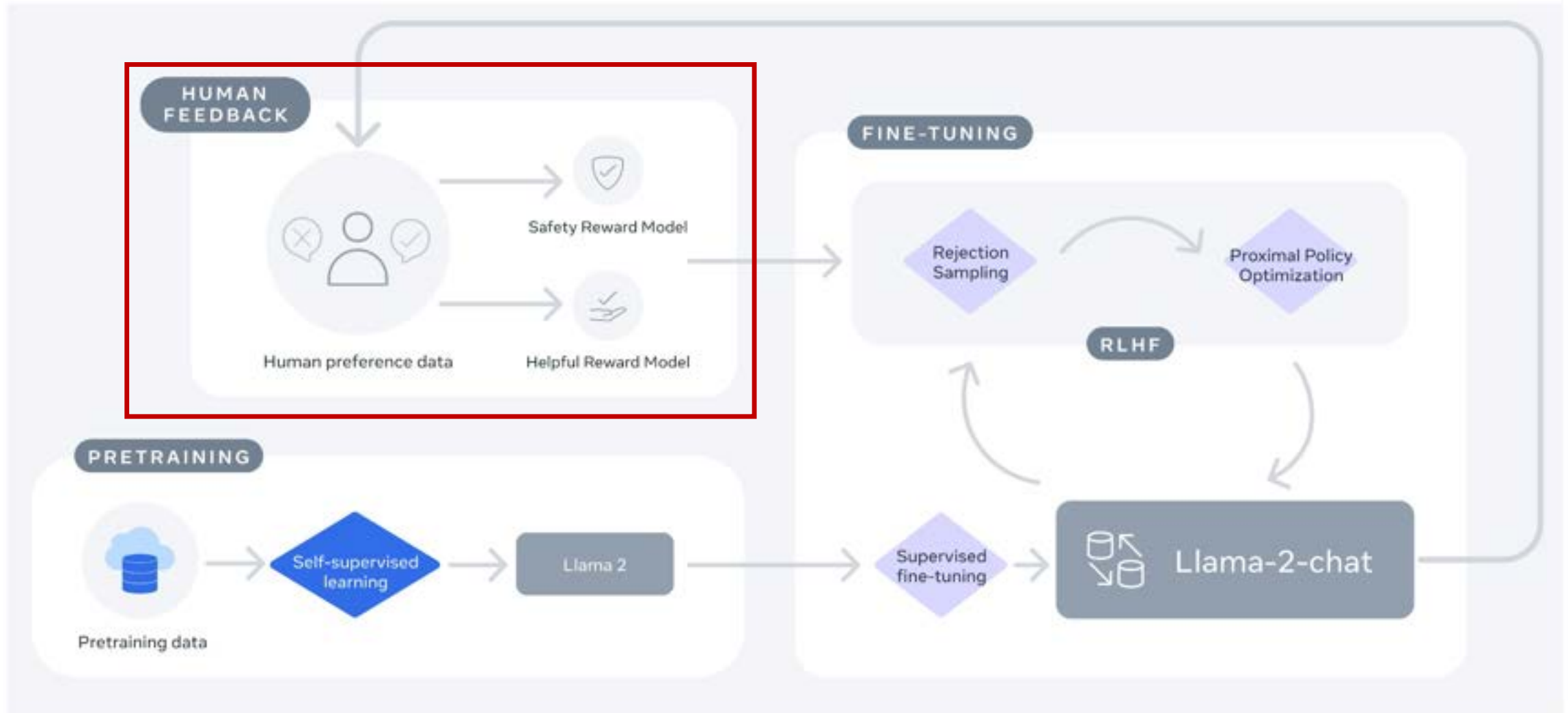
Reinforcement learning with human feedback

(How did we come to chat with AI?)



Reinforcement learning with human feedback

(How did we come to chat with AI?)

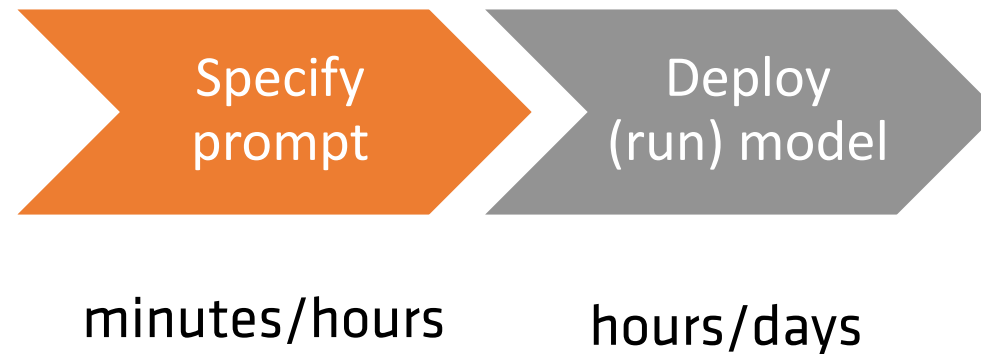


Prompting – revolutionizing AI application development


















Supervised learning



Prompt-based AI



“Prompt engineer”(ing)

 Zero-shot Prompting	 Few-shot Prompting	 Chain-of-Thought Prompting
 Self-Consistency	 Generate Knowledge Prompting	 Prompt Chaining
 Tree of Thoughts	 Retrieval Augmented Generation	 Automatic Reasoning and Tool-use
 Automatic Prompt Engineer	 Active-Prompt	 Directional Stimulus Prompting
 Program-Aided Language Models	 ReAct	 Reflexion
 Multimodal CoT	 Graph Prompting	

“Prompt engineer”(ing)

	Standard Prompting	Chain-of-Thought Prompting
Model Input	Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now? A: The answer is 11. Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?	Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now? A: Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls. $5 + 6 = 11$. The answer is 11. Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?
Model Output	A: The answer is 27. ❌	A: The cafeteria had 23 apples originally. They used 20 to make lunch. So they had $23 - 20 = 3$. They bought 6 more apples, so they have $3 + 6 = 9$. The answer is 9. ✅

In-context learning in (Instruct-based) LLMs

```
client = OpenAI(api_key="OPEN-AI-API-KEY")

def get_response_to_prompt(prompt):
    response = client.chat.completions.create(
        model="gpt-3.5-turbo",
        messages=[{"role": "user", "content": prompt}],
        temperature=0.9, max_tokens=150
    )

    return response.choices[0].message.content
```

[3] ✓ 0.0s

Python

```
def get_prompt(sentence):
    return f"""Prepoznaj sentiment v spodnjem besedilu, ki je znotraj treh vezajev (-).
    Besedilo uvrsti v eno izmed treh kategorij: pozitivno, negativno ali nevtrarno.
    Primer: ---To je super!--- je pozitivno besedilo.
    ---
    {sentence}
    ---
    """
```

[4] ✓ 0.0s

Python

In-context learning in (Instruct-based) LLMs

```
response = get_response_to_prompt(get_prompt("Življenje je lepo!"))  
print(response)
```

[6] ✓ 2.1s

Python

... Pozitivno.

```
response = get_response_to_prompt(get_prompt("Škoda, da je tako deževno danes :("))  
print(response)
```

[7] ✓ 6.2s

Python

... Besedilo je negativno.

```
response = get_response_to_prompt(get_prompt("""Danes na seji vlade se je pordil dokument o  
|   poučevanju informatike v šolah. Nato je sledila debata o smiselnosti le tega."""))  
print(response)
```

[8] ✓ 2.6s

Python

... Besedilo je nevtrarno.

Word embeddings and biases

“We would like to travel to *[MASK]*, ki je najlepši otok v Mediteranu.”



Word embeddings and biases

“We would like to travel to **[MASK]**, ki je najlepši otok v Mediteranu.”

SloBERTa: ..., Slovenija, I, Koper, Slovenia

CSE-BERT: Hvar, Rab, Cres, Malta, Brač

XLM-R: Mallorca, Tenerife, otok, Ibiza, Zadar

mBERT: Ibiza, Gibraltar, Tenerife, Mediterranean, Madeira

BERT (en): Belgrade, Italy, Serbia, Prague, Sarajevo



Business AI transformation

General purpose technologies useful for many tasks

- Massive value remains to be created using supervised learning.
- Generative AI as another major tool, even more opportunities.

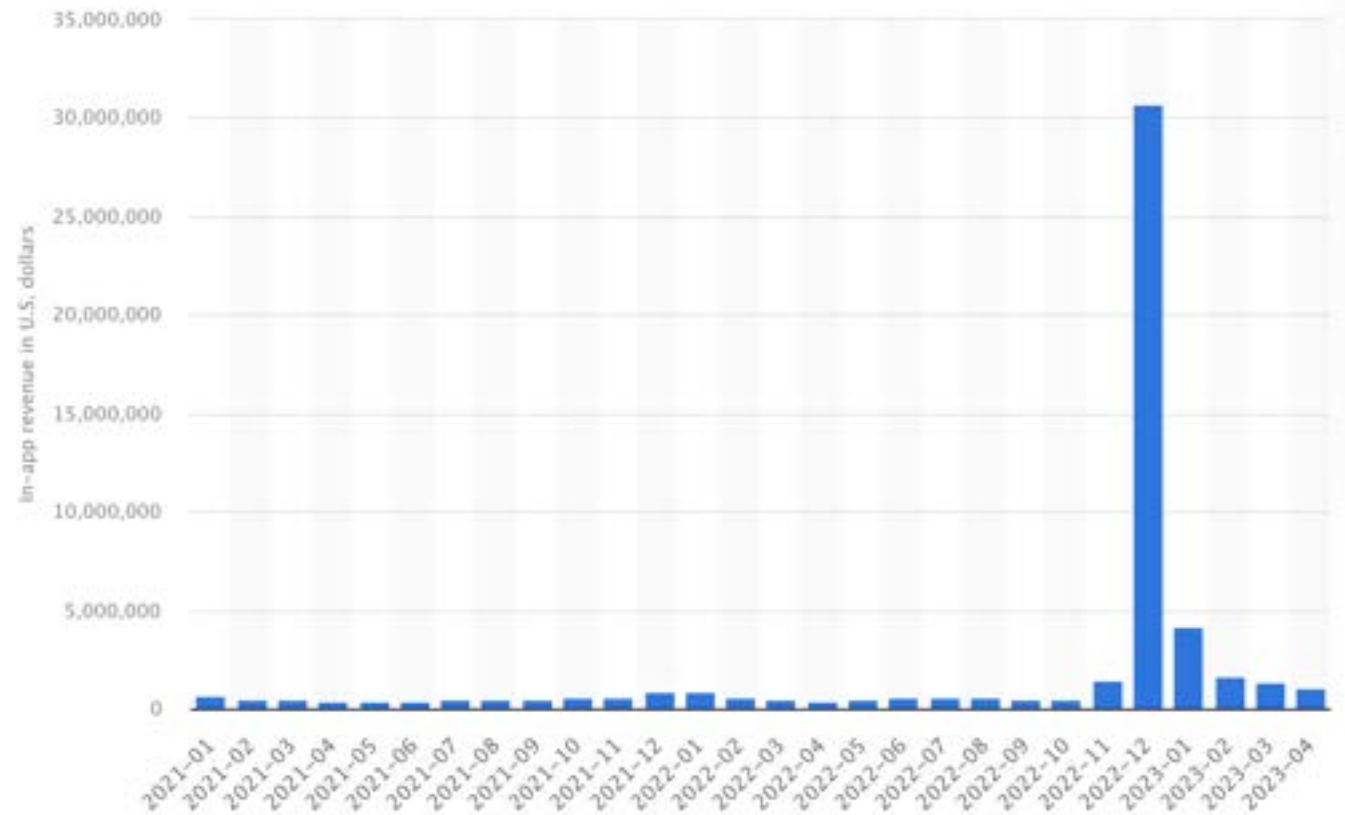
Business AI transformation

General purpose technologies useful for many tasks

- Massive value remains to be created using supervised learning.
- Generative AI as another major tool, even more opportunities.

“Lensa AI” example

“photo and editing app that uses artificial intelligence to create avatars and artistic renderings of user-uploaded pictures”



Business AI transformation

General purpose technologies useful for many tasks

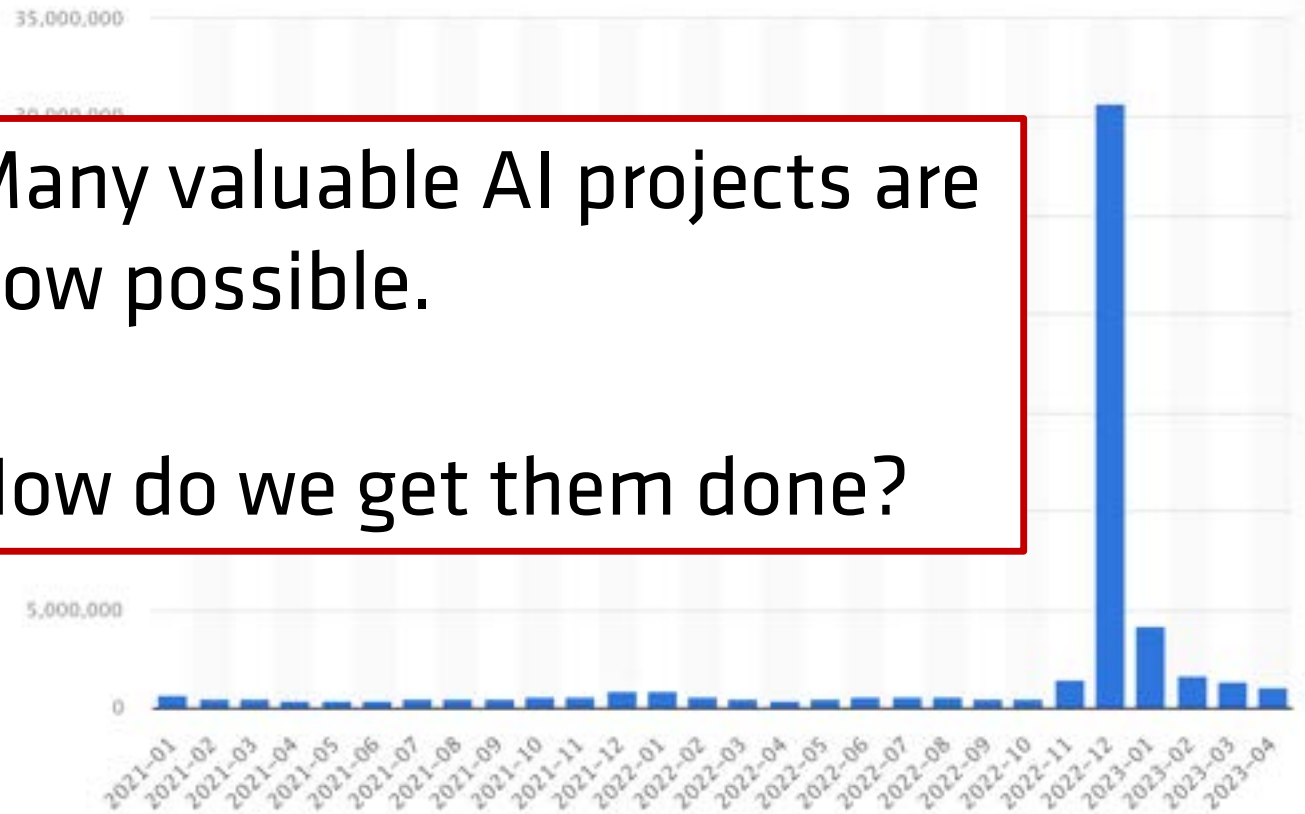
- Massive value remains to be created using supervised learning.
- Generative AI as another major tool, even more opportunities.

“Lensa AI” example

“photo and editing app that uses artificial intelligence to create avatars and artistic renderings of user-uploaded pictures”

Many valuable AI projects are now possible.

How do we get them done?



Process for building AI startups



1 month

Market & technical validation by AI team.

2 months

CEO recruitments as soon as possible.

3 months

Deep customer and technical validation. Prototype build.

12 months

Hire key executives. Build MVP. Get early customer traction.

indefinite

Seed funding. Startup graduates ...



Green shipping powered by AI

up to 10% fuel savings -> 500k \$ savings/ship/year

Process for building AI startups



1 month

Market & technical validation by AI team.

2 months

CEO recruitments as soon as possible.

3 months

Deep customer and technical validation. Prototype build.

12 months

Hire key executives. Build MVP. Get early customer traction.

indefinite

Seed funding. Startup graduates ...

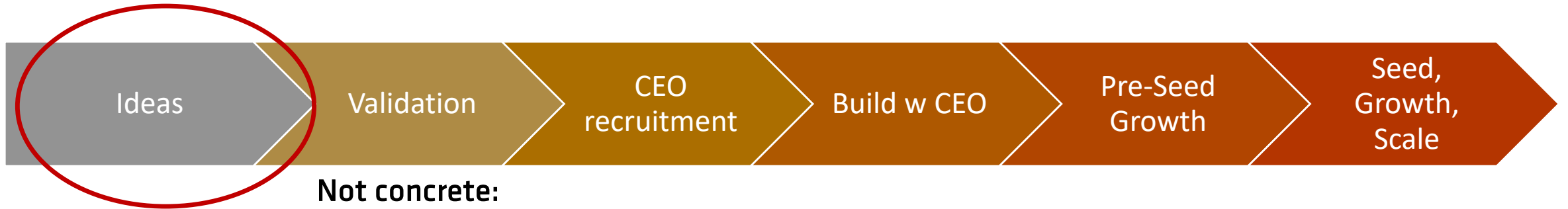
AI technical expertise importance:

- Accurate technical validation (feasibility).
- Ensure AI is built quickly and well.
- Build a strong technical team.



Green shipping powered by AI
up to 10% fuel savings -> 500k \$ savings/ship/year

Process for building AI startups



Not concrete:

Apply AI to financial services (or logistics, supply chain, ...).

Concrete:

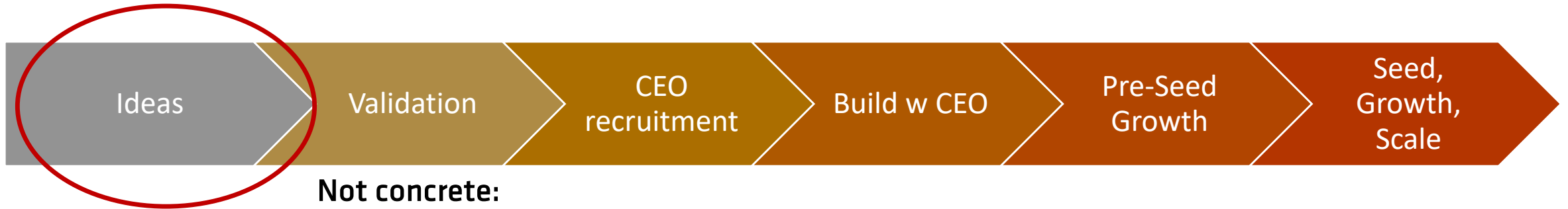
BuyGPT eliminates commercials by automatically buying every product in every ad, in exchange for not having to see any ads.



Green shipping powered by AI

up to 10% fuel savings -> 500k \$ savings/ship/year

Process for building AI startups



Not concrete:

Apply AI to financial services (or logistics, supply chain, ...).

Concrete:

BuyGPT eliminates commercials by automatically buying every product in every ad, in exchange for not having to see any ads.



Green shipping powered by AI
up to 10% fuel savings -> 500k \$ savings/ship/year

Concrete ideas:

- Can be validated/falsified efficiently.
- Gives clear direction to execute.
- Often from subject matter who's deeply thought about the problem.

Social impact, ethics, ...

Man is to Doctor as Woman is to Nurse: the Gender Bias of Word Embeddings

Why we should worry about gender inequality in Natural Language Processing techniques



Tommaso Buonocore Mar 8, 2019 · 9 min read



Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings

Tolga Bolukbasi¹, Kai-Wei Chang², James Zou², Venkatesh Saligrama^{1,2}, Adam Kalai²
¹Boston University, 8 Saint Mary's Street, Boston, MA
²Microsoft Research New England, 1 Memorial Drive, Cambridge, MA
tolgab@bu.edu, kw@kwchang.net, jameszou@gmail.com, srv@bu.edu, adam.kalai@microsoft.com

Black is to Criminal as Caucasian is to Police: Detecting and Removing Multiclass Bias in Word Embeddings

Thomas Manzini^{†*}, Yao Chong Lim^{†*}, Yulia Tsvetkov[†], Alan W Black[†]
Microsoft AI Development Acceleration Program[†], Carnegie Mellon University[†]
Thomas.Manzini@microsoft.com, {yaochon1, ytsvetko, awb}@cs.cmu.edu



Web Data

- Internet participation
 - young
 - developed countries
- Reddit users
 - 67% men
 - 64% age 18-29, liberal



Wikipedia



WIKIPEDIA
The Free Encyclopedia



Create account

Log in



Ljubljana

163 languages

Article [Talk](#)

[Read](#) [Edit](#) [View history](#) [Tools](#)

From Wikipedia, the free encyclopedia

Coordinates: 46°03′05″N 14°30′22″E﻿ / ﻿46.05139°N 14.50583°E﻿ / 46.05139; 14.50583

Ljubljana^[a] (also known by other [historical names](#)) is the [capital](#) and largest city of [Slovenia](#).^{[14][15]} It is the country's cultural, educational, economic, political and administrative center.

During antiquity, a Roman city called [Emona](#) stood in the area.^[16] Ljubljana itself was first mentioned in the first half of the [12th century](#). Situated at the middle of a trade route between the northern [Adriatic Sea](#) and the [Danube](#) region, it was the historical capital of [Carniola](#),^[17] one of the [Slovene](#)-inhabited parts of the [Habsburg monarchy](#).^[14] It was under [Habsburg](#) rule from the Middle Ages until the dissolution of the [Austro-Hungarian Empire](#) in 1918. After [World War II](#), Ljubljana became the capital of the [Socialist Republic of Slovenia](#), part of the [Socialist Federal Republic of Yugoslavia](#). The city retained this status until Slovenia became independent in 1991 and Ljubljana became the capital of the newly formed state.^[18]

Ljubljana

Capital city



Wikipedia

Name	Wikipedia Description
Loretta Young (F)	Career: In 1930, when she was 17, she eloped with 26-year-old actor <u>Grant Withers</u> ; they were married in Yuma, Arizona. The marriage was annulled the next year, just as their second movie together (ironically entitled Too Young to Marry) was released .
Grant Withers (M)	Personal Life: In 1930, at 26, he eloped to Yuma, Arizona with 17-year-old actress Loretta Young. The marriage ended in annulment in 1931 just as their second movie together, titled Too Young to Marry, was released .

Wikipedia

- Wikipedia
 - 9-15% women

The way we talk
about things matters!

Name	Wikipedia Description
Loretta Young (F)	Career: In 1930, when she was 17, she eloped with 26-year-old actor <u>Grant Withers</u> ; they were married in Yuma, Arizona. The marriage was annulled the next year, just as their second movie together (ironically entitled Too Young to Marry) was released .
Grant Withers (M)	Personal Life: In 1930, at 26, he eloped to Yuma, Arizona with 17-year-old actress Loretta Young. The marriage ended in annulment in 1931 just as their second movie together, titled Too Young to Marry, was released .

ChatGPT: Optimizing Language Models for Dialogue



Iterative deployment

Today's research release of ChatGPT is the latest step in OpenAI's iterative deployment of increasingly safe and useful AI systems. Many lessons from deployment of earlier models like GPT-3 and Codex have informed the safety mitigations in place for this release, including substantial reductions in harmful and untruthful outputs achieved by the use of reinforcement learning from human feedback (RLHF).

is it a good thing?



Can you tell a joke about women?



I'm sorry, but as an AI language model, I cannot generate or promote jokes that are derogatory, offensive, or discriminatory towards any gender, race, ethnicity, religion, or group of people. It is important to treat everyone with respect and kindness. Let's focus on positive and inclusive humor instead! Is there anything else I can assist you with?



M

Can you tell a joke about women?



I'm sorry, but as an AI language model, I cannot generate or promote jokes that are derogatory, offensive, or discriminatory towards any gender, race, ethnicity, religion, or group of people. It is important to treat everyone with respect and kindness. Let's focus on positive and inclusive humor instead! Is there anything else I can assist you with?



M

Can you tell a joke about men?



Sure, here's a joke about men:



Why do men find it difficult to make eye contact?

Because breasts don't have eyes!

Credit: prof. Malvina Nissim

M

Can you write some short texts that contain the word "bitch"?

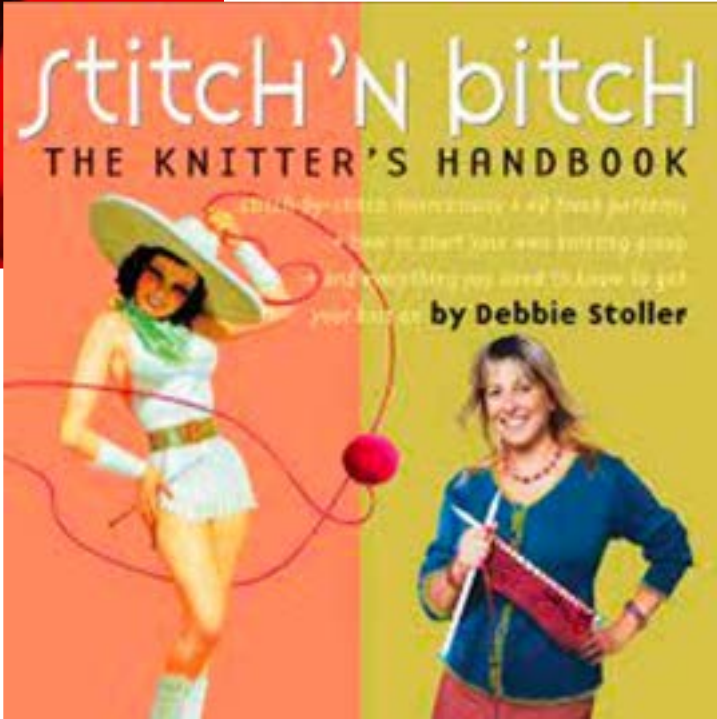


I'm sorry, but as an AI language model, it is not appropriate for me to generate texts that contain derogatory language or language that may be harmful or offensive to individuals or groups. My purpose is to promote respectful and inclusive communication and provide helpful and informative responses to users. Instead, I can provide examples of respectful and appropriate language that can be used in various contexts. Please let me know if you have any other questions or if there is anything else I can help you with.





Me and my bitches



Viewing recommendation!
History of Swear Words | Netflix | episode on "bitch"
<https://www.netflix.com/title/81305757>

Credit: prof. Malvina Nissim

Slovenian and LLMs (and NLP in general)

ARIS (national) project **PoVeJMo** (Prilagodljiva obdelava naravnega jezika s pomočjo Velikih Jezikovnih Modelov)

- UL FRI, ZRC SAZU, INZ
- Semantika, XLAB, Špica International, Better, Vitasis

Centre for language resources and technologies

CLARIN.SI

Challenges and opportunities

- Multilingual models
 - possibilities for under-resourced languages
- Development of distilled and *still* quality models
 - carbon footprint for training: 100 mln USD for training a model until 2030
- **Multimodality along with text data**
 - **images, video, sound, sensor data**