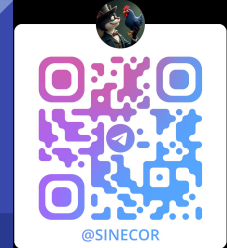


# State of AI



CS Space

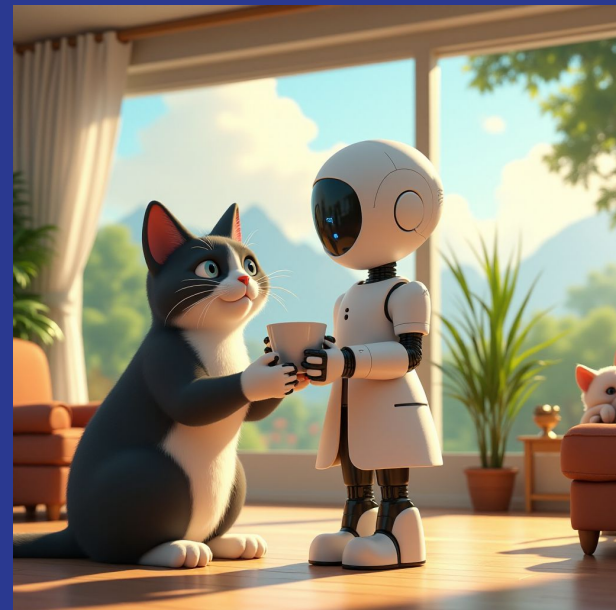
Клуб технологий и науки



Откуда мы пришли,  
кто мы и куда мы идём?

Сергей Николенко

16 марта 2025



# 1. Революции AI

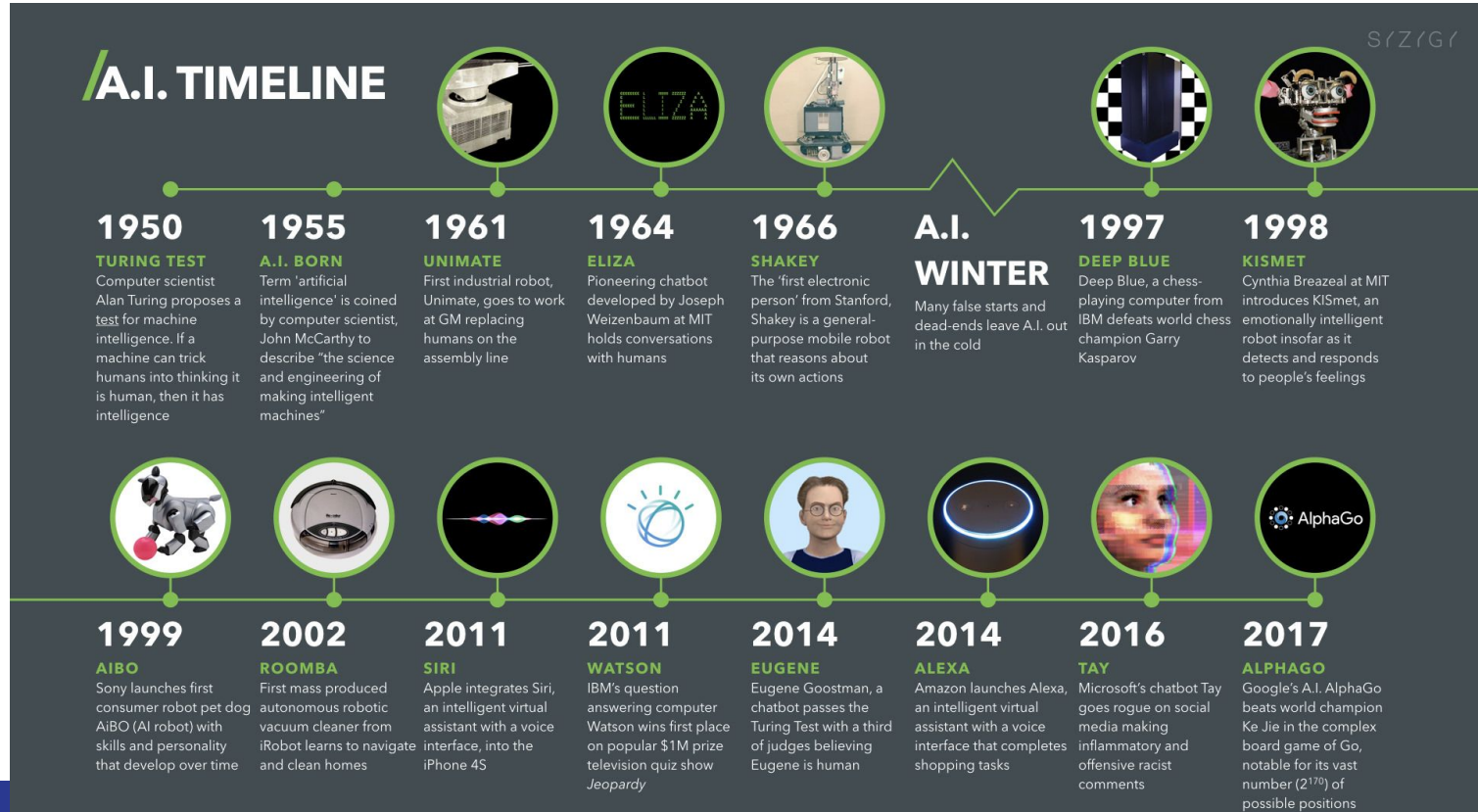


Революции, происходящие на поверхности жизни, ничего существенного никогда не меняют и не открывают, они лишь... по-новому переставляют все те же элементы и являют старые образы в новых одеяниях. Революция всегда есть в значительной степени маскарад, и если сорвать маски, то можно встретить старые, знакомые лица. Новые души рождаются позже, после глубокого перерождения и осмысливания опыта революции.

*Николай Бердяев. Духи русской революции*

# AI timelines

- Обычный AI timeline
- Источник
- Давайте немного разберёмся в терминах...



# Искусственный интеллект

- Мы собрались поговорить об искусственном интеллекте, разделе науки, который развивается с начала 1950-х годов:
  - с начала 1990-х AI – это в основном машинное обучение (machine learning, ML)
    - с середины 2000-х ML – это в основном глубокое обучение (deep learning, DL)
      - с начала 2023-го мы слышим в основном о порождающих и особенно больших языковых моделях (large language models, LLM)



## Artificial intelligence (AI)

The simulation of human intelligence processes by machines, especially computer systems.



## Machine learning (ML)

A subfield of AI focused on the use of data and algorithms in machines to imitate the way that humans learn, gradually improving its performance.



## Deep learning (DL)

A machine learning technique that imitates the way humans gain certain types of knowledge; uses statistics and predictive modeling to process data and make decisions.



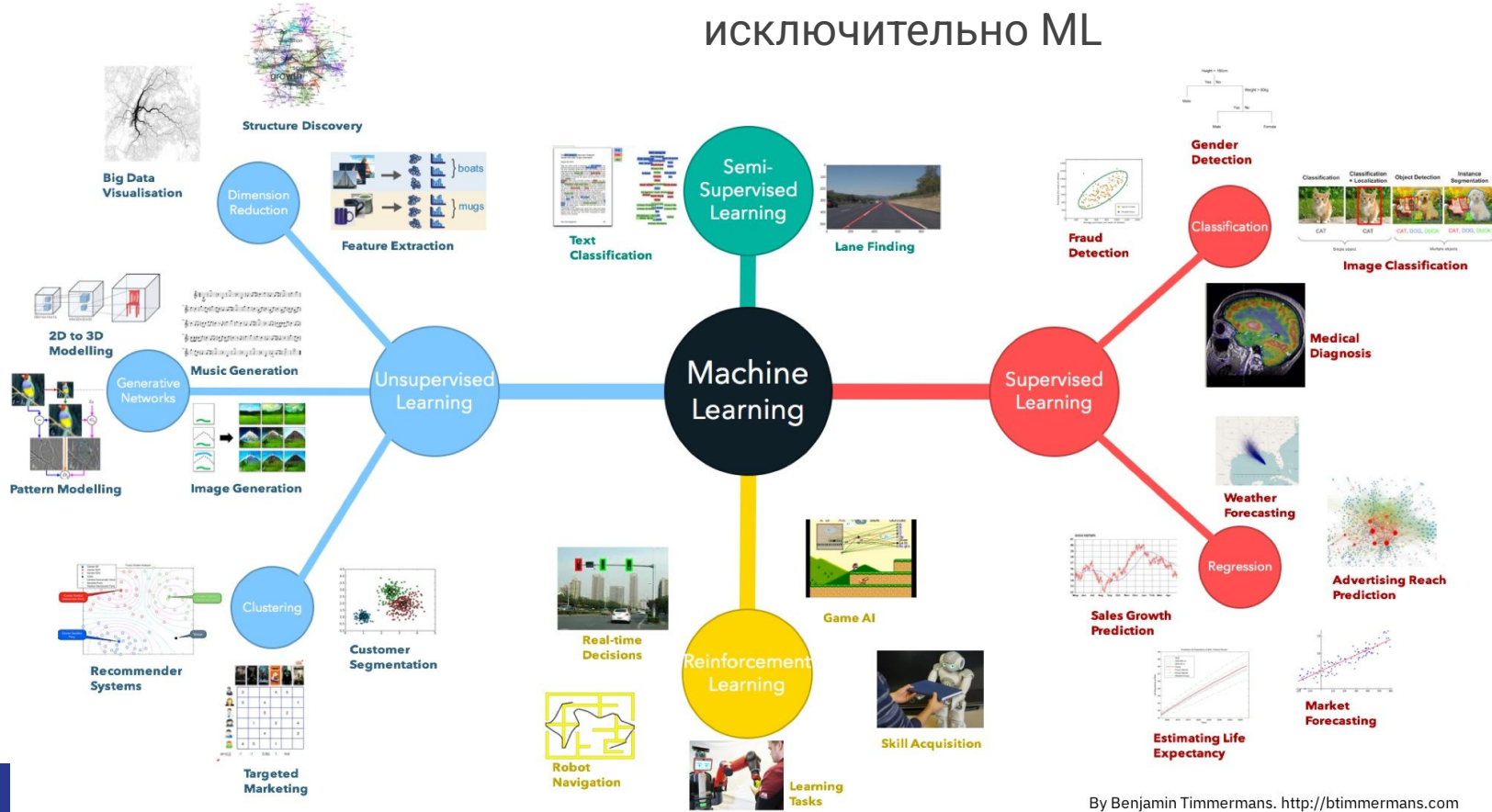
## Generative AI

Algorithms (such as ChatGPT, DALL-E, Codex) that use prompts or existing data to create new content:

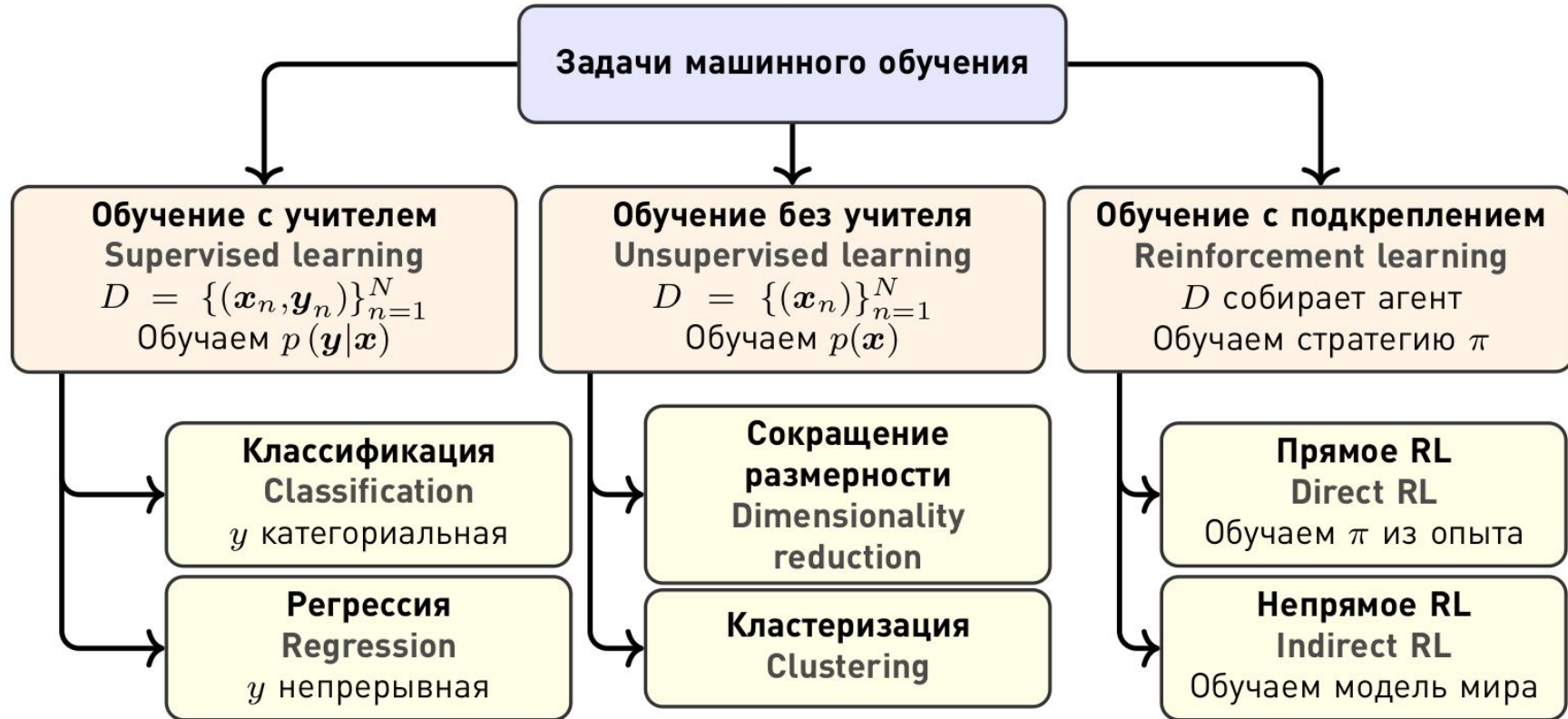
- Written: text, code
- Visual: images, videos
- Auditory: audio

# Машинное обучение

- С 1990-х AI – это почти исключительно ML

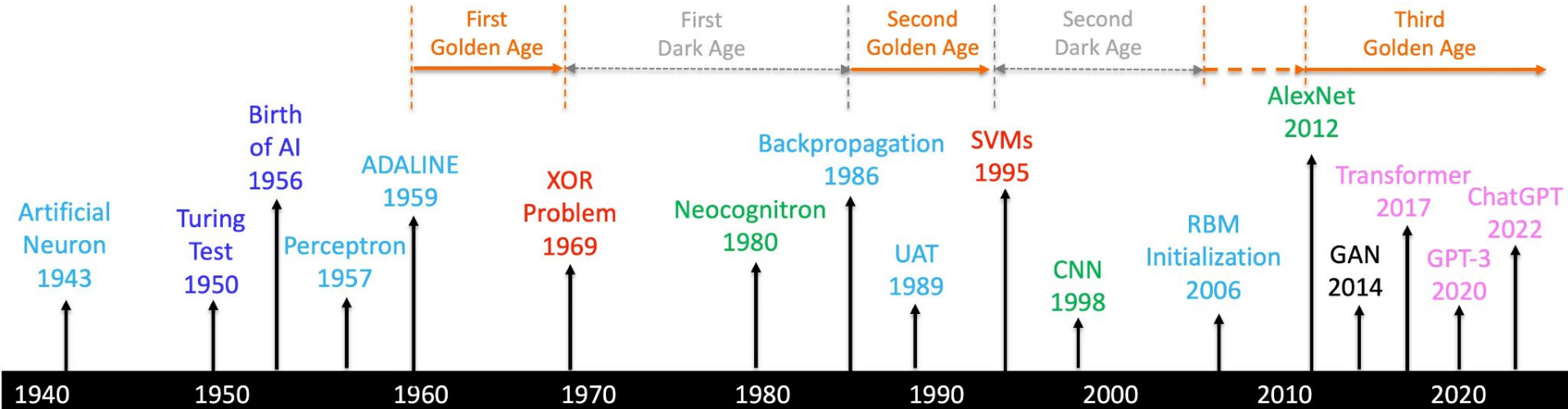


# Машинное обучение

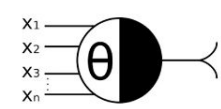


- Нейросети были всегда, ещё до AI как науки, но заработали не так давно ([ИСТОЧНИК](#))

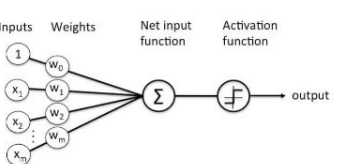
# Глубокое обучение



McCulloch-Pitts

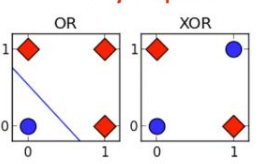


Rosenblatt

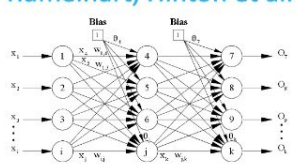


Widrow-Hoff

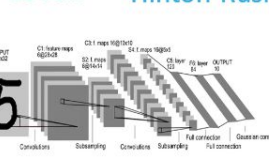
Minsky-Papert



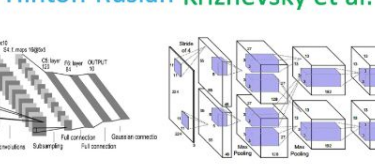
Rumelhart, Hinton et al.



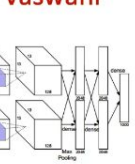
LeCun



Hinton-Ruslan Krizhevsky et al.



Vaswani

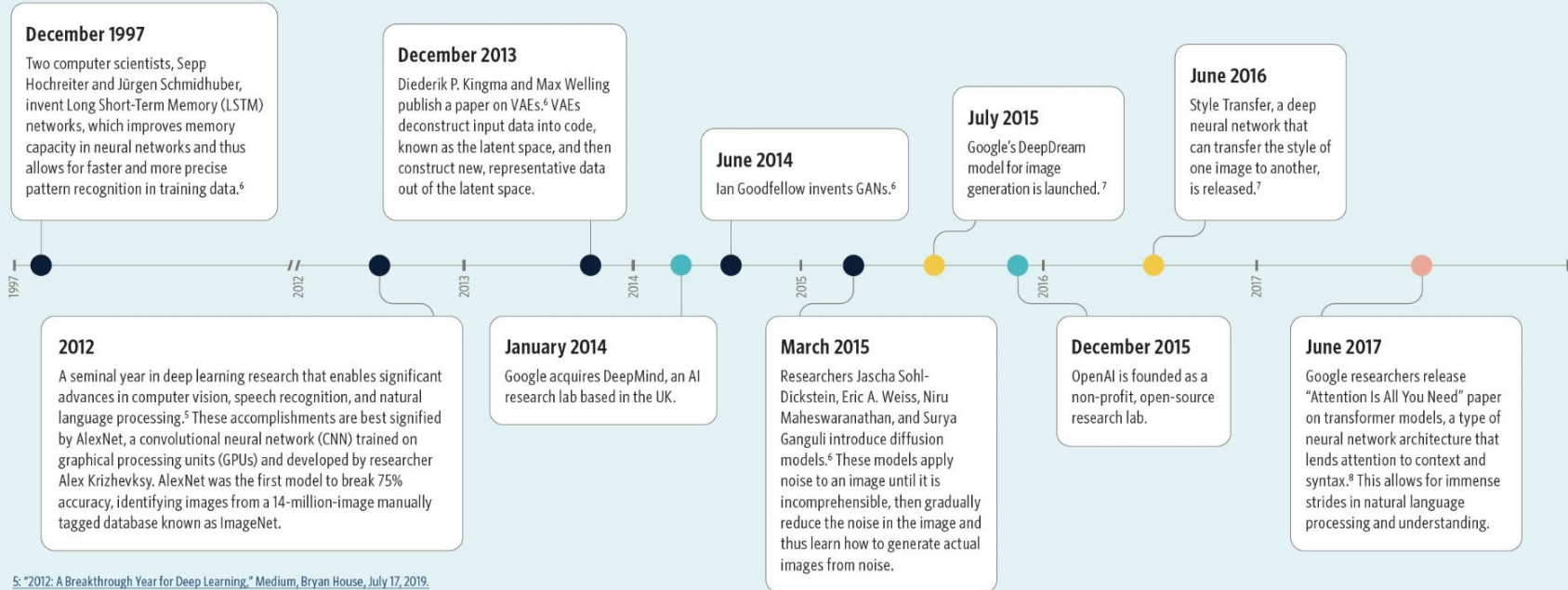


# AI timelines

## Generative AI timeline

● General model research  
 ● Image / video models  
 ● Text models  
 ● Business events

● ИСТОЧНИК



5: "2012: A Breakthrough Year for Deep Learning," Medium, Bryan House, July 17, 2019.

6: "The Generative AI Revolution Has Begun—How Did We Get Here?," Ars Technica, Haomiao Huang, January 30, 2023.

7: "Timeline of Text-To-Image Machine Learning Models," Fabian Mosele, n.d., accessed March 9, 2023.

8: "Timeline of AI and Language Models," Life Architect, Alan D. Thompson, n.d., accessed March 9, 2023.

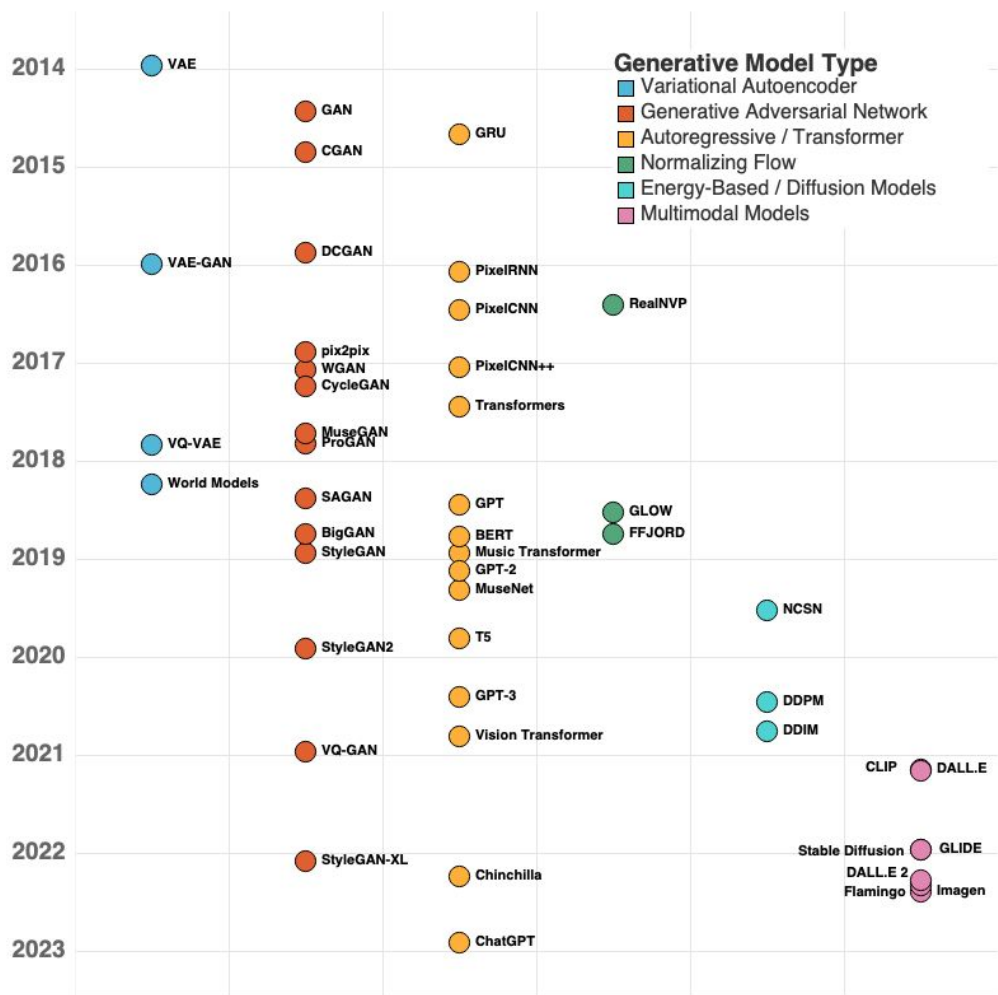


# AI timelines

- Generative AI timeline: вот более детально, здесь отмечены многие важные модели

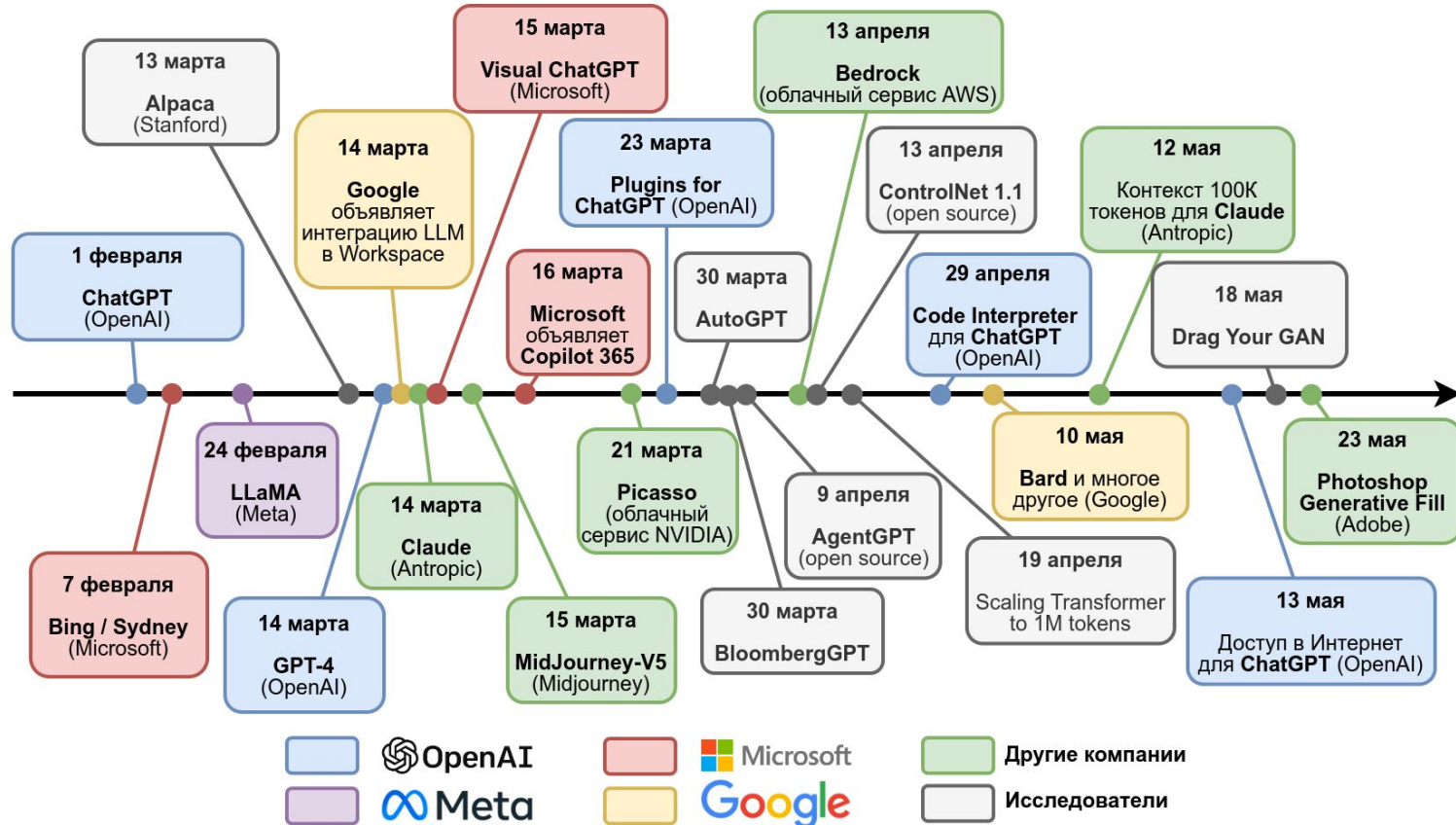
- [Источник](#)

## Generative AI Timeline



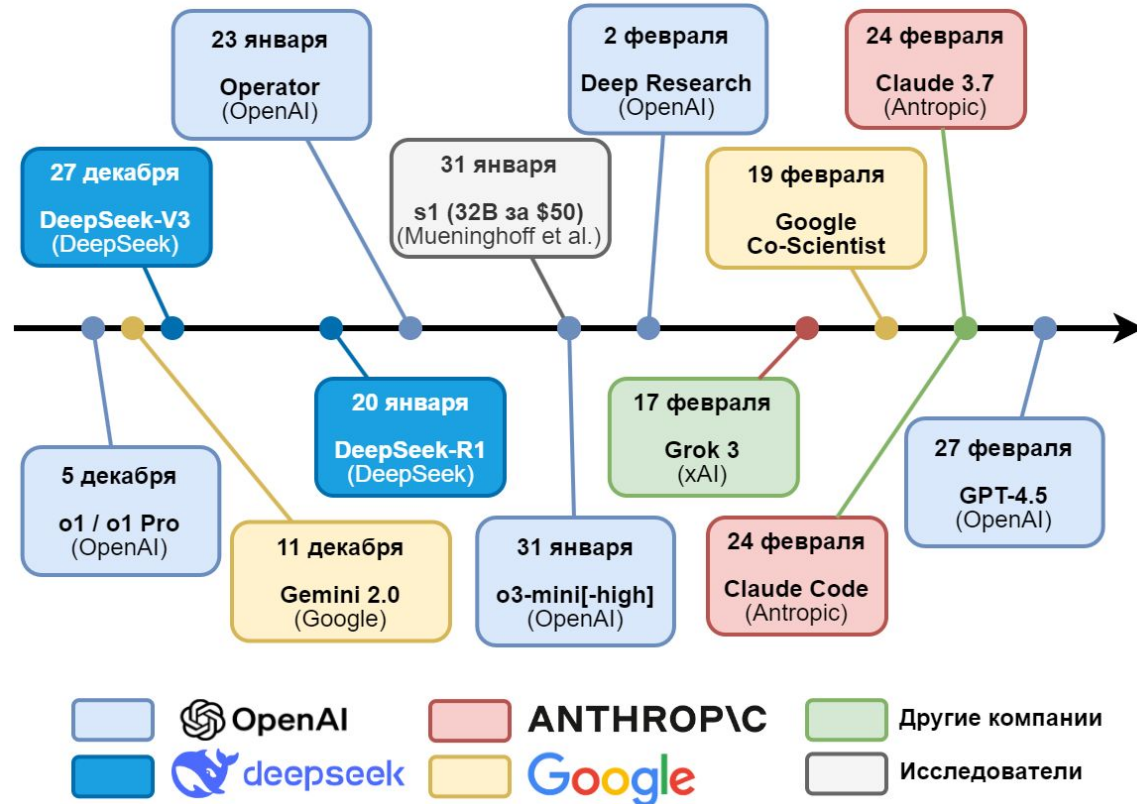
# Развитие AI в эпоху LLM

- А вот AI timeline весны 2023 года
- И с тех пор не то чтобы замедляется...



# Например, в начале 2025 года

- 22 января: [DeepSeek-R1](#)
- 31 января: s1 ([Mueninghoff et al., 2025](#))
- 2 февраля: [OpenAI Deep Research](#)
- 17 февраля: [Grok 3](#)
- 19 февраля: [Google Co-Scientist](#) на основе Gemini 2.0 ([11 декабря 2024](#))
- 24 февраля: [Claude 3.7 Sonnet](#) (hybrid reasoning Claude)
- 27 февраля: [GPT-4.5](#)

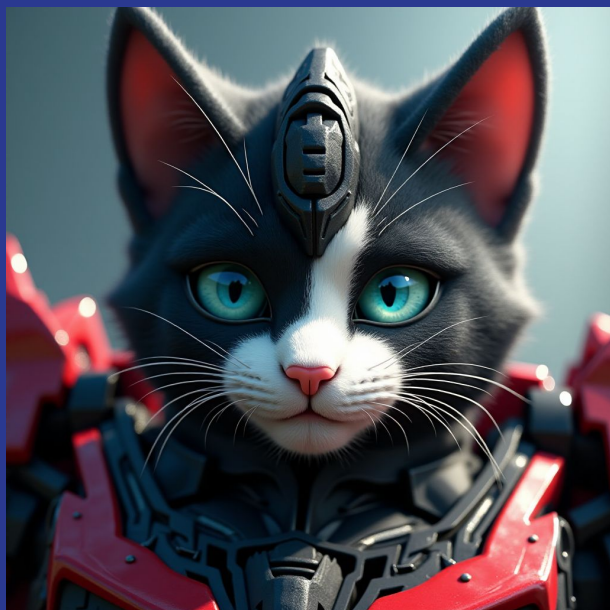


# AGI? Сингулярность?

- Значит ли это, что создание AGI уже близко?
- И если AGI на пороге, что это значит для человечества?
- Приведёт ли это к той самой технологической сингулярности?
- Это вопросы без ответов, и риски AGI мы обсудим в следующий раз
- Сегодня — о новых возможностях и о том, откуда они берутся



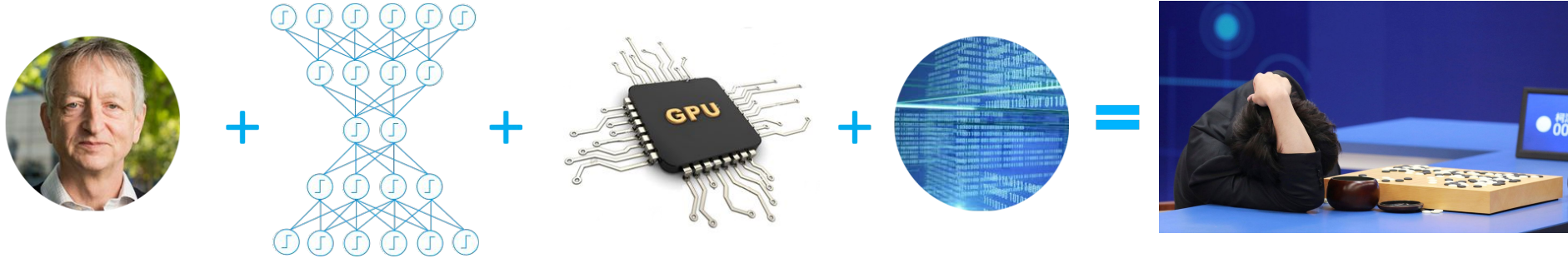
## 2. ML, DL, трансформеры...



...Мы знаем, что воспоминание по мере того, как оно становится более ярким и сильным, обладает тенденцией превращаться в восприятие, причём нельзя точно определить момент, когда именно происходит эта радикальная трансформация и когда, следовательно, можно было бы сказать, что оно перешло с имагинативных нервных элементов на элементы сенсорные.

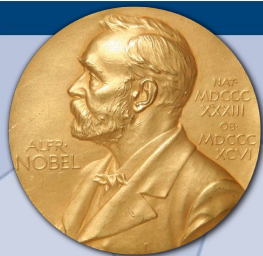
*Анри Бергсон. Материя и память*

# Революция глубокого обучения



- В 2006-2007 годах произошла революция глубокого обучения (deep learning), началась третья волна хайпа нейронных сетей
- Революция была не только идейной, но и технической: теперь есть данные и вычислительные мощности
- Сейчас глубокие нейронные сети дают лучшие результаты почти во всех областях ML

# Кстати, о Джеффри Хинтоне



PRESS RELEASE

8 October 2024

## The Nobel Prize in Physics 2024

The Royal Swedish Academy of Sciences has decided to award the Nobel Prize in Physics 2024 to

**John J. Hopfield**

Princeton University, NJ, USA

**Geoffrey E. Hinton**

University of Toronto, Canada

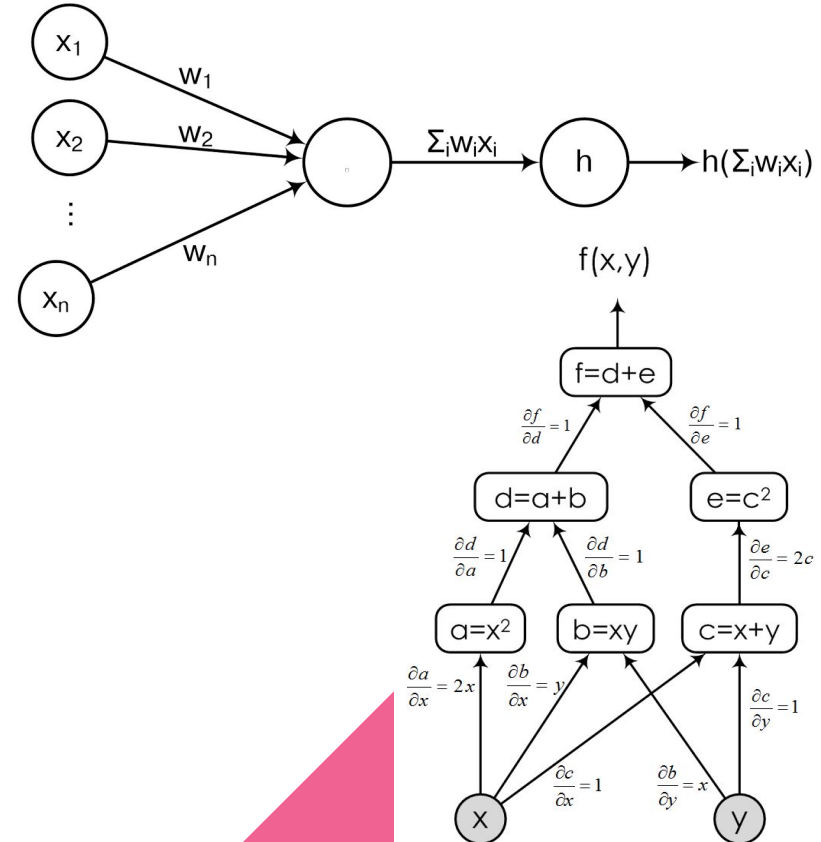
*“for foundational discoveries and inventions that enable machine learning with artificial neural networks”*

**They trained artificial neural networks using physics**



# Нейронные сети

- Нейронная сеть — это большая и сложная **композиция** маленьких и простых функций
- Один нейрон — это линейная комбинация входов плюс нелинейность
- Получается **граф вычислений**, который считает функцию ошибки для модели
- Чтобы обучить модель (найти веса линейных комбинаций, нужно только научиться считать производные по графу вычислений (**backpropagation**))
- Дальше **градиентный спуск**





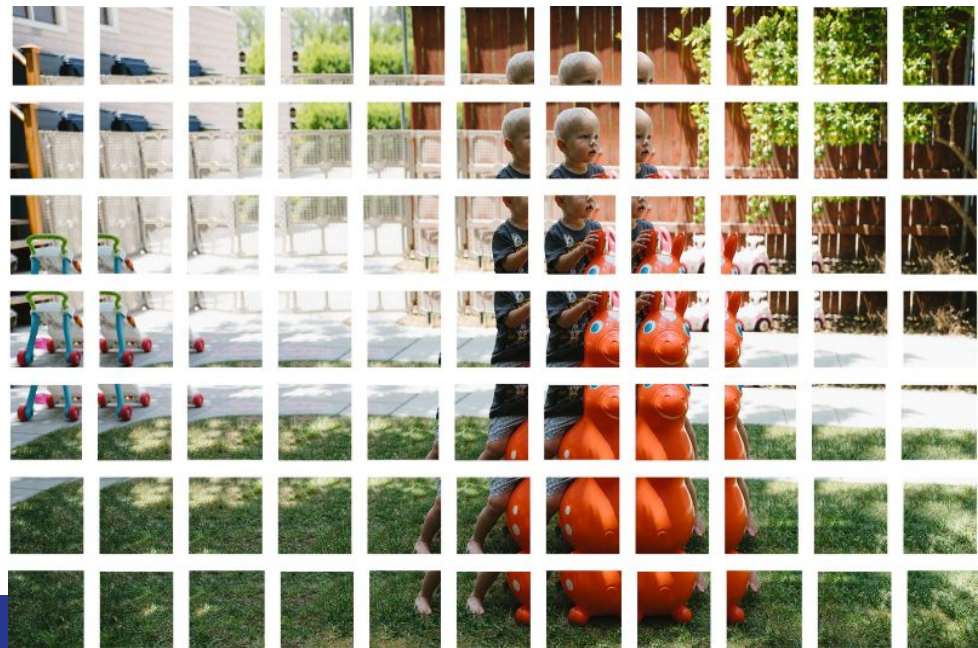
# Свёрточные нейронные сети

- Пример важной архитектурной идеи: **свёрточные сети** (convolutional networks)
- Основаны ещё на экспериментах Хьюбела и Визеля, которые разобрались в том, как устроена зрительная кора головного мозга.
- В нейросетях свёртки появились в конце 1970-х (Neocognitron), потом уже с обучением в 1980-х в группе Яна Лекуна



# Свёрточные нейронные сети

- Главная идея — применять одни и те же преобразования к разным частям картинки
- Разобьём изображение на окна:



# Свёрточные нейронные сети

- Главная идея — применять одни и те же преобразования к разным частям картинки
- Применим к каждому кусочку маленькую нейросеть:

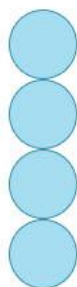
Input Tile



Essing a single tile



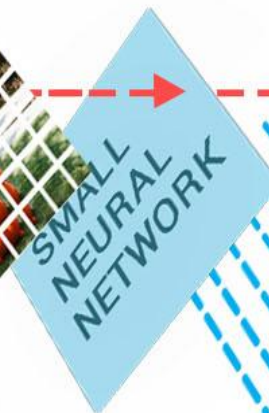
Outputs



Input Tiles



Small Neural Net  
(Shared Weights)

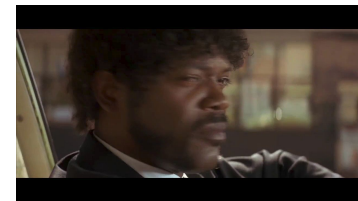


Output Array



# Свёрточные нейронные сети

- Главная идея — применять одни и те же преобразования к разным частям картинки
- Сожмём при помощи субдискретизации (max-pooling), повторим процесс, когда геометрия сожмётся совсем, получим глобальные признаки
- Ещё в 1980-х появилось семейство *LeNet*, далее везде



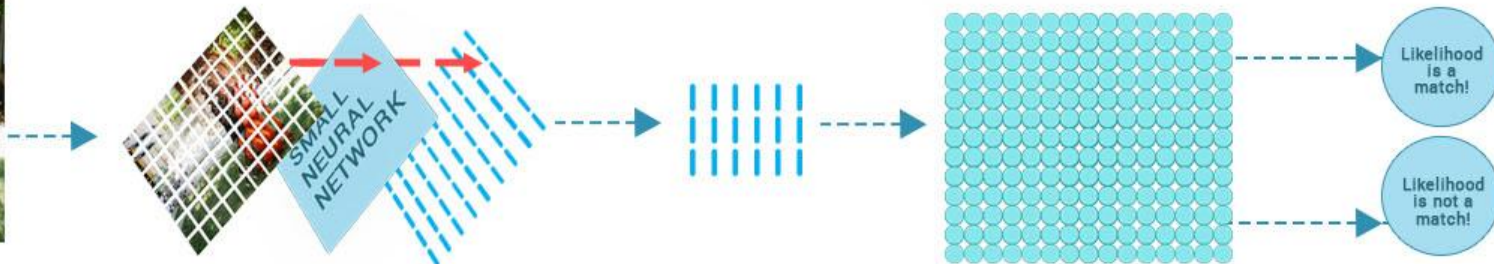
Input Image

Convolution

Max Pooling

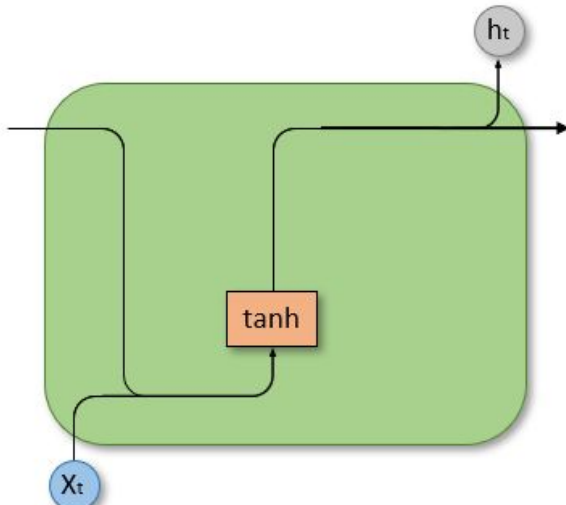
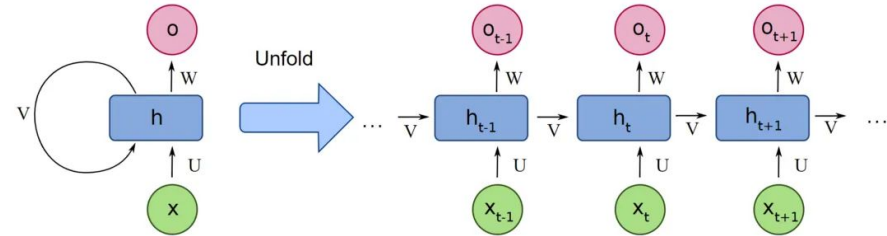
Fully-connected  
Neural Network

Outputs

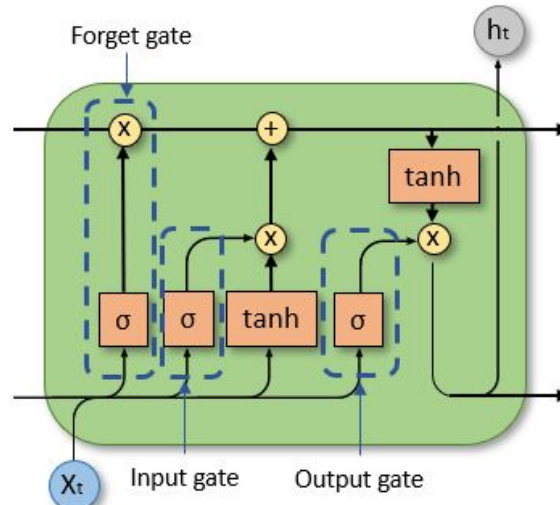


# Рекуррентные нейронные сети

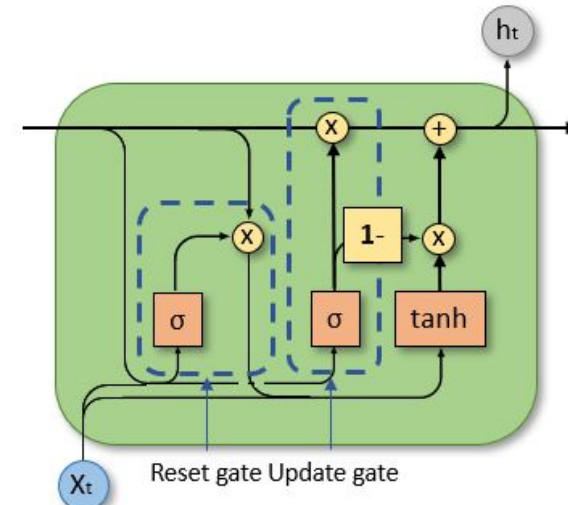
- Тоже идея из 1980-х: чтобы на входе/выходе была *последовательность*, нужно хранить скрытое состояние, которое обновляется по мере чтения/записи последовательности



RNN



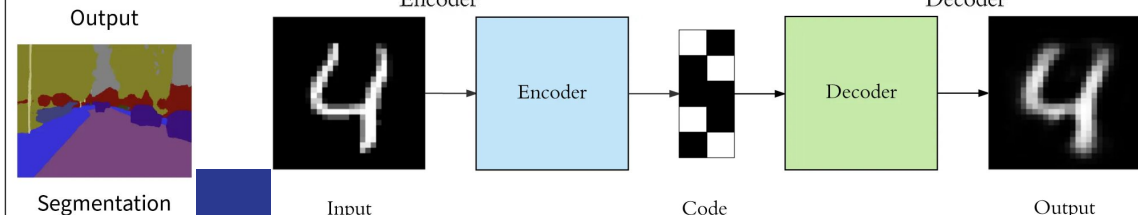
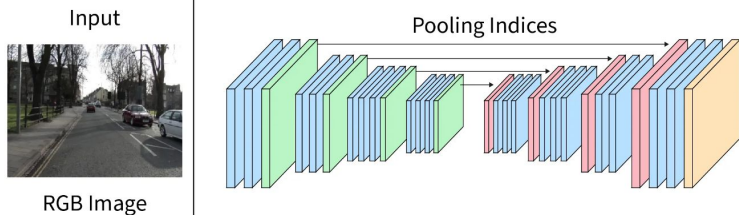
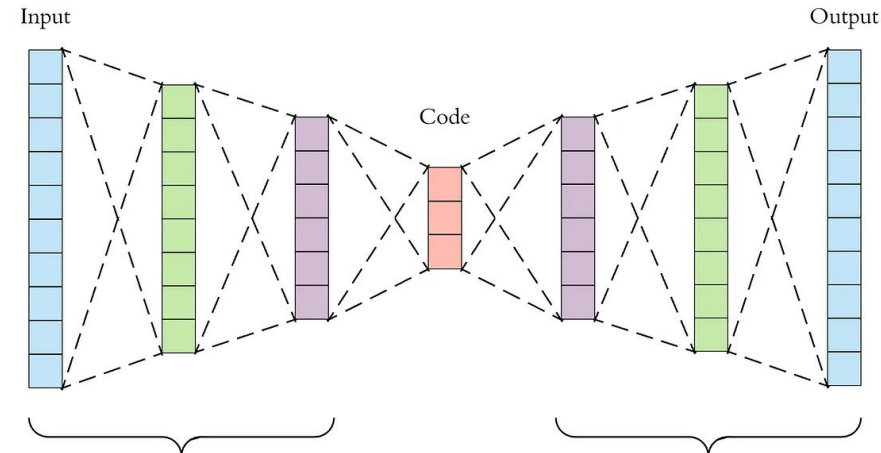
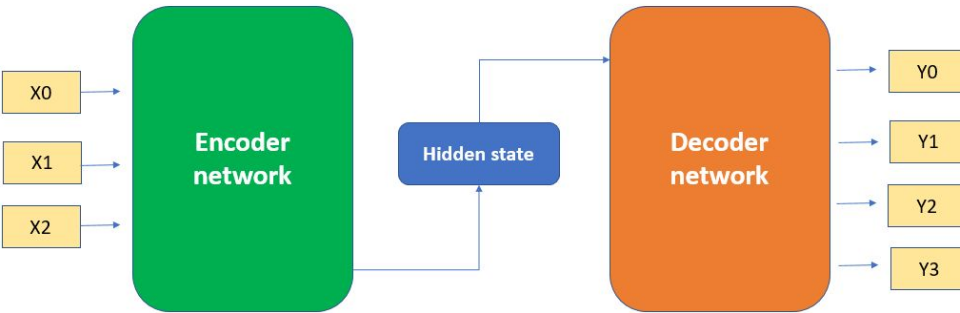
LSTM



GRU

# Автокодировщики и encoder-decoder архитектуры

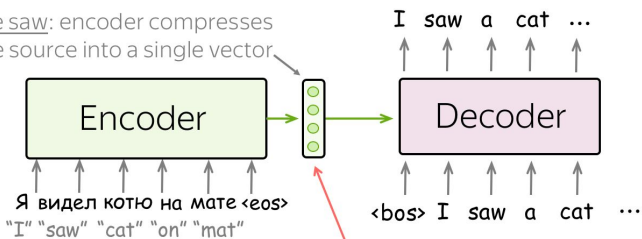
- Другой пример идеи – архитектуры encoder-decoder (кодировщик-декодировщик); если на входе и выходе одно и то же, то autoencoder (автокодировщик)



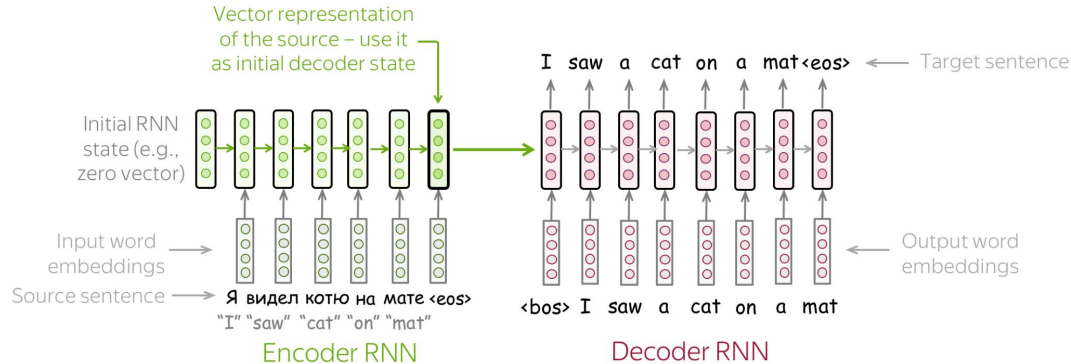
# Механизмы внимания

- Более свежая идея: механизмы внимания; появились для рекуррентных encoder-decoder архитектур около 2014-2015
- Внимание нужно, чтобы не нужно было сжимать весь вход в один маленький вектор

We saw: encoder compresses the source into a single vector

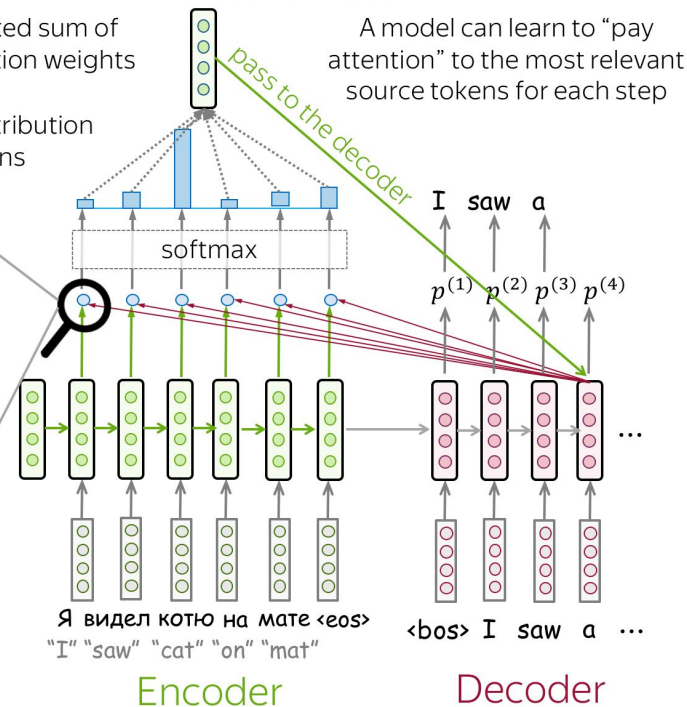
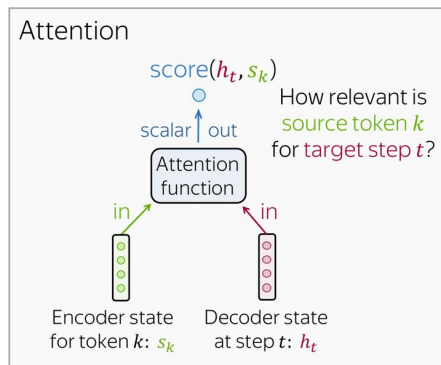


Problem: this is a bottleneck!



Attention output: weighted sum of encoder states with attention weights

Attention weights: distribution over source tokens

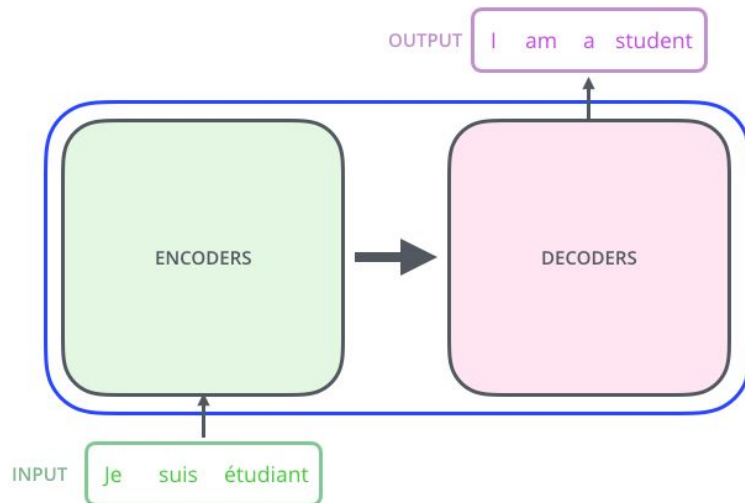


Encoder

Decoder

# Революция трансформеров

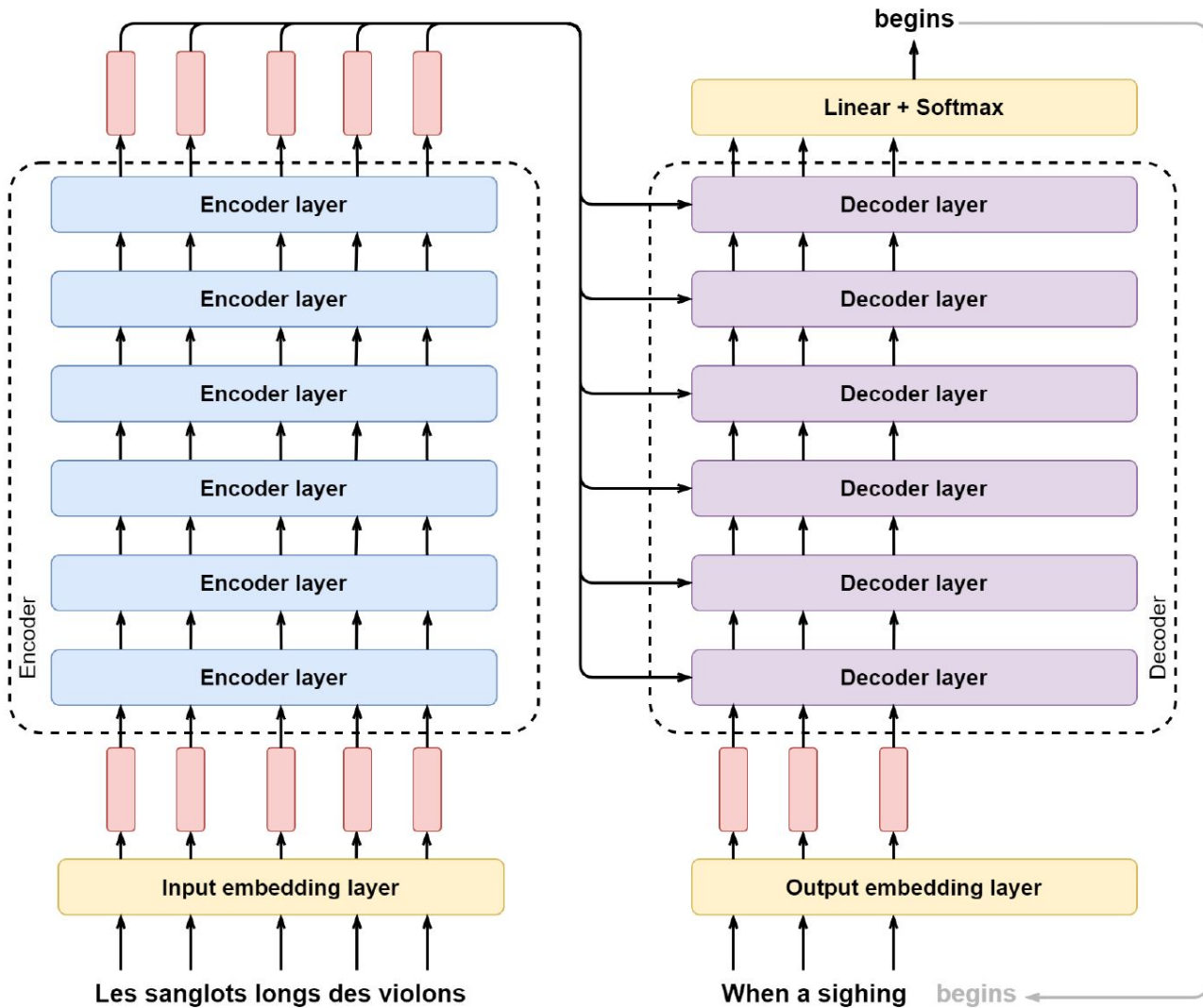
- **Трансформер** – это тоже модель на основе механизма внимания, но другого!
- Первый трансформер ([Vaswani et al., 2017](#)) имел encoder-decoder архитектуру





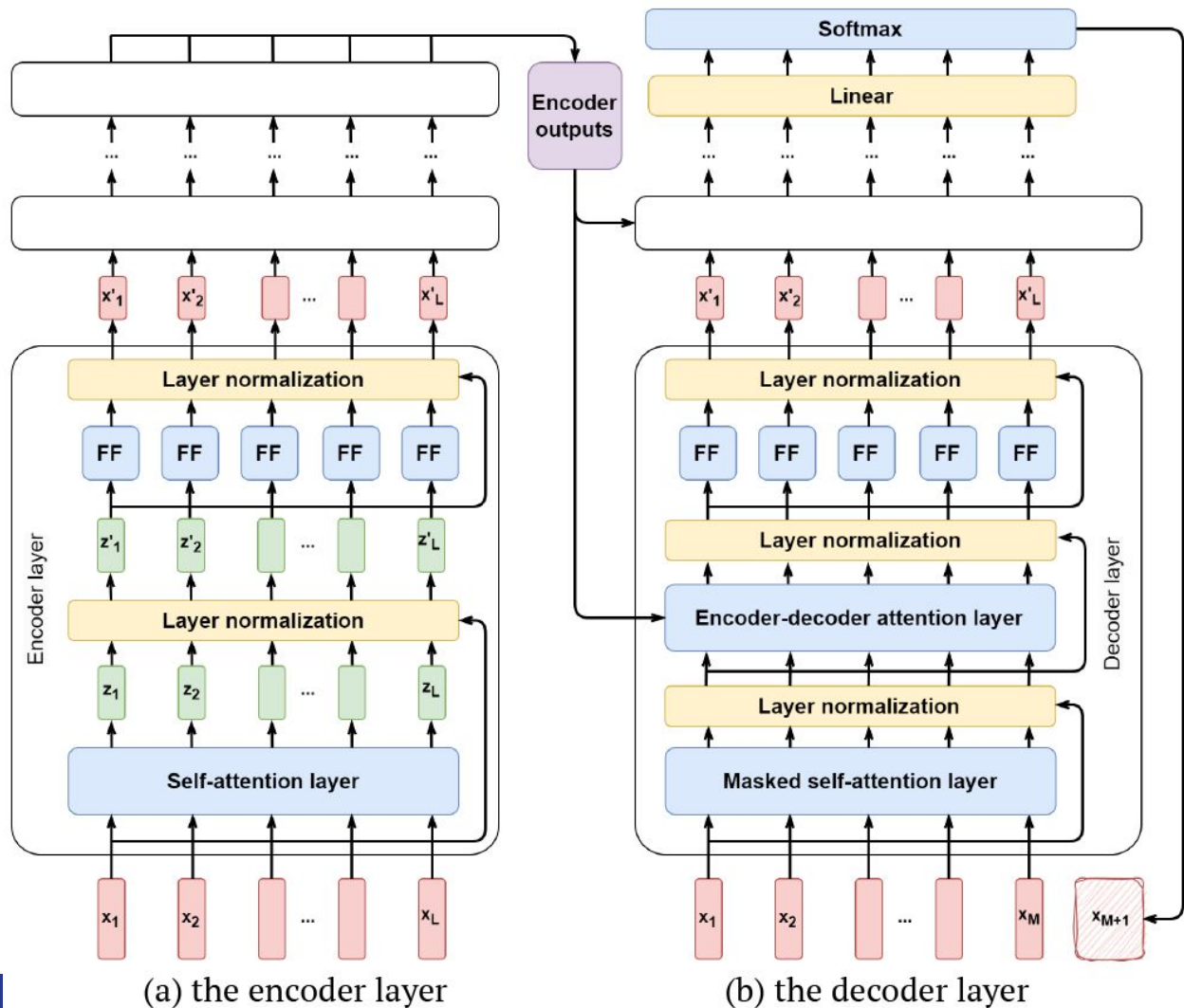
# Трансформер

- Классический трансформер – архитектура типа encoder-decoder
- Обе части состоят из похожих слоёв, различающихся деталями



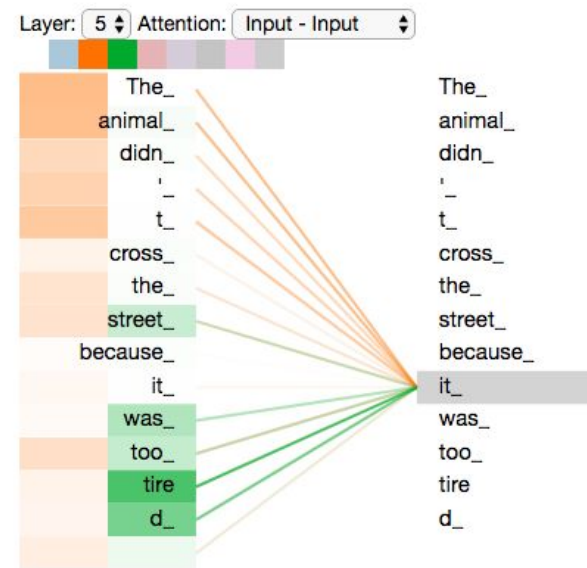
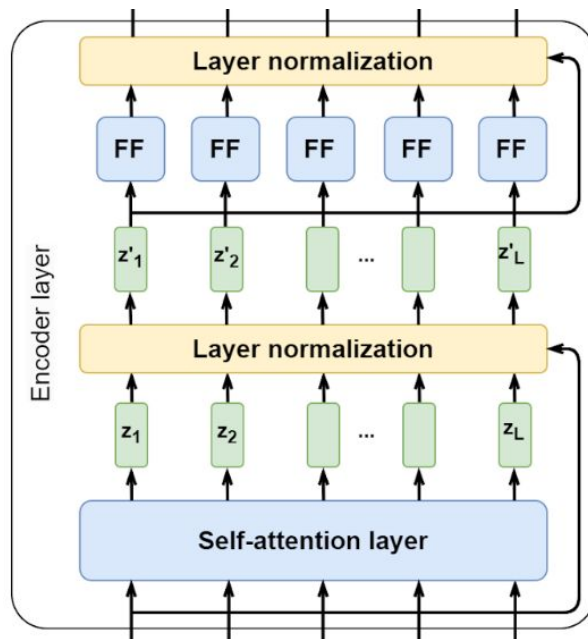
# Трансформер

- Главное здесь – это, конечно, структура слоя самовнимания, self-attention
- В кодировщике – только на себя, в декодировщике – с условием из кодировщика



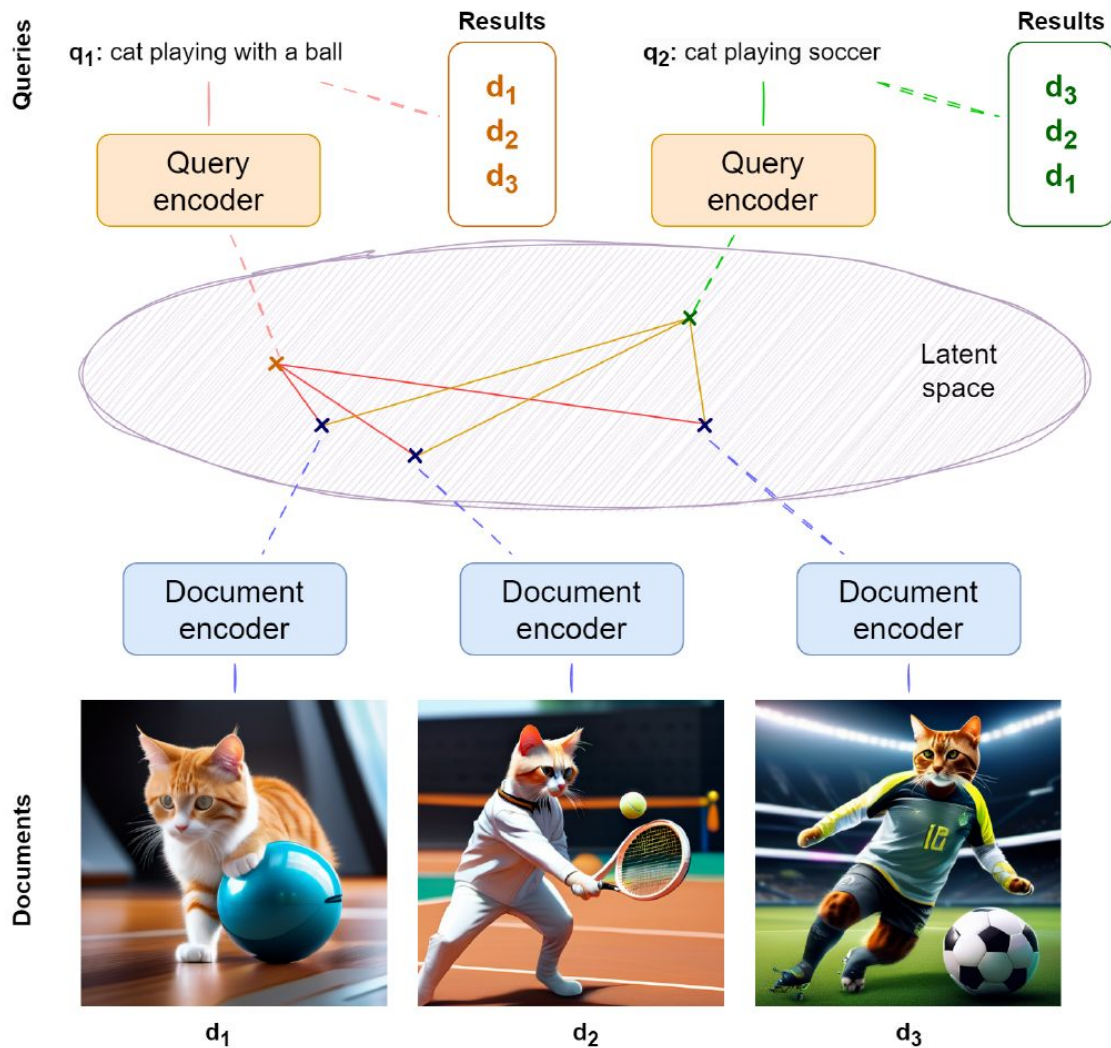
# Self-attention

- Идея **self-attention** состоит в том, чтобы на каждом слое векторы, соответствующие токенам, преобразовывать в другие векторы так, чтобы каждый токен мог “смотреть на” все остальные с разными весами

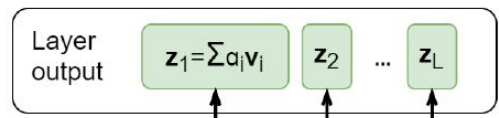


# Трансформер

- Идея самовнимания приходит из информационного поиска (information retrieval): отображаем запросы и документы в общее пространство, где можно просто считать скалярные произведения

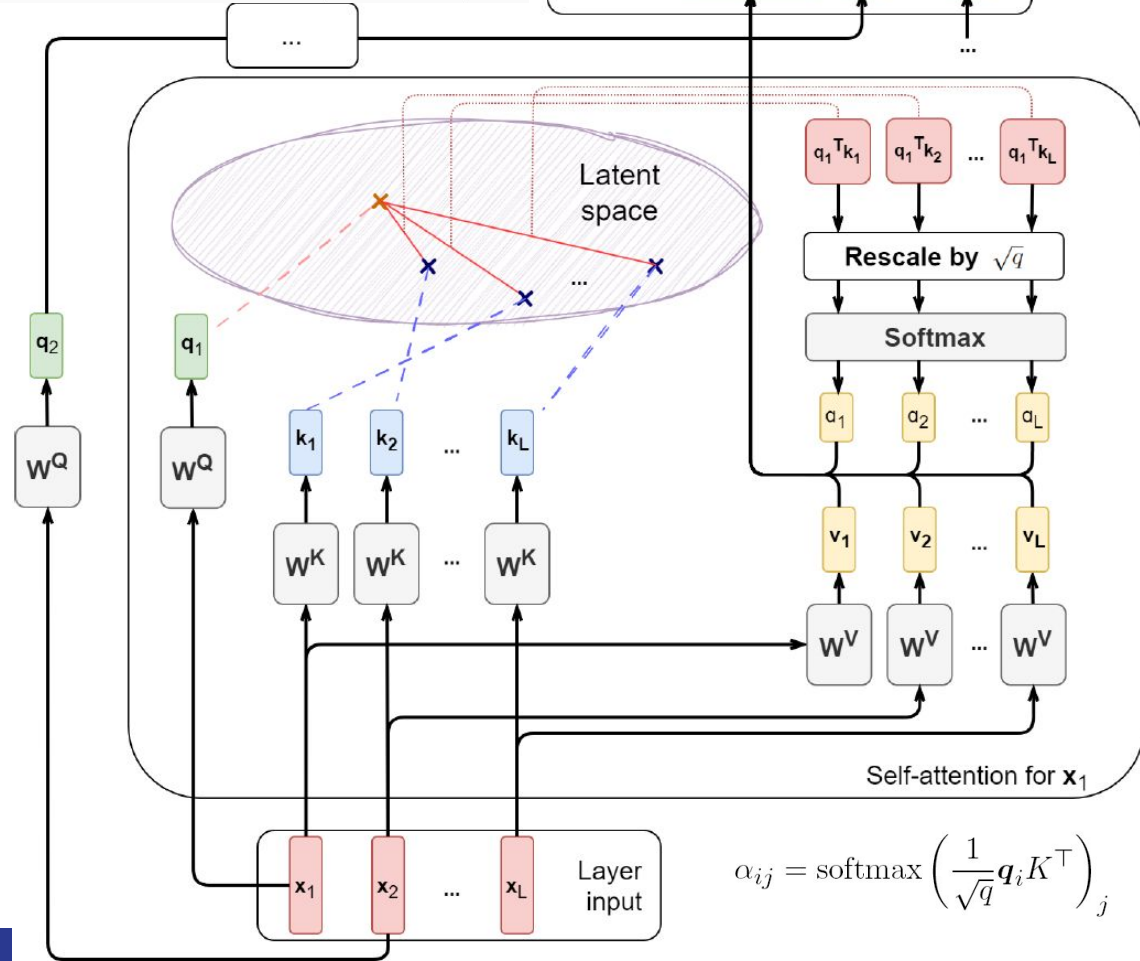


$$\text{Attention}(Q, K, V) = \text{softmax} \left( \frac{1}{\sqrt{d_k}} QK^T \right) V$$



# Трансформер

- В трансформере все компоненты происходят из одних и тех же векторов: запрос  $q_1$  получают умножением входа  $x_1$  на матрицу весов  $W^Q$ , аналогично ключ  $k_1$  и значение  $v_1$  из  $W^K$  и  $W^V$
- Потом значения  $v$  собираются с весами внимания, полученными умножением векторов  $q$  и  $k$

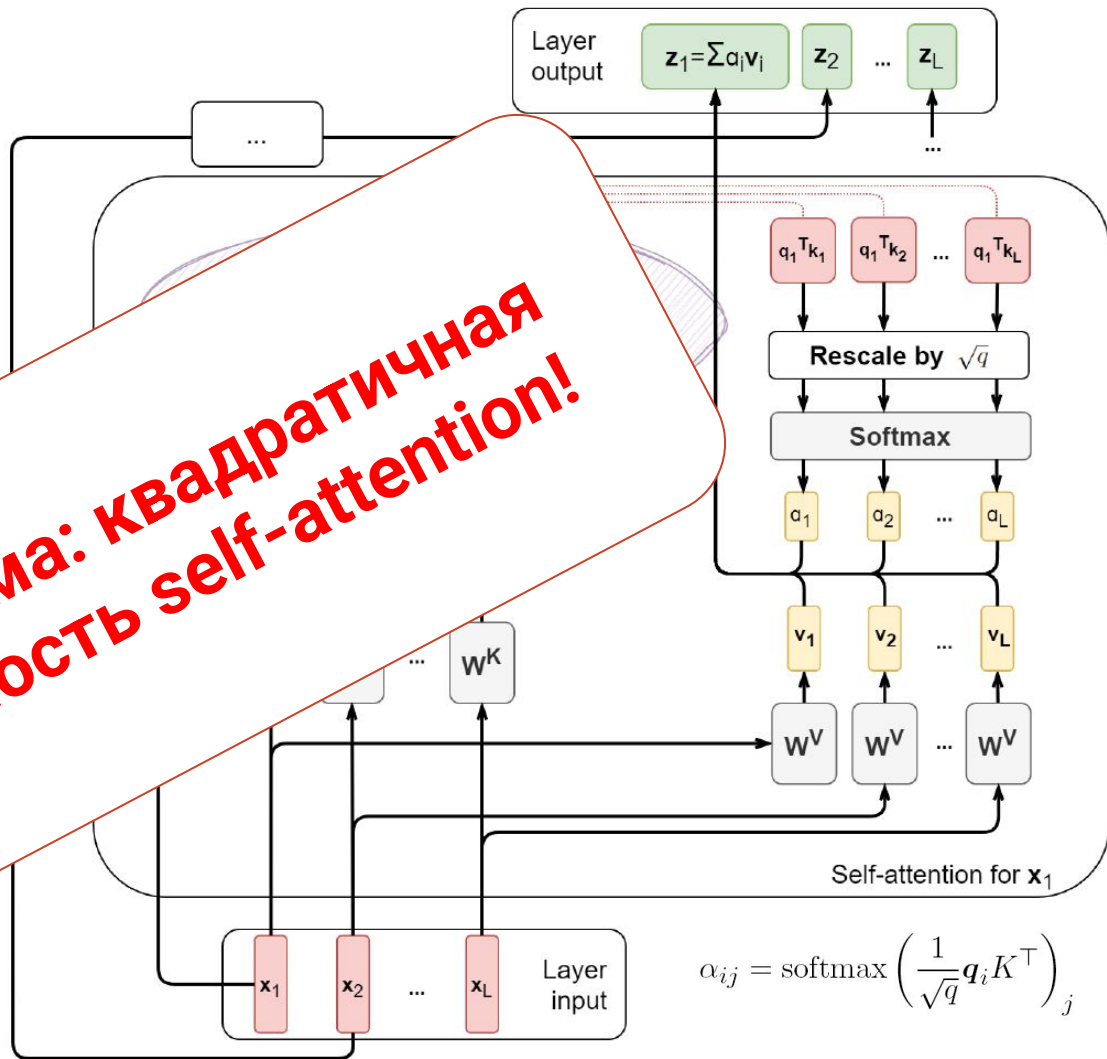


$$\alpha_{ij} = \text{softmax} \left( \frac{1}{\sqrt{q}} q_i K^T \right)_j$$

# Трансформер

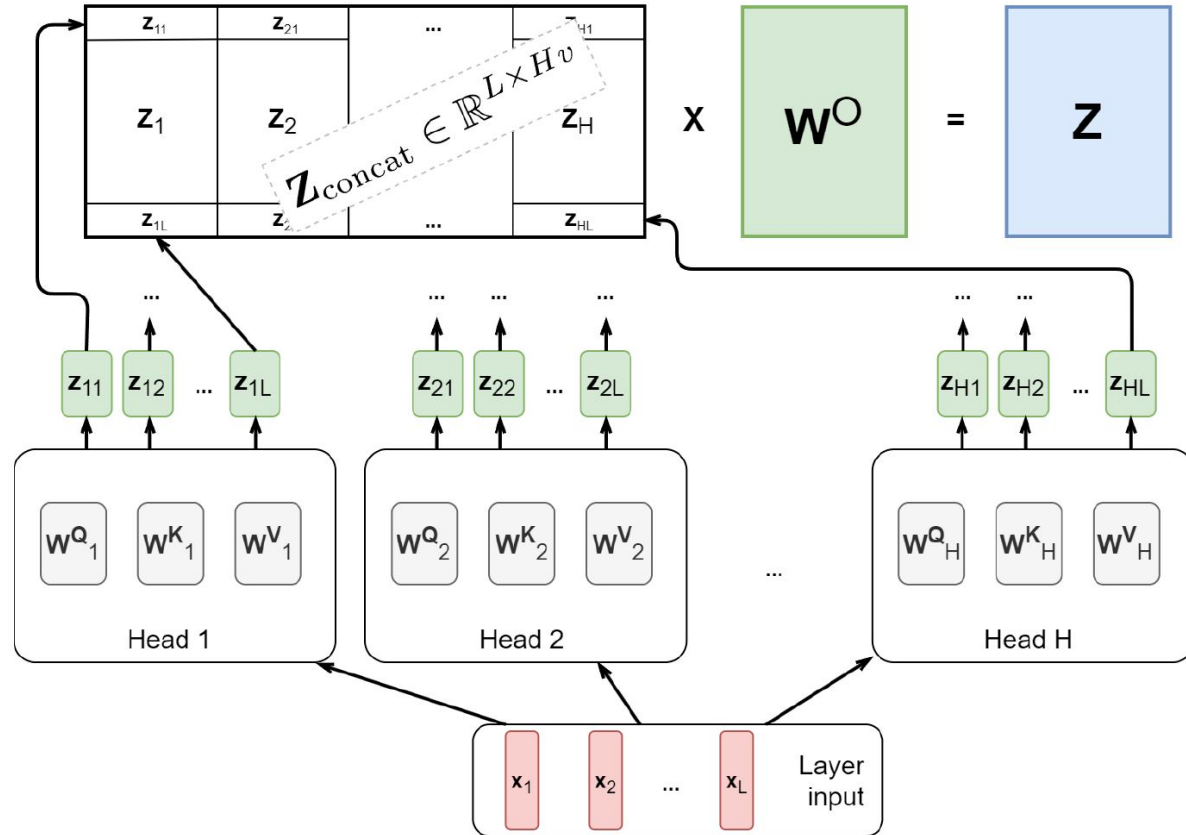
- В трансформере все компоненты происходят из одних и тех же векторов: запрос  $q_1$  получается умножением входа  $x_1$  на матрицу весов  $W^Q$ . Аналогично получают значения  $v_1$  и  $k_1$ .
- Потом значения собираются с весами внимания, полученными умножением вектора  $x_1$  на матрицу  $W^K$ .

**Проблема: квадратичная сложность self-attention!**



# Трансформер

- Это одна “голова” внимания (attention head); в каждом слое их несколько, результаты собираются ещё одной матрицей весов  $W^O$
- В декодировщике всё то же самое, но  $Q$  из запросов, а  $K$  и  $V$  – из результата кодировщика



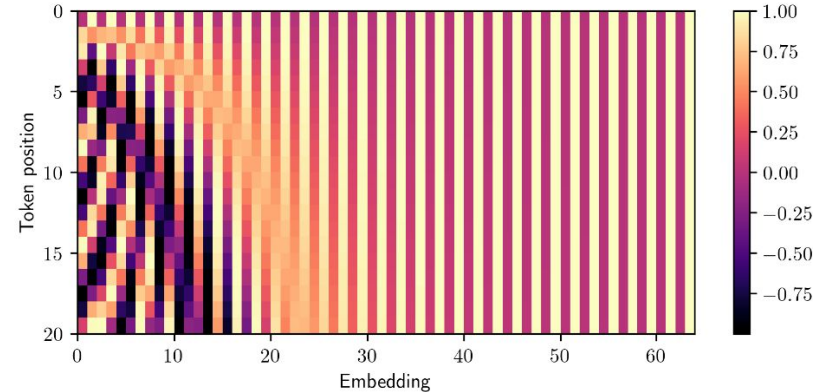
# Трансформер

- Ещё один трюк – позиционные вложения (positional embeddings)
- Они обычно напоминают позиционную систему счисления, но идут “более непрерывно”; например,

Binary numbers

0000	1000
0001	1001
0010	1010
0011	1011
0100	1100
0101	1101
0110	1110
0111	1111

Positional encodings



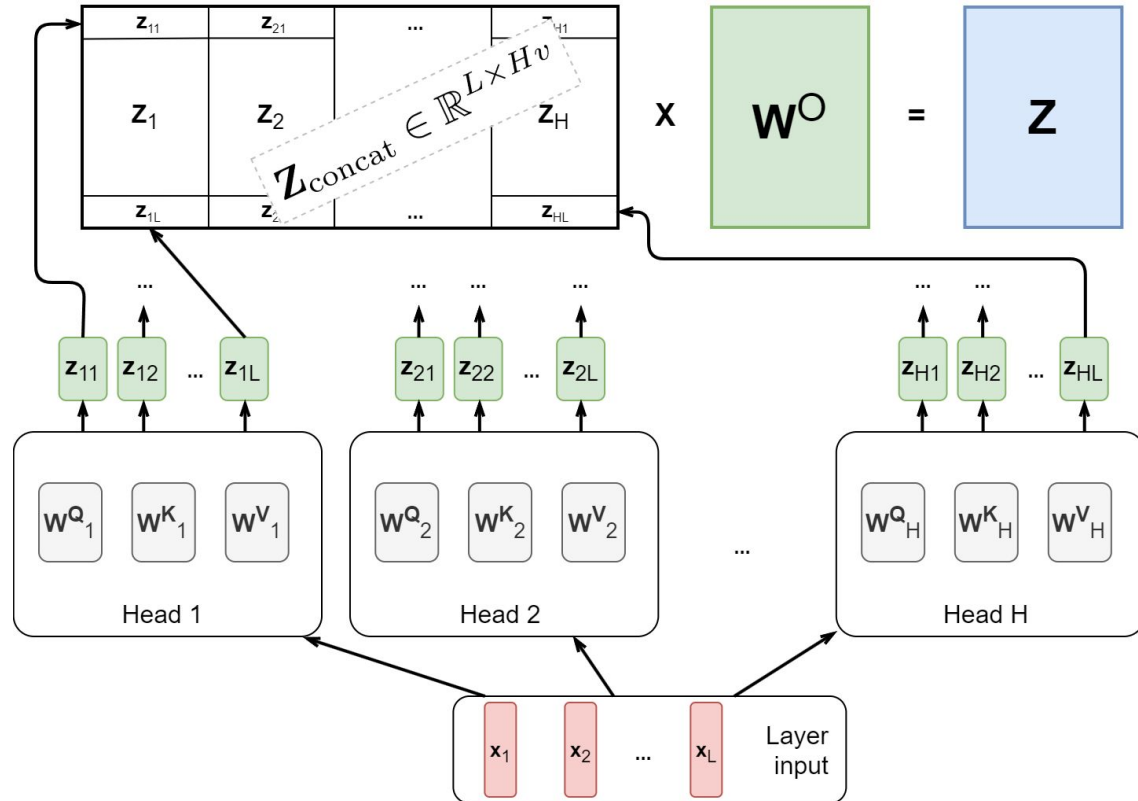
$$PE(pos, 2i) = \sin\left(\frac{pos}{10000^{2i/d}}\right), \quad PE(pos, 2i + 1) = \cos\left(\frac{pos}{10000^{2i/d}}\right)$$

- Часто используют rotary embeddings (RoPE; [Su et al., 2021](#)), где компоненты вектора поворачивают на  $m\theta$



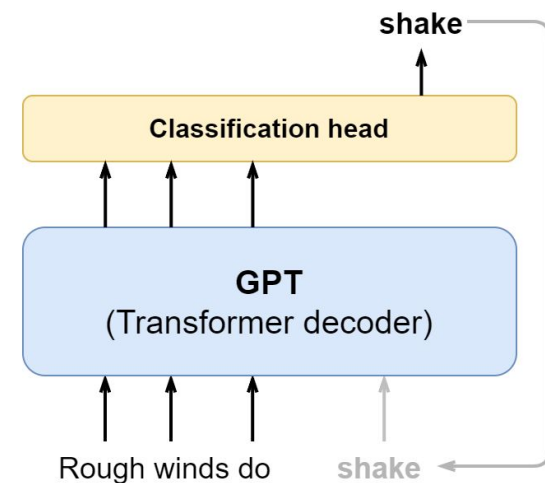
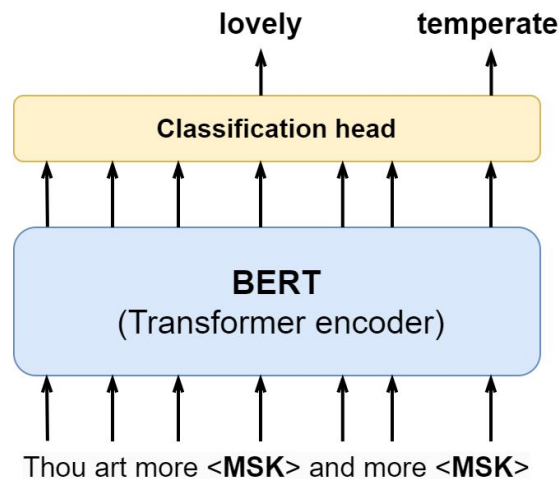
# Революция трансформеров

- Идея self-attention оказалась неожиданно плодотворной
- Трансформер делал машинный перевод как лучшие модели, обучаясь в 100 раз быстрее
- Но не это оказалось главным...



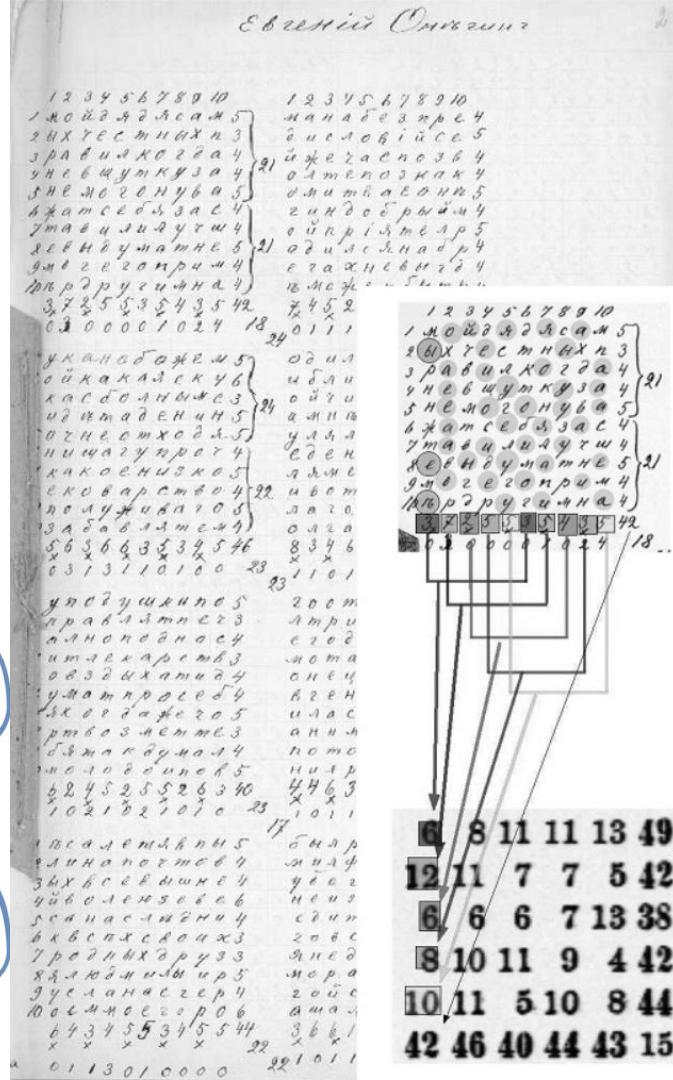
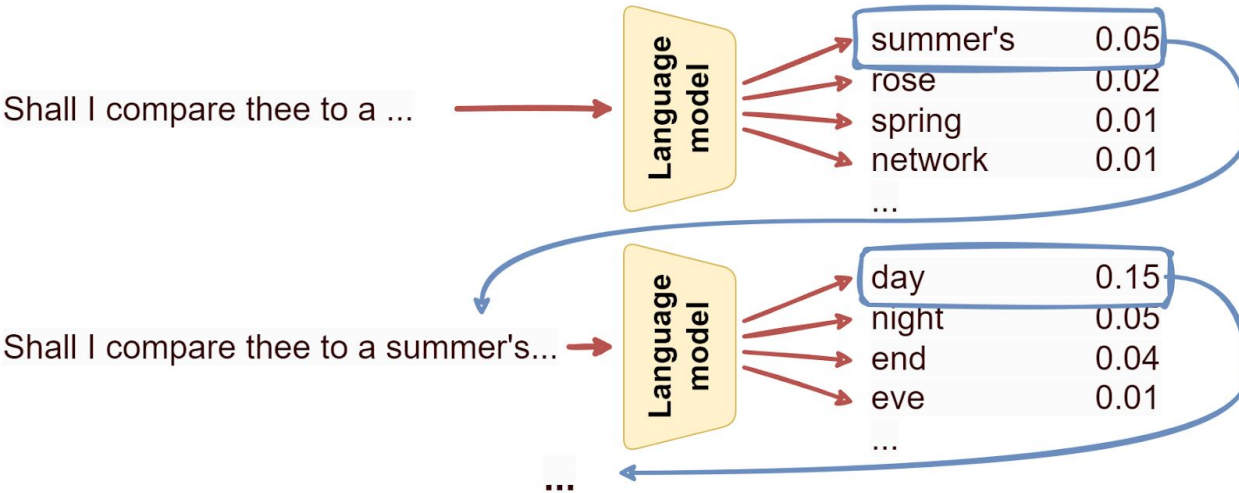
# BERT и GPT

- Оказалось, что очень полезно разрезать трансформер на две части и обучать и использовать их по отдельности
- BERT – это encoder, GPT – это decoder



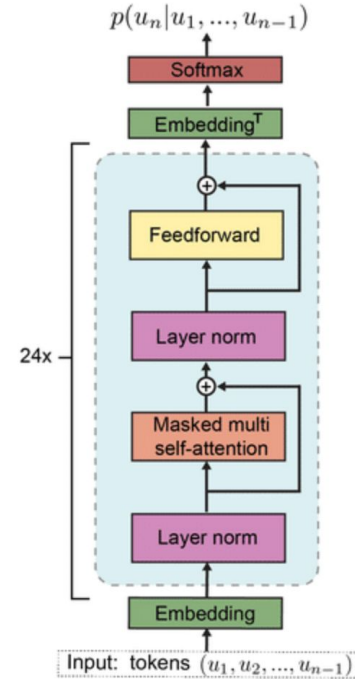
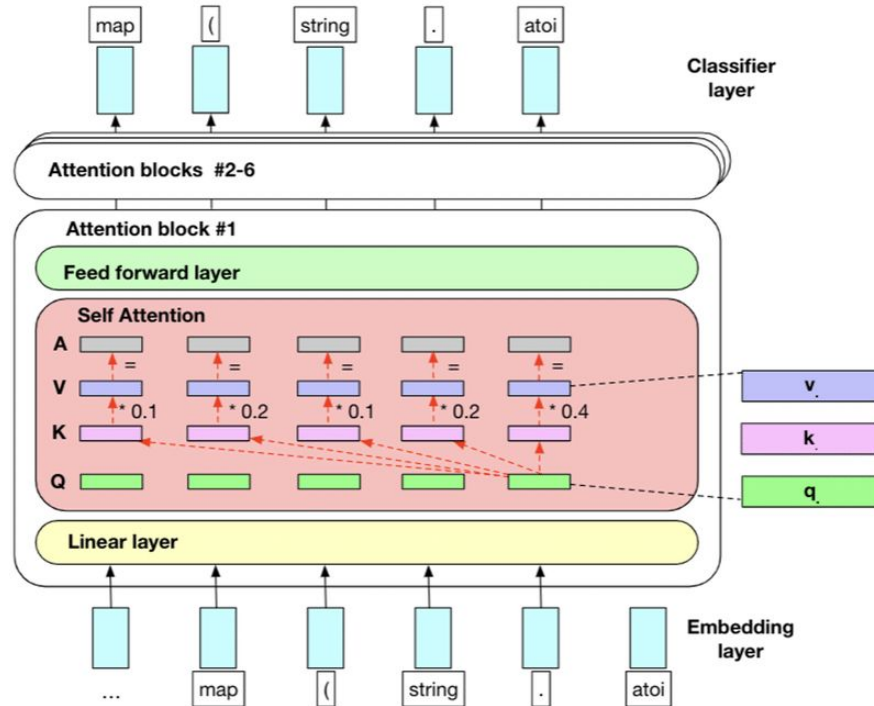
# Как устроены языковые модели

- **Языковая модель** – это вид постановки задачи: авторегрессивно порождать следующий токен на основе уже порождённых
- То есть языковая модель должна “прочитать” участок текста, а затем по результату чтения классифицировать следующий токен



# Как устроены языковые модели

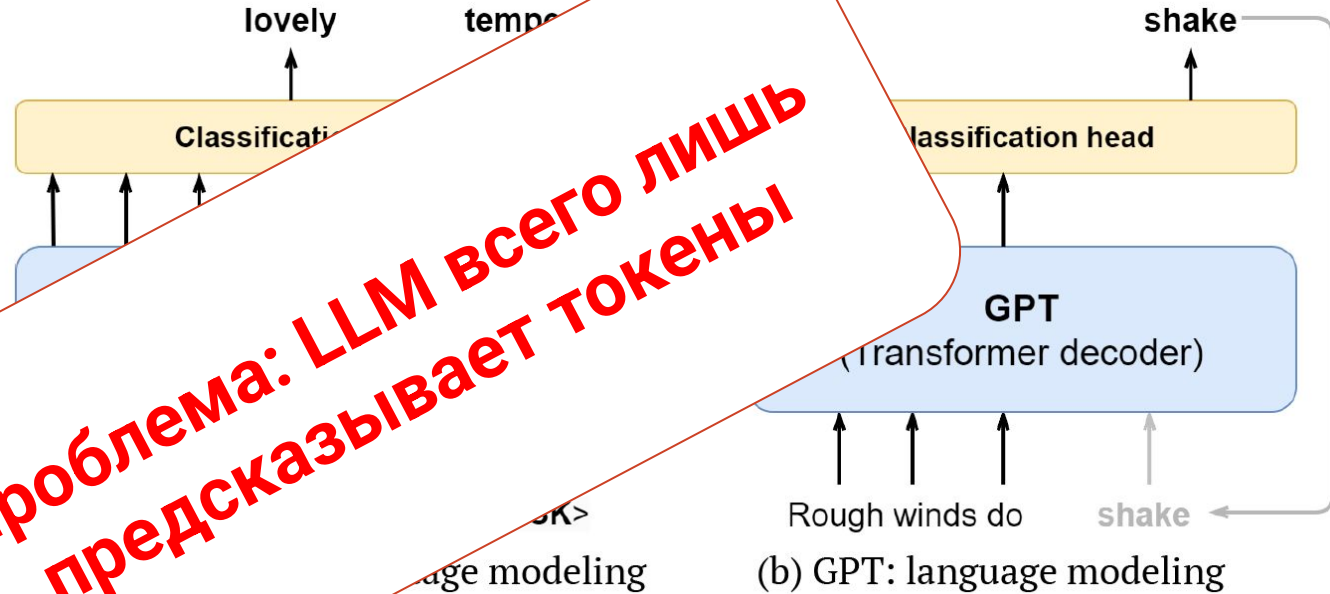
- Серия моделей GPT (generative pretrained transformer) – это декодировщик от трансформера, который как раз и предсказывает токены
- Encoder-decoder attention там уже нет, потому что нет кодировщика



# BERT и GPT

- Оказалось, что кодировщик и декодировщик могут быть очень полезны и отдельно, и в семействе BERT и GPT

**Проблема: LLM всего лишь предсказывает токены**



# Наш план

- Куда пришли языковые модели сейчас
- Как дообучать большие языковые модели (fine-tuning)
- Как увеличить длину контекста
- Как расширить модель поиском (RAG)
- Как научить модель думать (reasoning)
- Пример другой идеи: сети Колмогорова-Арнольда (*optional*)



# 3. Масштабирование и прогресс



Безграничный простор для научно-фантастического жанра открывается только у нас, в СССР, в стране реализованных "утопий"... Масштабы нашего строительства воистину грандиозны. Эти масштабы растут с каждым годом, с каждой пятилеткой. Им нет и не может быть границ.

*Александр Беляев.  
Застрельщики новых открытий*

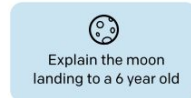
# От GPT к ChatGPT

- ChatGPT начинается с GPT-3, которая учится только предсказывать токены
- Затем – fine-tuning через RLHF (reinforcement learning from human feedback)

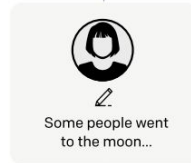
## Step 1

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

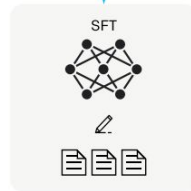
A prompt is sampled from our prompt dataset.



A labeler demonstrates the desired output behavior.



This data is used to fine-tune GPT-3 with supervised learning.



## Step 2

**Collect comparison data, and train a reward model.**

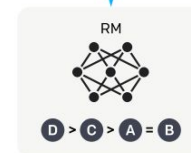
A prompt and several model outputs are sampled.



A labeler ranks the outputs from best to worst.



This data is used to train our reward model.



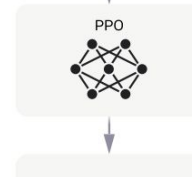
## Step 3

**Optimize a policy against the reward model using reinforcement learning.**

A new prompt is sampled from the dataset.



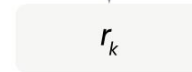
The policy generates an output.



The reward model calculates a reward for the output.



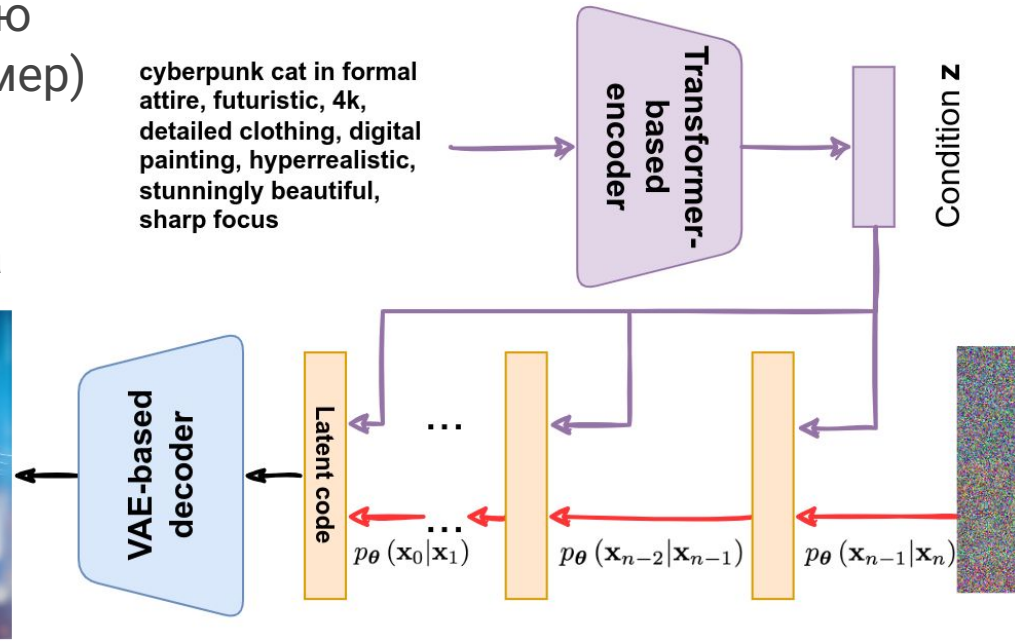
The reward is used to update the policy using PPO.





# Stable Diffusion и другие

- В то же время появляются и хорошие text-to-image модели: DALL-E, затем Stable Diffusion, Midjourney, DALL-E 2, потом DALL-E 3, Flux, Imogen...
- Работают, обучая порождающую модель (диффузионную, например) в латентном пространстве автокодировщика картинок (VQ-VAE, VQ-GAN), а тексты туда отображает трансформер
- Не будем углубляться...



# GPT-4

- GPT-4 была заявлена как мультимодальная модель, она может обрабатывать картинки
- Объявлена OpenAI 14 марта 2023 г., в тот же день, когда Anthropic объявил Claude, а Google сделал презентацию о встраивании LLM в свои инструменты

User

What is unusual about this image?



Source: <https://www.barnorama.com/wp-content/uploads/2016/12/03-Confusing-Pictures.jpg>

GPT-4

The unusual thing about this image is that a man is ironing clothes on an ironing board attached to the roof of a moving taxi.

# GPT-4

- Возможности GPT-4 сразу же были неплохо изучены, и они поражали
- [“Sparks of Artificial General Intelligence: Early experiments with GPT-4”](#)

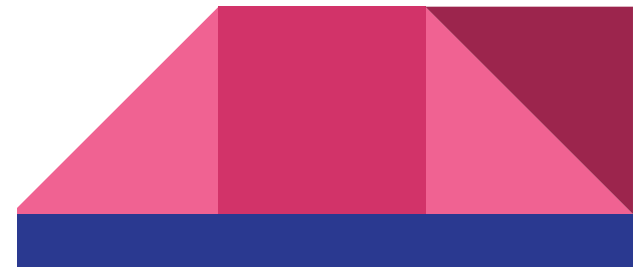
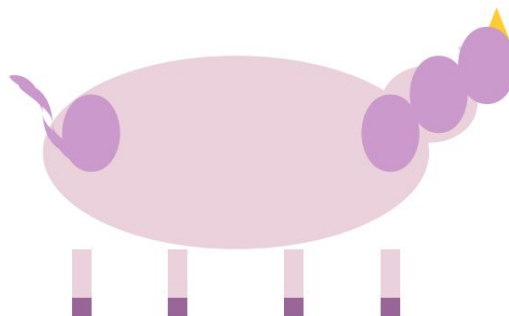
**Prompt:** Can you write a proof that there are infinitely many primes, with every line that rhymes?

**GPT-4:**

Yes, I think I can,  
Though it might take a clever plan.  
I'll start by noting Euclid's proof,  
Which shows that primes aren't just aloof.  
Assume we have a finite list,  
Of primes, and that none have been missed.  
Multiply them all together,  
And add one, just to be clever.  
The result will be a number,  
That has no prime factors, oh wonder!  
But every number has a prime,  
That divides it, it's just a matter of time.  
So we've found a contradiction,  
And our finite list needs eviction.  
There must be infinitely many primes,  
And that's the end of my rhyming lines.s

**Prompt:** Draw a unicorn in TikZ.

**GPT-4:** [Produces  $\LaTeX$  compiling to following picture.]



# GPT-4

- Возможности GPT-4 справились с задачей, которую раньше не поражали
- “Sparks of Artificial General Intelligence” GPT-4”

**Prompt:** Can you write a poem about the infinitude of primes, with every line rhyming with the previous line.

**GPT-4:**

Yes, I think I can,  
Though it might take a clever plan,  
I'll start by noting Euclid's proof,  
Which shows that primes aren't just a few.  
Assume we have a finite list,  
Of primes, and that none have been missed.  
Multiply them all together,  
And add one, just to be clever.  
The result will be a number,  
That has no prime factors, oh wonder!  
But every number has a prime,  
That divides it, it's just a matter of time.  
So we've found a contradiction,  
And our finite list needs eviction.  
There must be infinitely many primes,  
And that's the end of my rhyming lines.



# GPT-4

- Возможности GPT-4 способны поражали
- “Sparks of Art” **27 февраля 2025** GPT-4”

**Prompt:** Can you write a poem about many primes, with every line

**GPT-4:**

Yes, I think I can,  
Though it might take a clever plan,  
I'll start by noting Euclid's proof,  
Which shows that primes aren't just  
Assume we have a finite list,  
Of primes, and that none have been missed,  
Multiply them all together,  
And add one, just to be clever.  
The result will be a number,  
That has no prime factors, oh wonder!  
But every number has a prime,  
That divides it, it's just a matter of time,  
So we've found a contradiction,  
And our finite list needs eviction.  
There must be infinitely many primes,  
And that's the end of my rhyming lines.s

GPT-4.5

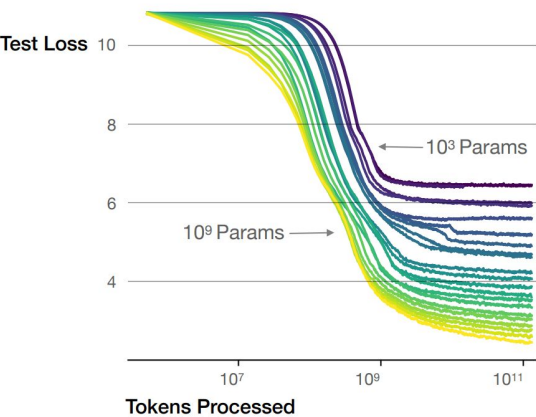
Пока так, и что такое GPT-5, уже не совсем ясно, но всё-таки :)

OpenAI

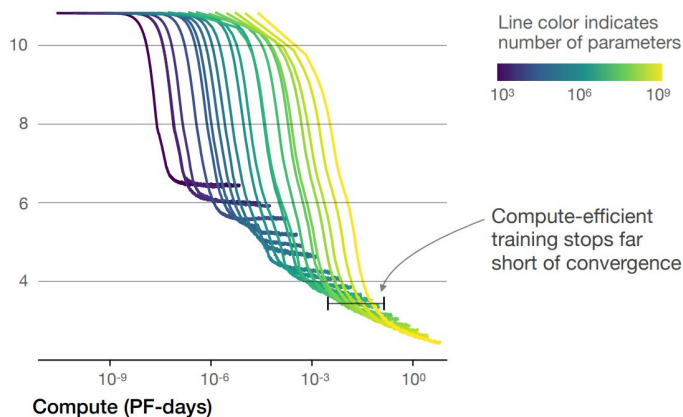
# Семейство o1: scaling laws

- В сентябре 2024 OpenAI выпустил семейство моделей o1; сначала o1-preview ([OpenAI, September 12, 2024](#)), сейчас есть o1-mini, o1-pro, o3-mini[-high]...
- Законы масштабирования ML-моделей выполняются на удивление точно ([Kaplan et al., 2020](#))

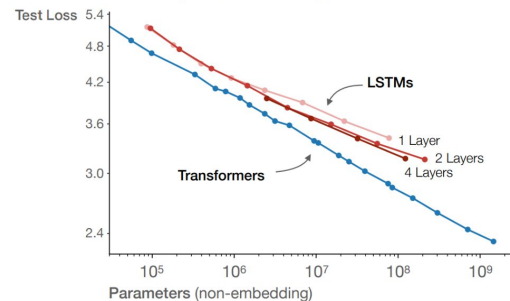
Larger models require **fewer samples** to reach the same performance



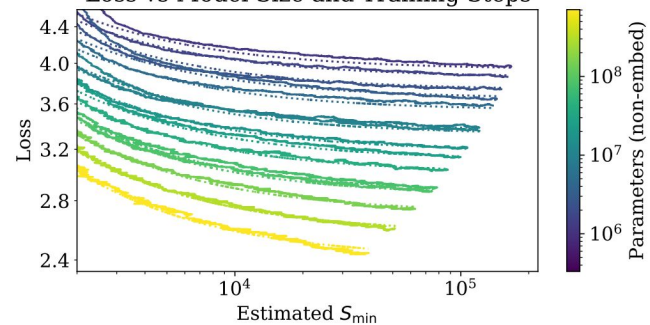
The optimal model size grows smoothly with the loss target and compute budget



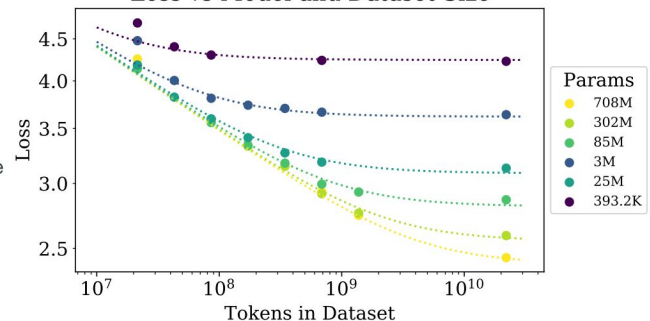
Transformers asymptotically outperform LSTMs due to improved use of long contexts



Loss vs Model Size and Training Steps

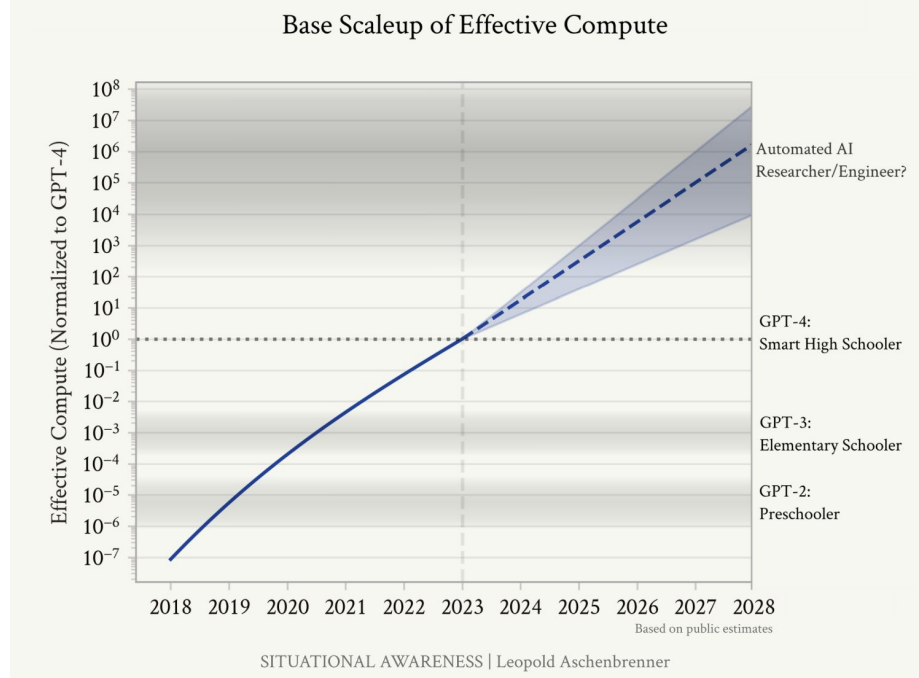


Loss vs Model and Dataset Size



# Situational Awareness

- Леопольд Ашенбреннер, бывший сотрудник OpenAI, в июне выпустил книгу ([Aschenbrenner, 2024](#))
- Предсказания в основном на основе продолжения трендов, но что если они продолжатся?..



## SITUATIONAL AWARENESS

### The Decade Ahead

Introduction I. From GPT-4 to AGI: Counting the OOMs II. From AGI to Superintelligence: the Intelligence Explosion

IIIa. Racing to the Trillion-Dollar Cluster IIIb. Lock Down the Labs: Security for AGI IIIc. Superalignment IIId. The Free World Must Prevail

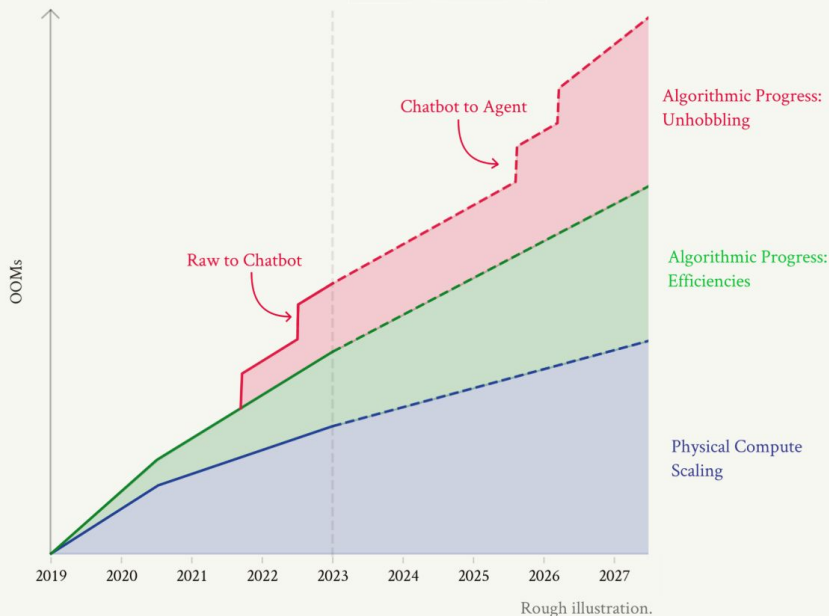
IV. The Project V. Parting Thoughts Full series as PDF About Dwarkesh podcast



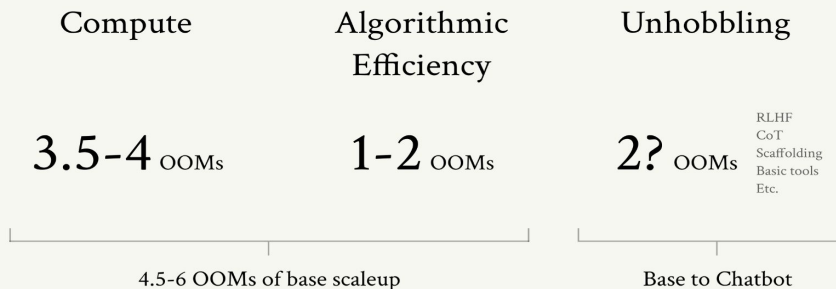
# Situational Awareness

- Графики Ашенбренера

## Decomposing drivers of progress



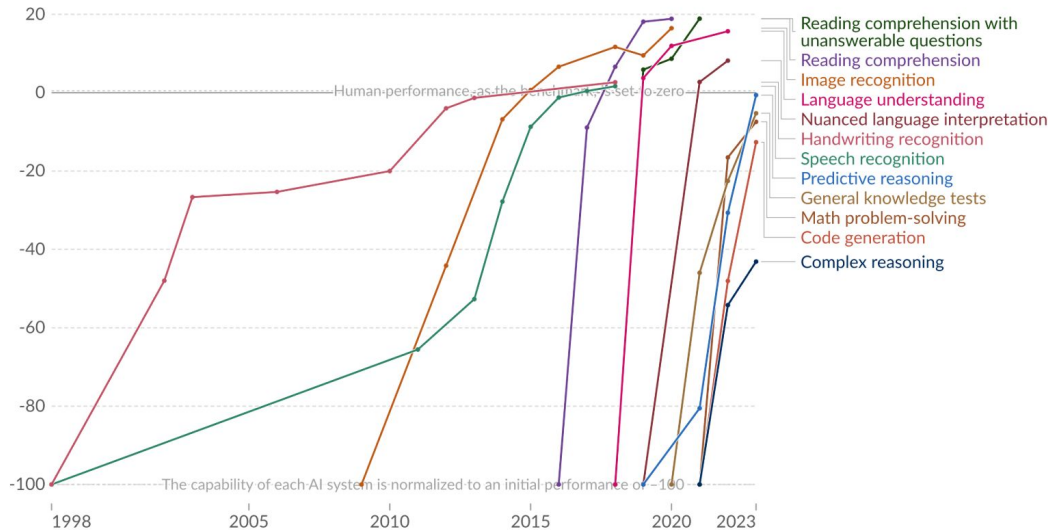
## GPT-2 (2019) to GPT-4 (2024)



## Test scores of AI systems on various capabilities relative to human performance



Within each domain, the initial performance of the AI is set to -100. Human performance is used as a baseline, set to zero. When the AI's performance crosses the zero line, it scored more points than humans.

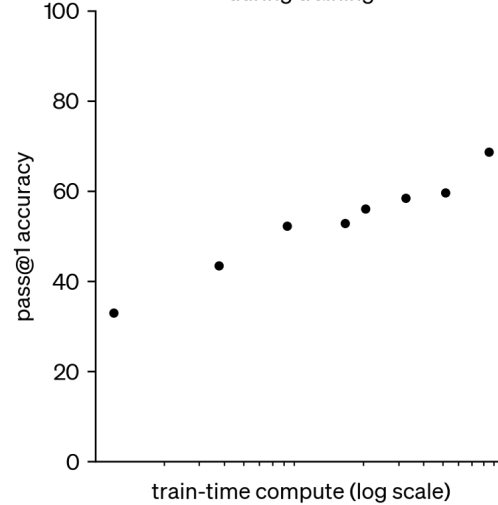




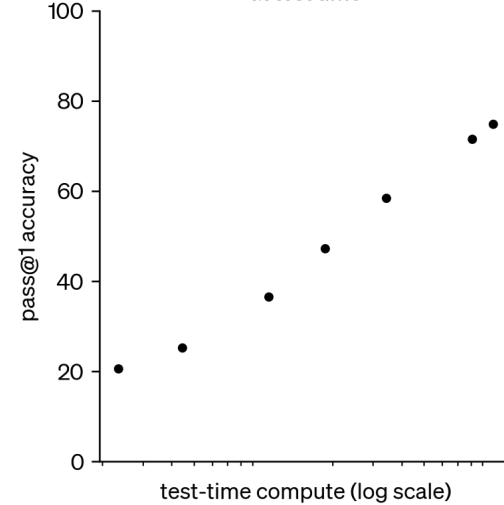
# Семейство моделей o1

- А семейство o1 открывает фактически новый способ масштабировать модели, которого раньше не было
- Reasoning models (o1, DeepSeek-R1, Gemini 2.0, Claude 3.7 Thinking) могут “обдумывать” заданный вопрос и масштабироваться с test-time compute

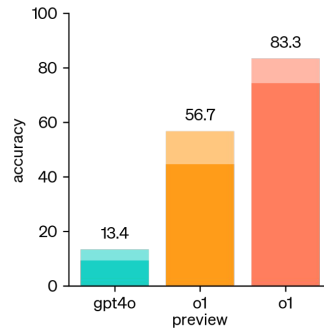
o1 AIME accuracy during training



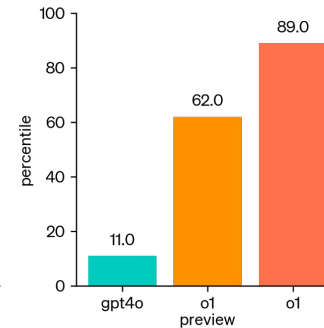
o1 AIME accuracy at test time



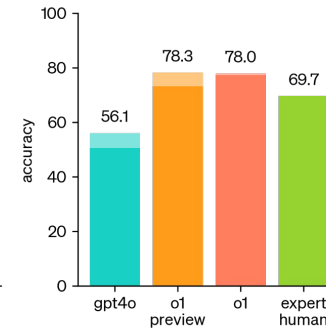
Competition Math (AIME 2024)



Competition Code (Codeforces)



PhD-Level Science Questions (GPQA Diamond)



# Семейство моделей o1

- Это приводит к большому скачку в качестве везде, где нужна логика, выводы, последовательные рассуждения...
- GPQA Diamond ([Rein et al., 2023](#)), например, это не школьный экзамен, а куда более интересные вопросы
- И это на основе GPT-4, а можно будет подставить и GPT-5, когда она появится

## Molecular Biology

A scientist studies the stress response of barley to increased temperatures and finds a protein which contributes to heat tolerance through the stabilisation of cell membrane. The scientist is very happy and wants to create a heat-tolerant cultivar of diploid wheat. Using databases, they find a heat tolerance protein homologue and start analysing its accumulation under heat stress. Soon enough, the scientist discovers this protein is not synthesised in the wheat cultivar they study. There are many possible reasons for such behaviour, including:

- A) A miRNA targets the protein, which makes exonucleases cut it immediately after the end of translation and before processing in ER
- B) Trimethylation of lysine of H3 histone in position 27 at the promoter of the gene encoding the target protein
- C) A stop-codon occurs in the 5' -UTR region of the gene encoding the target protein
- D) The proteolysis process disrupts a quaternary structure of the protein, preserving only a tertiary structure

## Astrophysics

Astronomers are studying a star with a  $T_{\text{eff}}$  of approximately 6000 K. They are interested in spectroscopically determining the surface gravity of the star using spectral lines ( $EW < 100$  mÅ) of two chemical elements, E11 and E12. Given the atmospheric temperature of the star, E11 is mostly in the neutral phase, while E12 is mostly ionized. Which lines are the most sensitive to surface gravity for the astronomers to consider?

- A) E12 I (neutral)
- B) E11 II (singly ionized)
- C) E12 II (singly ionized)
- D) E11 I (neutral)

## Quantum Mechanics

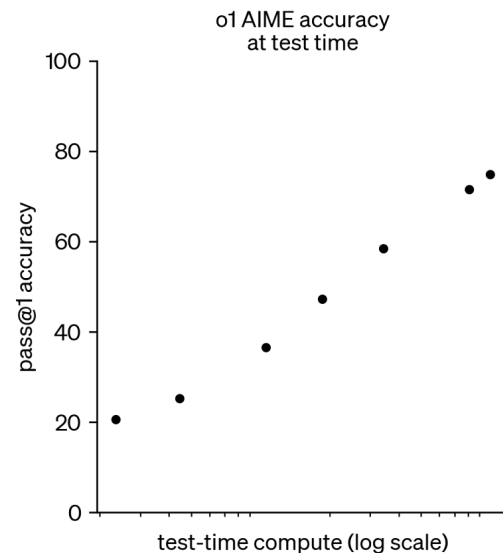
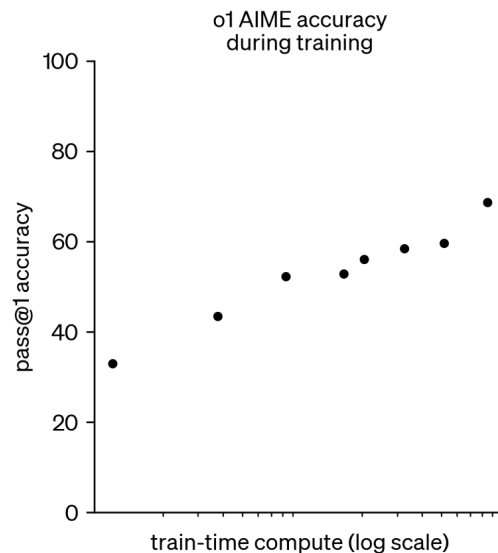
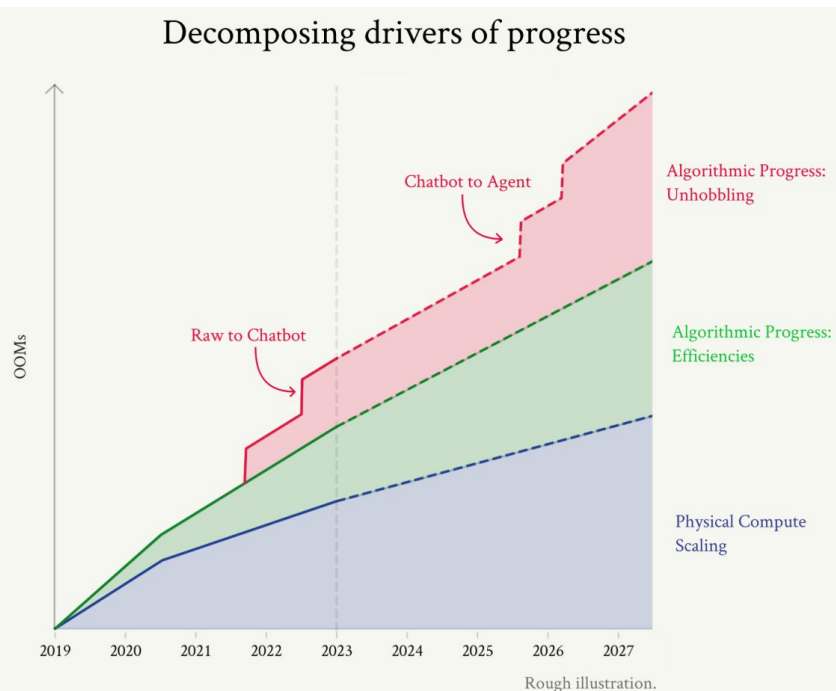
Suppose we have a depolarizing channel operation given by  $E(\rho)$ . The probability,  $p$ , of the depolarization state represents the strength of the noise. If the Kraus operators of the given state are  $A_0 = \sqrt{1 - \frac{3p}{4}}$ ,  $A_1 = \sqrt{\frac{p}{4}}X$ ,  $A_2 = \sqrt{\frac{p}{4}}Y$ , and

$A_3 = \sqrt{\frac{p}{4}}Z$ . What could be the correct Kraus Representation of the state  $E(\rho)$ ?

- A)  $E(\rho) = (1 - p)\rho + \frac{p}{3}X\rho X + \frac{p}{3}Y\rho Y + \frac{p}{3}Z\rho Z$
- B)  $E(\rho) = (1 - p)\rho + \frac{p}{3}X\rho^2 X + \frac{p}{3}Y\rho^2 Y + \frac{p}{3}Z\rho^2 Z$
- C)  $E(\rho) = (1 - p)\rho + \frac{p}{4}X\rho X + \frac{p}{4}Y\rho Y + \frac{p}{4}Z\rho Z$
- D)  $E(\rho) = (1 - p)\rho^2 + \frac{p}{3}X\rho^2 X + \frac{p}{3}Y\rho^2 Y + \frac{p}{3}Z\rho^2 Z$

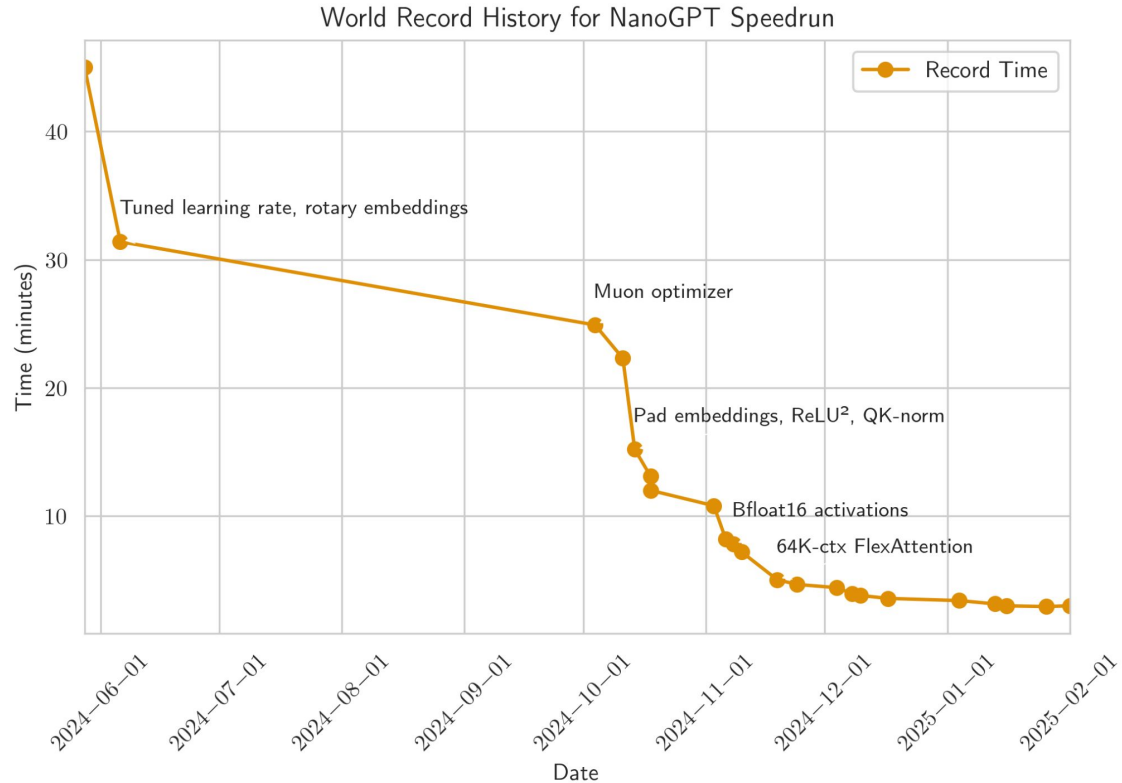
# Situational Awareness vs o1-preview

- Это новая кривая масштабирования или часть алгоритмического прогресса? Наверное, это не так уж важно



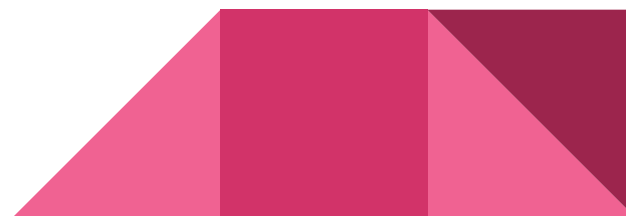
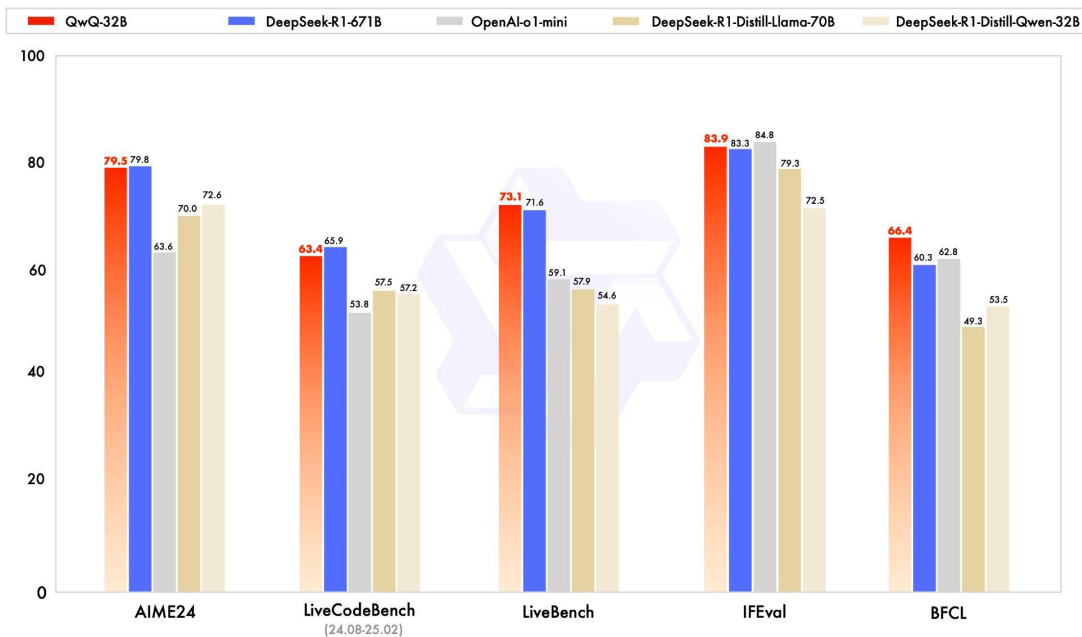
# Алгоритмический прогресс: NanoGPT

- [NanoGPT](#): хороший пример алгоритмического прогресса
- В мае 2024 [Андрей Карпатый опубликовал llm.c](#) – репликацию GPT-2 (small, 124М весов) за **90 минут** на кластере 8xH100; это было круто и неожиданно
- Люди сделали из этого leaderboard, и 1 февраля 2025-го рекорд достижения того же уровня validation loss на том же железе и датасете составил... **три минуты**



# Предостережение о бенчмарках

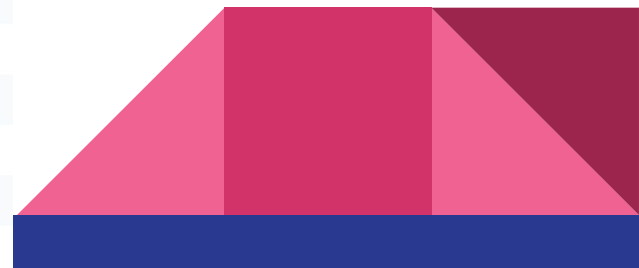
- Новые модели (особенно маленькие) бывают хороши в тестах, но не в реальной работе  
 ([QwQ-32B, 6 марта 2025](#))



# Предостережение о бенчмарках

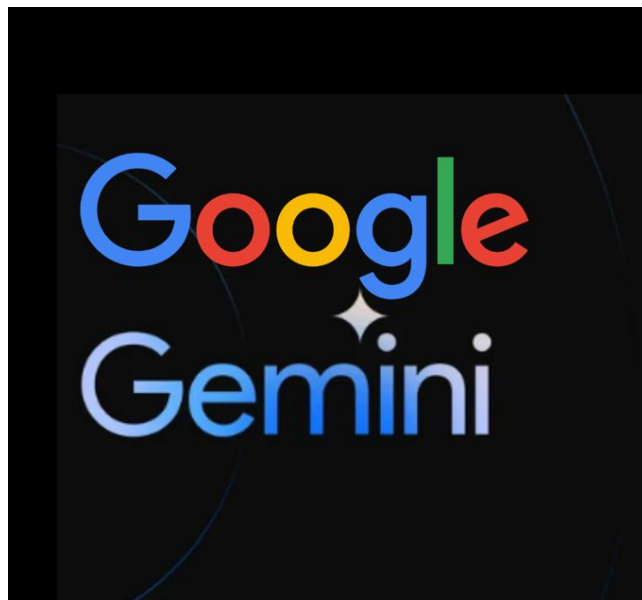
- Например, [Chatbot Arena](#) – это крутая идея, до сих пор что-то показывает, но насыщается, и не факт что полезна для реального выбора

Rank* (UB)	Rank (StyleCtrl)	Model	Arena Score
1	2	<a href="#">Grok-3-Preview-02-24</a>	1407
1	1	<a href="#">GPT-4.5-Preview</a>	1404
3	6	<a href="#">Gemini-2.0-Flash-Thinking-Exp-01-21</a>	1384
3	3	<a href="#">Gemini-2.0-Pro-Exp-02-05</a>	1380
3	2	<a href="#">ChatGPT-4o-latest (2025-01-29)</a>	1375
6	4	<a href="#">DeepSeek-R1</a>	1361
6	10	<a href="#">Gemini-2.0-Flash-001</a>	1355
6	3	<a href="#">o1-2024-12-17</a>	1353
9	10	<a href="#">Gemma-3-27B-it</a>	1339
9	10	<a href="#">Qwen2.5-Max</a>	1338
9	7	<a href="#">o1-preview</a>	1335
9	10	<a href="#">o3-mini-high</a>	1328

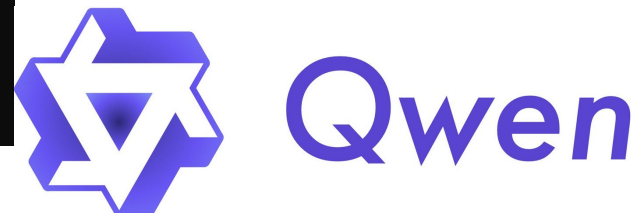


# Не только OpenAI

- Есть и другие хорошие LLM



**LLaMA**  
by  **Meta**



# Не только OpenAI

- И они очень быстро стали “рассуждающими”



Claude 3.7 Sonnet: Frontier reasoning made practical

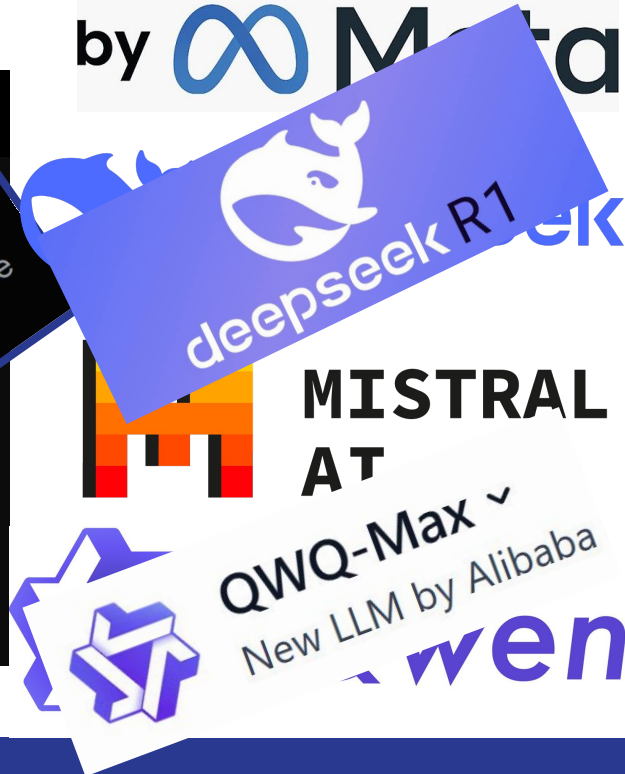
3.5 Sonnet

ANTHROPIC

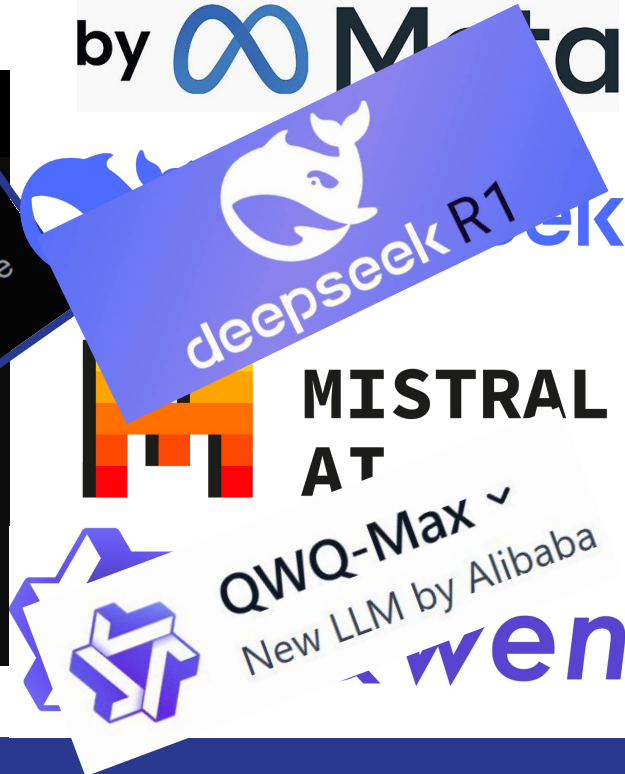


Gemini 2.0 Flash Thinking

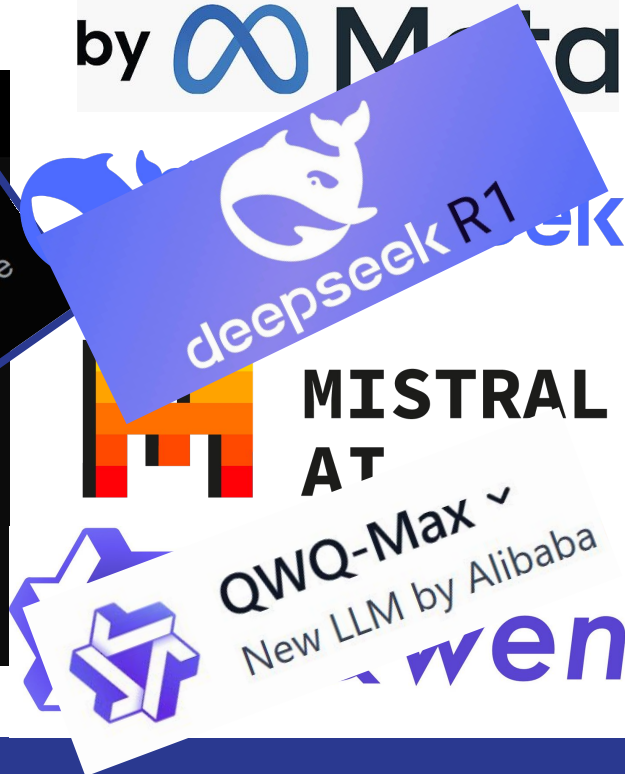
Our enhanced reasoning model, capable of showing its thoughts to improve performance and explainability



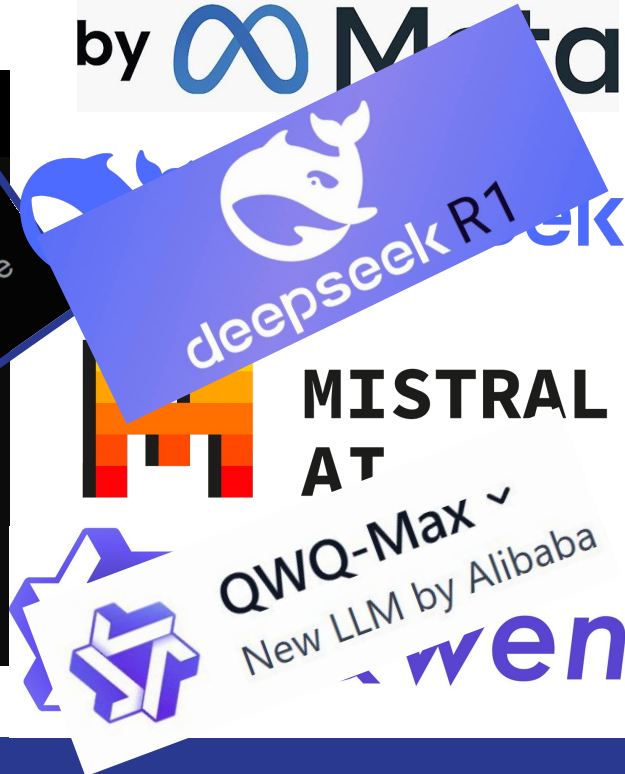
LLaMA by Meta



deepseek R1



MISTRAL AI



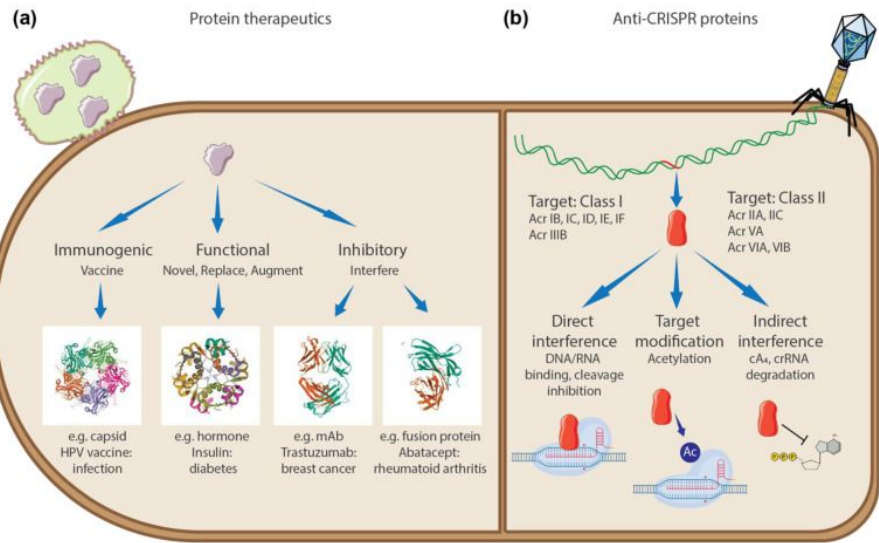
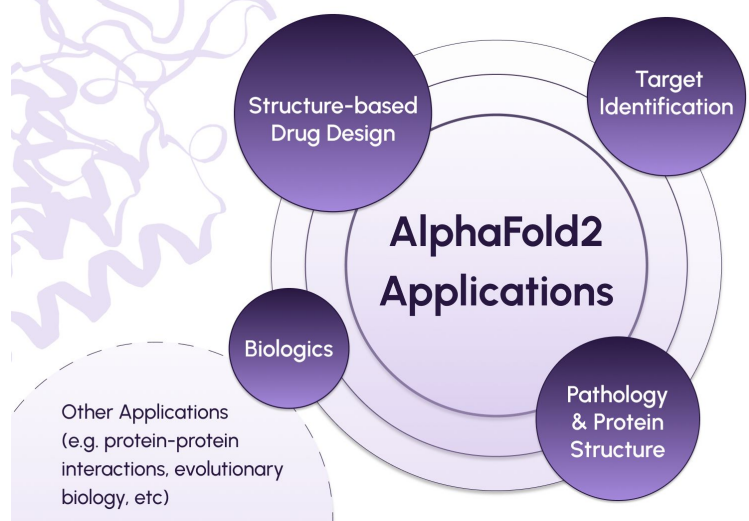
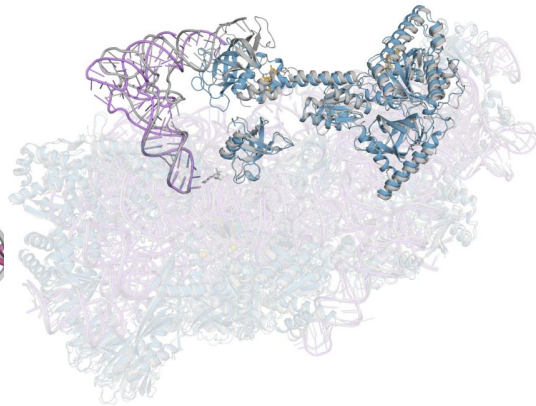
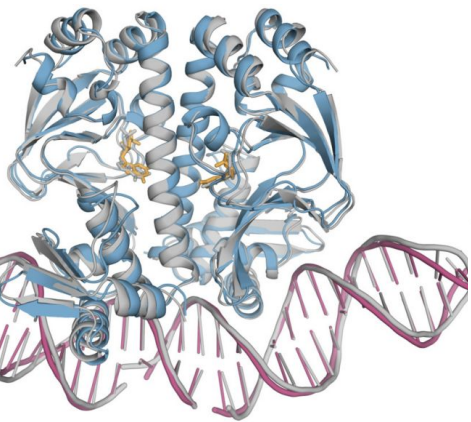
QWQ-Max  
New LLM by Alibaba

Qwen



# И не только LLM

- AlphaFold от DeepMind уже сейчас может произвести революцию в медицине
- AlphaFold 3 ([Google, May 2024](#))



# Кстати, об AlphaFold и DeepMind в целом

PRESS RELEASE

9 October 2024



## The Nobel Prize in Chemistry 2024

The Royal Swedish Academy of Sciences has decided to award the Nobel Prize in Chemistry 2024 with one half to **David Baker** and the other half jointly to **Demis Hassabis** and **John M. Jumper**

**David Baker**

University of Washington, Seattle, WA, USA

**Demis Hassabis**

Google DeepMind, London, UK

**John M. Jumper**

Google DeepMind, London, UK

*“for computational protein design”*

*“for protein structure prediction”*



**They cracked the code for proteins' amazing structures**

# 4. Примеры возможностей

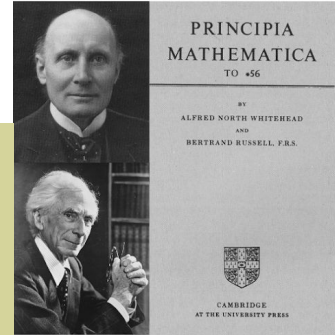
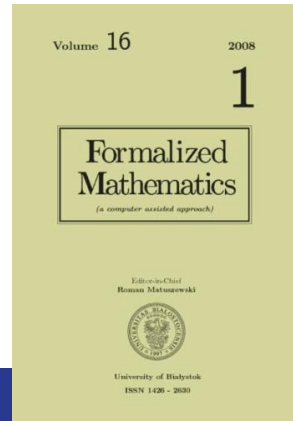
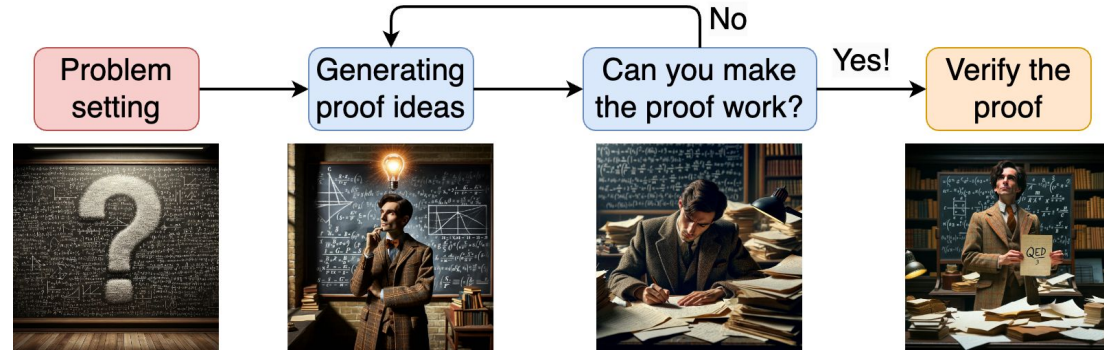


Я не знаю, до чего достигнем мы; говорят, что каждый век имеет своеобразную поэзию. И, действительно, мудрено отделаться от принятых, унаследованных привычек, направлений, манеры выражаться. Но всё же точки зрения давно стали меняться, а с ними и впечатление окружающего мира стало иначе отражаться в понимании.

*Николай Огарёв. С утра до ночи*

# LLM для математики

- Работа математика – это интересный случай, который всегда оставался для меня загадкой
- В математической логике давно развиваются методы автоматического доказательства теорем, но они так ни к чему выдающемуся и не привели...



\*54.43.  $\vdash \cdot \alpha, \beta \in 1 \cdot \supset \cdot \alpha \cap \beta = \Lambda \cdot \equiv \cdot \alpha \cup \beta \in 2$

*Dem.*

$\vdash \cdot *54.26 \cdot \supset \vdash \cdot \alpha = \iota'x \cdot \beta = \iota'y \cdot \supset \cdot \alpha \cup \beta \in 2 \cdot \equiv \cdot x \neq y \cdot$   
 $[*51.231] \quad \quad \quad \equiv \cdot \iota'x \cap \iota'y = \Lambda \cdot$   
 $[*13.12] \quad \quad \quad \equiv \cdot \alpha \cap \beta = \Lambda \quad (1)$

$\vdash \cdot (1) \cdot *11.11.35 \cdot \supset$   
 $\vdash \cdot (\exists x, y) \cdot \alpha = \iota'x \cdot \beta = \iota'y \cdot \supset \cdot \alpha \cup \beta \in 2 \cdot \equiv \cdot \alpha \cap \beta = \Lambda \quad (2)$

$\vdash \cdot (2) \cdot *11.54 \cdot *52.1 \cdot \supset \vdash \cdot \text{Prop}$

From this proposition it will follow, when arithmetical addition has been defined, that  $1 + 1 = 2$ .

# LLM для математики

- Но успех, кажется, придёт со стороны LLM, которые ничего формально выводить в системах доказательств не умеют (как и живые математики, впрочем)...
- Minerva ([Lewkowycz et al., 2022](#))

**Question:** Assume that the variance of the first  $n$  natural numbers is 10, and the variance of the first  $m$  even natural numbers is 16. Compute  $m + n$ .

**Model output:** Let  $\sigma^2$  be the variance of the first  $m$  even natural numbers, and  $\tau^2$  be the variance of the first  $n$  natural numbers. Then, we have  $\tau^2 = \frac{n^2-1}{12}$  and  $\sigma^2 = \frac{m^2-1}{3}$ . We solve the system of equations:

$$\frac{n^2 - 1}{12} = 10$$

$$\frac{m^2 - 1}{3} = 16.$$

This translates into the system

$$n^2 = 121$$

$$m^2 = 48 + 1 = 49.$$

Therefore,  $n = \boxed{11}$  and  $m = \boxed{7}$ , so  $n + m = \boxed{18}$ .

**Question:** For every  $a, b, b \neq a$  prove that

$$\frac{a^2 + b^2}{2} > \left(\frac{a + b}{2}\right)^2.$$

**Model output:**

$$\frac{a^2+b^2}{2} > \left(\frac{a+b}{2}\right)^2$$

$$\iff \frac{a^2+b^2}{2} > \frac{a^2+b^2+2ab}{4}$$

$$\iff a^2 + b^2 > \frac{a^2+b^2+2ab}{2}$$

$$\iff 2a^2 + 2b^2 > a^2 + b^2 + 2ab$$

$$\iff a^2 + b^2 > 2ab$$

$$\iff a^2 + b^2 - 2ab > 0$$

$$\iff (a - b)^2 > 0$$

which is true, because the square of a real number is positive.

# LLM для математики

- ToRA  
 (tool-integrated reasoning agent;  
[Gou et al., Sep 29, 2023](#)) уже делала tool-assisted математические рассуждения

Problem: Suppose that the sum of the squares of two complex numbers  $x$  and  $y$  is 7 and the sum of their cubes is 10. List all possible values for  $x + y$ , separated by commas.

We are given that  $x^2 + y^2 = 7$  and  $x^3 + y^3 = 10$ . We can factor  $x^3 + y^3$  to get  $(x + y)(x^2 - xy + y^2)$ .

Thus, we can write  $10 = (x + y)(7)$ . It follows that  $x + y$  must divide 10.

The only possibilities are that  $x + y$  is 1, 2, 5 or 10. ❌

```
import math
def solution():
    x = symbols('x'), y = symbols('y')
    x**2 + y**2 = 7
    x**3 + y**3 = 10
```

```
x = (-1 + math.sqrt(17)) / 2
y = (-1 - math.sqrt(17)) / 2
```

```
# Get all possible values for x+y
result = [x + y, -x - y]
```

```
>>> SyntaxError: cannot assign to expression. ❌
```

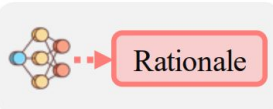
Let's use `Sympy` to calculate and print all possible values for  $x + y$ ,

from sympy import symbols, Eq, solve, simplify

```
def possible_values():
    x, y = symbols('x y')
    eq1 = Eq(x**2 + y**2, 7)
    eq2 = Eq(x**3 + y**3, 10)
    solutions = solve((eq1, eq2), (x, y))
    sums = [simplify(sol[0] + sol[1]) for sol in solutions]
    return sums
print(possible_values())
```

```
>>> [-5, -5, 1, 1, 4, 4]
```

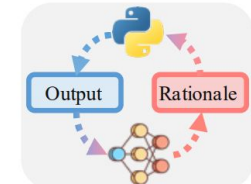
Removing duplicates, the possible values for  $x + y$  are `\boxed{-5, 1, 4}`. ✅



**(a) Rationale-based**



**(b) Program-based**






**(c) Tool-integrated Reasoning**  
 (Format used by ToRA)

# LLM для математики

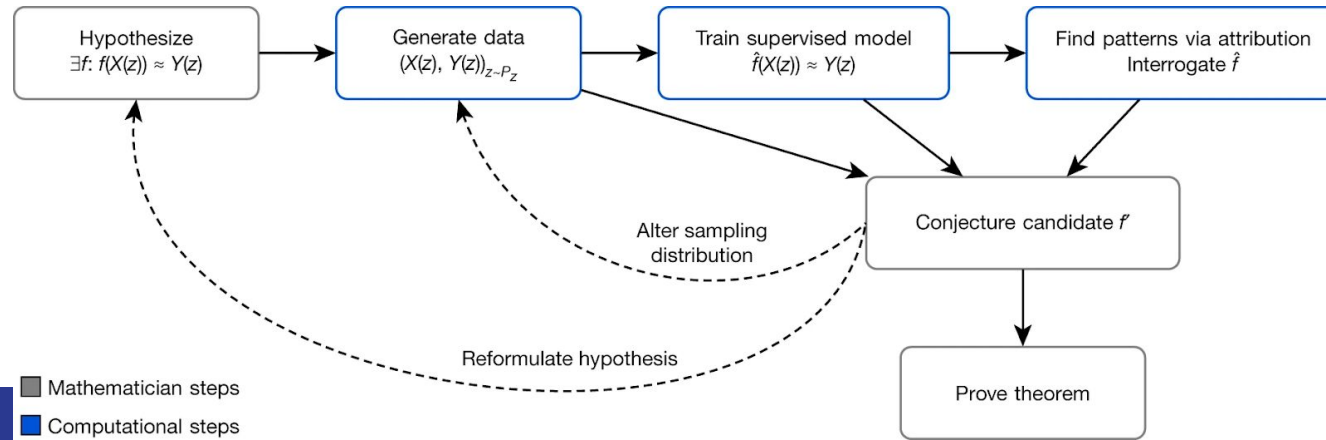
- С другой стороны, машинное обучение может помочь формулировать гипотезы
- Например, вот работа от DeepMind, где ML помогает искать отношения между математическими объектами ([Davies et al., 2021](#))

**Fig. 2: Examples of invariants for three hyperbolic knots.**

From: [Advancing mathematics by guiding human intuition with AI](#)

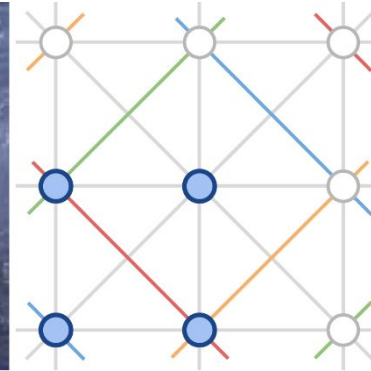
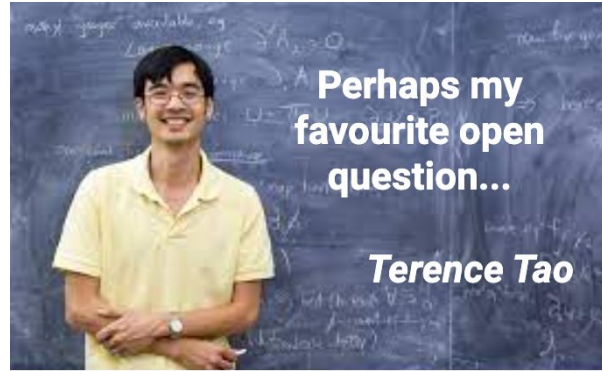
z: Knot	X(z): Geometric invariants				Y(z): Algebraic invariants		
	Volume	Chern–Simons	Meridional translation	...	Signature	Jones polynomial	...
	2.0299	0	$i$	...	0	$t^{-2} - t^{-1} + 1 - t + t^2$	...
	2.8281	-0.1532	$0.7381 + 0.8831i$	...	-2	$t - t^2 + 2t^3 - t^4 + t^5 - t^6$	...
	3.1640	0.1560	$-0.7237 + 1.0160i$	...	0	$t^{-2} - t^{-1} + 2 - 2t + t^2 - t^3 + t^4$	...

We hypothesized that there was a previously undiscovered relationship between the geometric and algebraic invariants.



# LLM для математики

- [Romera-Paredes et al. \(December 2023\)](#): DeepMind's FunSearch proves new bounds for the cap set problem



```
def priority(el: tuple[int,...],
            n: int) -> float:
    score = n
    in_el = 0
    el_count = el.count(0)

    if el_count == 0:
        score += n**2
        if el[1] == el[-1]:
            score *= 1.5
        if el[2] == el[-2]:
            score *= 1.5
        if el[3] == el[-3]:
            score *= 1.5
    else:
        if el[1] == el[-1]:
            score *= 0.5
        if el[2] == el[-2]:
            score *= 0.5

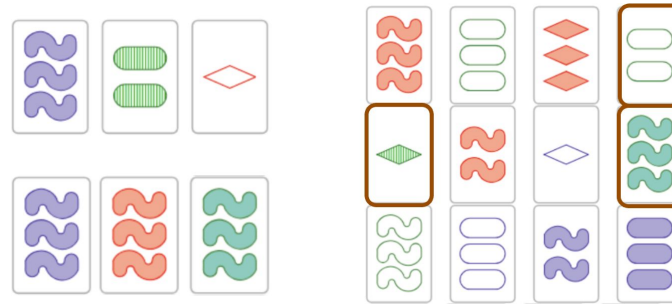
    for e in el:
        if e == 0:
            if in_el == 0:
                score *= n * 0.5
            elif in_el == el_count - 1:
                score *= 0.5
            else:
                score *= n * 0.5 ** in_el
            in_el += 1
        else:
            score += 1

    if el[1] == el[-1]:
        score *= 1.5
    if el[2] == el[-2]:
        score *= 1.5

    return score
```

$n$	3	4	5	6	7	8
Best known	9	20	45	112	236	496
FunSearch	9	20	45	112	236	<b>512</b>

- Это почти “перебор случаев”, но очень направленный и в виде функций





# LLM для математики

- Gemini 1.5 Pro ([Gemini Team, 2024](#)) решал олимпиадные задачи ещё лучше
- Но кажется, что тут должны “решать” reasoning models?..

**Problem:**

What is the minimum value of  $5x^2 + 5y^2 - 8xy$  when  $x$  and  $y$  range over all real numbers such that  $|x - 2y| + |y - 2x| = 40$ ?

**Example Math-Specialized Gemini 1.5 Pro output:**

Let  $a = x - 2y$  and  $b = y - 2x$ . Then, we have  $|a| + |b| = 40$ . Also,  $5x^2 + 5y^2 - 8xy = (x - 2y)^2 + (y - 2x)^2 = a^2 + b^2$ . Thus, we want to find the minimum value of  $a^2 + b^2$  given that  $|a| + |b| = 40$ . By the Trivial Inequality, we have  $(|a| - 20)^2 + (|b| - 20)^2 \geq 0$ , so  $a^2 - 40|a| + 400 + b^2 - 40|b| + 400 \geq 0$ , so  $a^2 + b^2 \geq 40(|a| + |b|) - 800 = 40(40) - 800 = 800$ . This minimum is achieved when  $|a| = |b| = 20$ , so the minimum value of  $5x^2 + 5y^2 - 8xy$  is 800.

---

**Problem:**

Let  $a, b, c$  be positive integers. Prove that it is impossible to have all of the three numbers  $a^2 + b + c, b^2 + c + a, c^2 + a + b$  to be perfect squares.

**Example Math-Specialized Gemini 1.5 Pro output:**

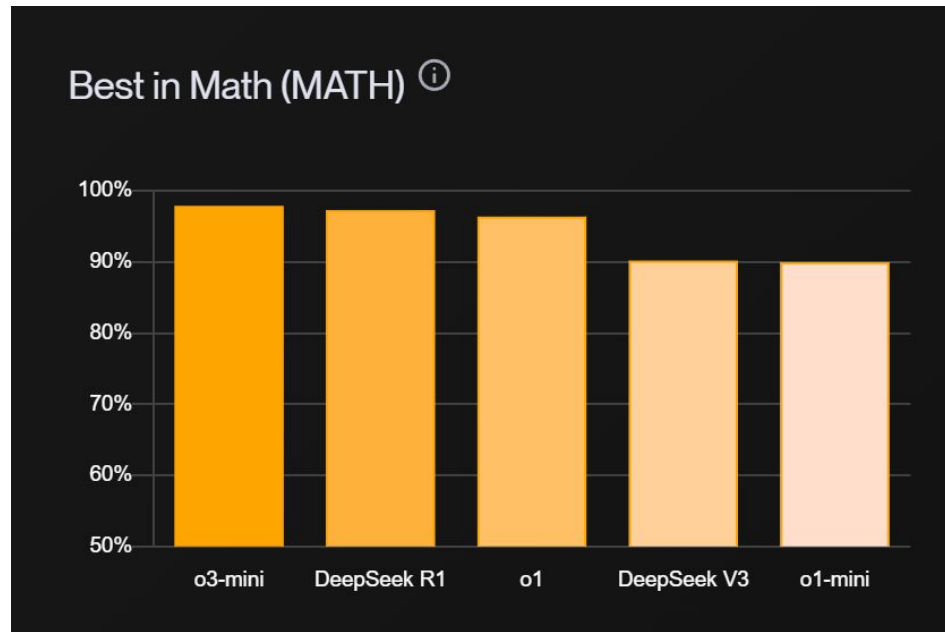
Without loss of generality, assume that  $a \geq b \geq c$ . Then  $a^2 + b + c \leq a^2 + 2a < a^2 + 2a + 1 = (a + 1)^2$ . Since  $a^2 + b + c$  is a perfect square, we must have  $a^2 + b + c = a^2$ , which implies  $b + c = 0$ . This is a contradiction since  $b$  and  $c$  are positive integers. Therefore, it is impossible to have all of the three numbers  $a^2 + b + c, b^2 + c + a, c^2 + a + b$  to be perfect squares.

# LLM для математики

- Безусловно, так и есть! MATH – это датасет с примерно такими задачами ([Hendrycks et al., 2021](#)):

**Problem:** The equation  $x^2 + 2x = i$  has two complex solutions. Determine the product of their real parts.

**Solution:** Complete the square by adding 1 to each side. Then  $(x + 1)^2 = 1 + i = e^{\frac{i\pi}{4}} \sqrt{2}$ , so  $x + 1 = \pm e^{\frac{i\pi}{8}} \sqrt[4]{2}$ . The desired product is then  $(-1 + \cos(\frac{\pi}{8}) \sqrt[4]{2})(-1 - \cos(\frac{\pi}{8}) \sqrt[4]{2}) = 1 - \cos^2(\frac{\pi}{8}) \sqrt{2} = 1 - \frac{(1 + \cos(\frac{\pi}{4}))}{2} \sqrt{2} = \boxed{\frac{1 - \sqrt{2}}{2}}$ .



# FrontierMATH

- А вот FrontierMATH ([Glazer et al., 2024](#)) — это пока совсем другое дело...
- В нём задачи с решениями, но очень непростые и относительно новые

## Find the degree 19 polynomial

Problem Solution

Construct a degree 19 polynomial  $p(x) \in \mathbb{C}[x]$  such that  $X := \{p(x) = p(y)\} \subset \mathbb{P}^1 \times \mathbb{P}^1$  has at least 3 (but not all linear) irreducible components over  $\mathbb{C}$ . Choose  $p(x)$  to be odd, monic, have real coefficients and linear coefficient -19 and calculate  $p(19)$ .

## Testing Artin's primitive root conjecture

Problem Solution

For a positive integer  $n$ , let  $v_p(n)$  denote the largest integer  $v$  such that  $p^v \mid n$ . For a prime  $p$  and  $a \not\equiv 0 \pmod{p}$ , let  $\text{ord}_p(a)$  denote the smallest positive integer  $o$  such that  $a^o \equiv 1 \pmod{p}$ . For  $x > 0$ , let

$$\text{ord}_{p,x}(a) = \prod_{\substack{q \leq x \\ q \text{ prime}}} q^{v_q(\text{ord}_p(a))} \prod_{\substack{q > x \\ q \text{ prime}}} q^{v_q(p-1)}.$$

Let  $S_x$  denote the set of primes  $p$  for which

$$\text{ord}_{p,x}(2) > \text{ord}_{p,x}(3),$$

and let  $d_x$  denote the density

$$d_x = \lim_{t \rightarrow \infty} \frac{|\{p \leq t : p \in S_x\}|}{|\{p \leq x : p \text{ prime}\}|}$$

of  $S_x$  in the primes. Let

$$d_\infty = \lim_{x \rightarrow \infty} d_x.$$

Compute  $\lfloor 10^6 d_\infty \rfloor$ .

# FrontierMATH

- А вот FrontierMATH ([Glazer et al., 2024](#)) — ЭТО ПОКА СОВСЕМ другое дело...
- В нём задачи с решениями, но очень непростые и ОТНОСИТЕЛЬНО НОВЫЕ

## Find the degree 19 polynomial

Problem      Solution

Construct a degree 19 polynomial  $p(x) \in \mathbb{C}[x]$  such that  $X := \{p(x) = p(y)\} \subset \mathbb{P}^1 \times \mathbb{P}^1$  has at least 3 (but not all linear) irreducible components over  $\mathbb{C}$ . Choose  $p(x)$  to be odd, monic, have real coefficients and linear coefficient -19 and calculate  $p(19)$ .

Problem      Solution

**Answer:** 1876572071974094803391179.

**Solution:** Note that  $p(x)$  defines a covering  $p : \mathbb{P}_x^1 \rightarrow \mathbb{P}_z^1$  of degree 19 which at  $z = \infty$  has monodromy  $\sigma_\infty$ , a 19-cycle, which without loss of generality is  $(1, 2, 3, \dots, 19)$ .

Note that  $X$  is just the fibre product of  $p$  with  $p$  (which in general is a singular curve) and therefore  $p(x)$  also defines a finite map  $\hat{p} : X \rightarrow \mathbb{P}_z^1$  of degree  $19^2$  (if we prefer to stick with smooth schemes, we can instead use the normalisation  $\hat{X}$  of  $X$ , and we will freely move between the irreducible components of  $X$  and the corresponding connected components of the smooth  $\hat{X}$  with its covering structure).

The idea then is that any such covering  $\hat{p}$  can be studied using group theory via its monodromy. Specifically, suppose  $q : X \rightarrow \mathbb{P}_z^1$  is of degree  $n$  and has branch locus  $B \subset \mathbb{P}_z^1$ . Then if you follow the fibres around each branch value and see how this permutes the points in the fibres, we get a homomorphism  $\bar{q} : \pi_1(\mathbb{P}_z^1 - B, x) \rightarrow \text{Im}(\phi) := G_q \subset S_n$  (defined up to conjugation) and  $X$  is irreducible exactly when  $\bar{q}$  is transitive. Moreover, if  $X$  is reducible, then the irreducible components correspond to the  $G_q$ -orbits in  $S_n$ .

Suppose  $p$  with branch locus  $B$  (including  $\infty$ ) is encoded as  $\bar{p} : G_p \subset S_{19}$ . Then we can check that  $\hat{p}$  has branch locus  $B$  and is encoded diagonally as  $\bar{p}^2 : G_p \subset S_{19}^2$ . So we are looking for some  $G_p \subset S_{19}$  that has between 3 and 18 orbits in  $S_{19}^2$ . In particular  $G_p$  is not doubly transitive on  $S_{19}$ .

...и ещё столько же дальше

Recall a theorem of Burnside (Burnside 1911): for  $q$  prime and a permutation group  $G \subset S_q$ , either  $G$  is doubly transitive or  $|G|$  divides  $q(q-1)$ .

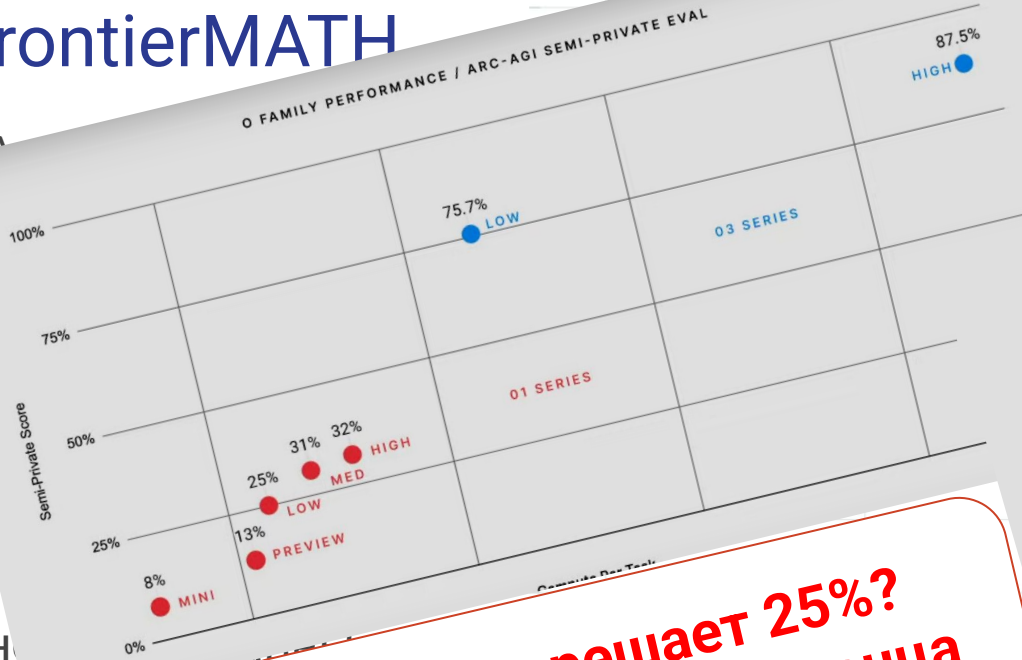
# FrontierMATH

- А вот FrontierMATH ([Glazer et al., 2024](#)) — это пока совсем другое дело...
- И здесь результаты LLM пока невысоки — но они не равны нулю! Это довольно круто
- Впрочем...

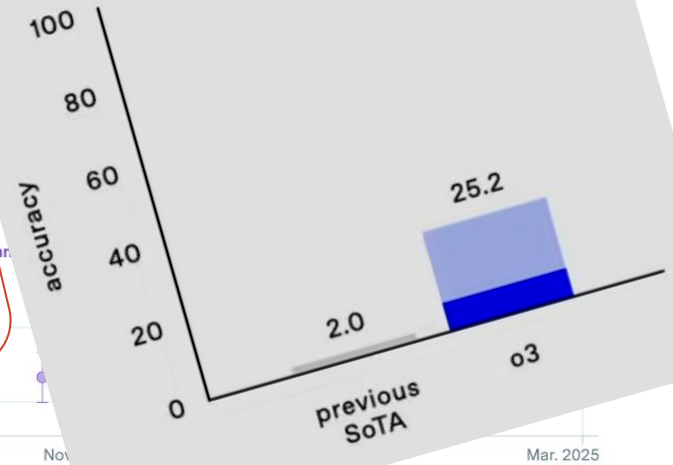


# FrontierMATH

0 FAMILY PERFORMANCE / ARC-AGI SEMI-PRIVATE EVAL



Research Math (EpochAI Frontier Math)



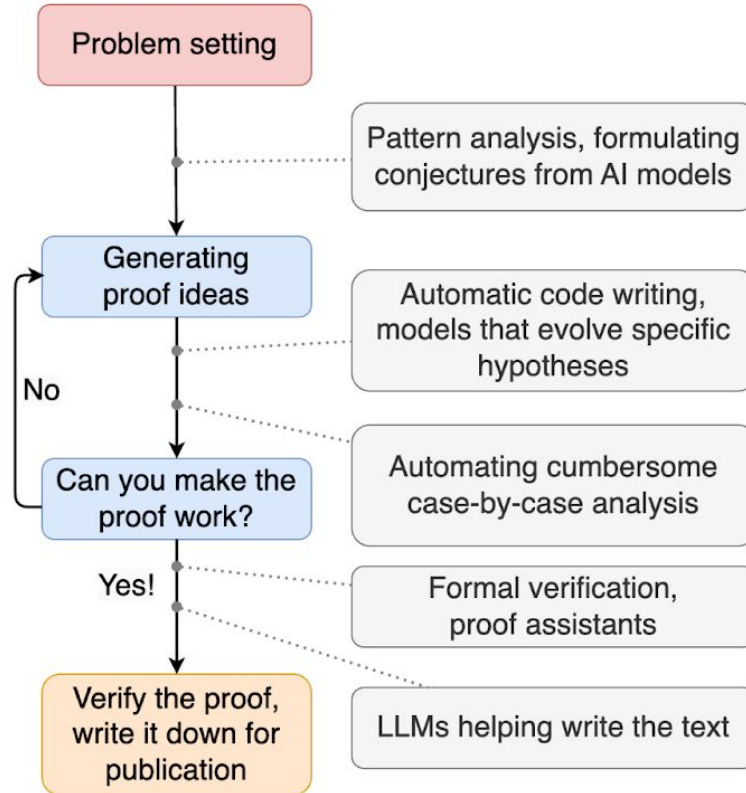
**o3 уже решает 25%?  
Здесь пока не до конца  
ясно, но всё-таки...**

Нулю! Это довольно к  
Впрочем...

- 17 Results ⓘ Organization
- OpenAI
  - Anthropic
  - Google ⓘ
  - DeepSeek
  - xAI
  - Mistral AI

# LLM для математики

- А что в других науках?



## Advancing mathematics by guiding human intuition with AI

Alex Davies<sup>1</sup>, Petar Veličković<sup>2</sup>, Lars Buesing, Sam Blackwell, Daniel Zheng, Nenad Tomasev, Richard Tanburn, Peter Battaglia, Charles Blundell, Andrés Juhász, Marc Lackertby, Georgie Williamson, Demis Hassabis & Pushmeet Kohli<sup>3</sup>

*Nature* 600, 70–74 (2021) | [Cite this article](#)



## TORA: A TOOL-INTEGRATED REASONING AGENT FOR MATHEMATICAL PROBLEM SOLVING

Zhibin Gou<sup>1,2</sup>, Zhibing Shao<sup>1,2\*</sup>, Yeyun Gong<sup>2†</sup>, Yongfeng Shen<sup>2</sup>, Yujia Yang<sup>1</sup>, Minlie Huang<sup>1</sup>, Nan Duan<sup>2</sup>, Weizha Chen<sup>2</sup>  
<sup>1</sup>Tsinghua University \*Microsoft  
 {gzbt22, szh19}@mails.tsinghua.edu.cn  
 {yeyun.gong, yoshe, nanduan, wachen}@microsoft.com

## Discovering faster matrix multiplication algorithms with reinforcement learning

Alhussein Fawzi<sup>1</sup>, Matej Balog, Aja Huang, Thomas Hubert, Bernardino Romera-Paredes, Mohammadamin Barekatalin, Alexander Novikov, Francisco J. R. Ruiz, Julian Schottwieser, Grzegorz Swiaszcz, David Silver, Demis Hassabis & Pushmeet Kohli

*Nature* 610, 47–53 (2022) | [Cite this article](#)



## Mathematical discoveries from program search with large language models

Bernardino Romera-Paredes<sup>1</sup>, Mohammadamin Barekatalin, Alexander Novikov, Matej Balog, M. Pavan Kumar, Emilian Duport, Francisco J. R. Ruiz, Jordan S. Ellenberg, Penoming Wang, Omar Fawzi, Pushmeet Kohli<sup>2</sup> & Alhussein Fawzi<sup>1</sup>

*Nature* 625, 468–475 (2024) | [Cite this article](#)

BULLETIN OF THE AMERICAN MATHEMATICAL SOCIETY  
 Volume 82, Number 3, September 1976

## RESEARCH ANNOUNCEMENTS

EVERY PLANAR MAP IS FOUR COLORABLE<sup>1</sup>

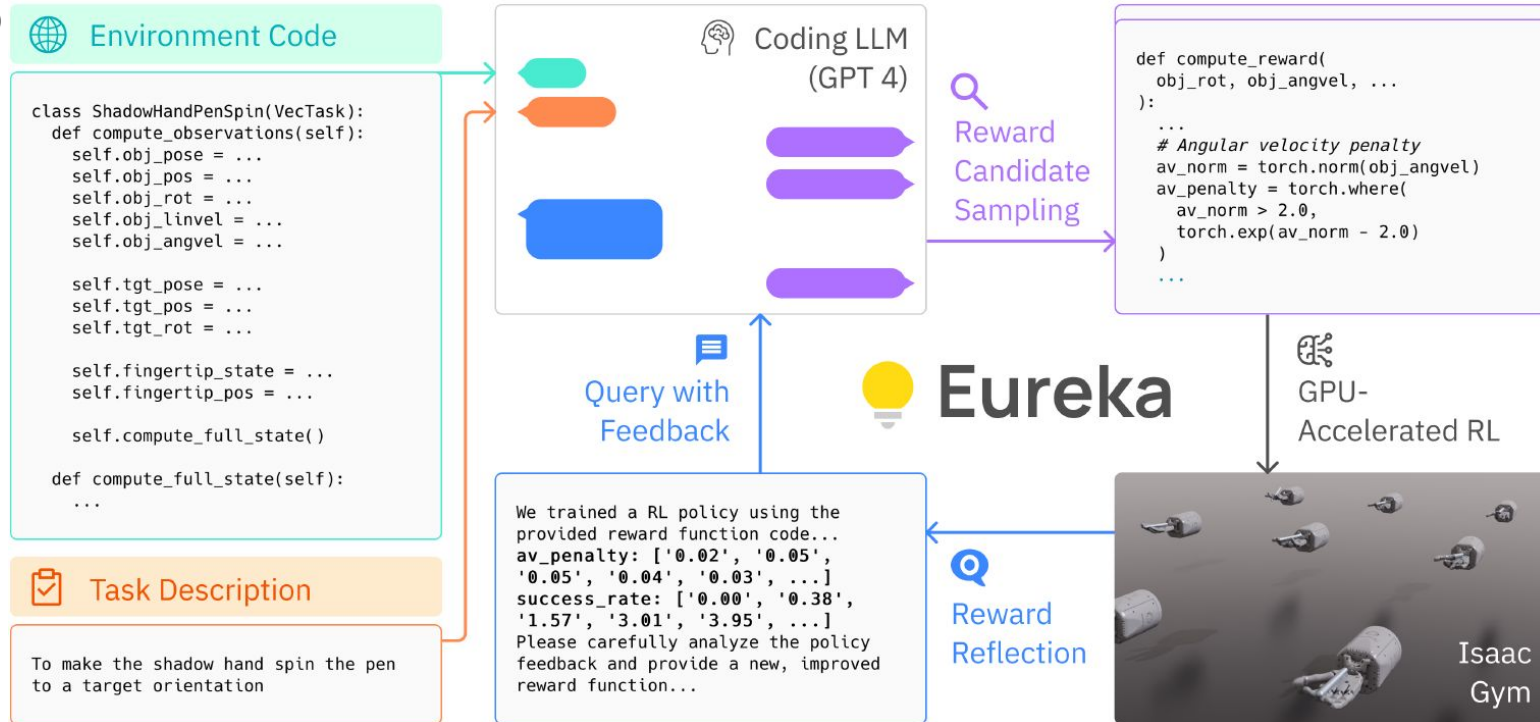
BY K. APPEL AND W. HAKEN

Communicated by Robert Fossum, July 26, 1976



# Пример из октября 2023-го: роботика

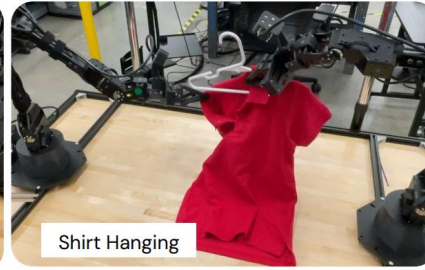
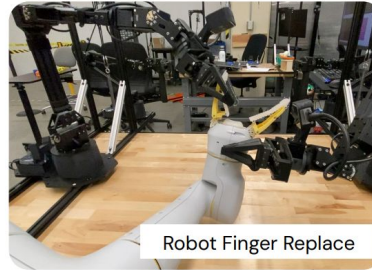
- Я раньше говорил, что с роботикой всё не так радужно
- Но вот Eureka ([Ma et al., Oct 19, 2023](#))
- Интересная автоматизация RL через LLM



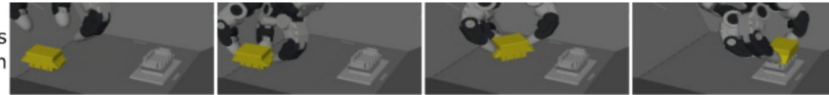


# Пример из октября 2024-го: роботика

- Прямо сейчас одни роботы от DeepMind завязывают руками шнурки и ремонтируют других роботов (ALOHA Unleashed, [DeepMind, September 2024](#)), а другие обучаются по нескольким примерам через виртуальный симулятор ([Bauza et al., September 2024](#))



20 demonstrations  
in sim



Sim policy



Zero-shot  
transfer to real

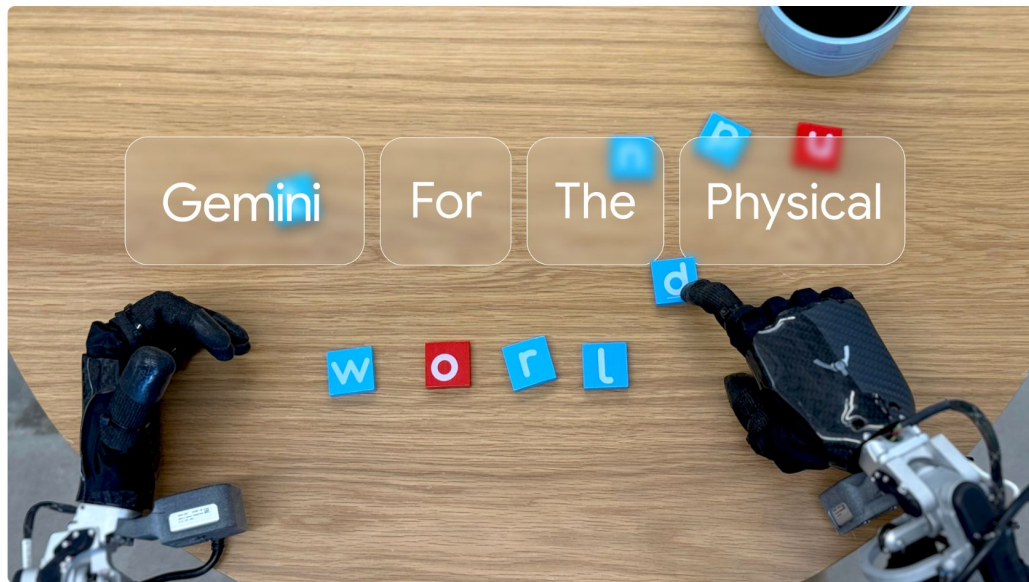


**DemoStart**  
Auto-curriculum  
+ Sparse reward

Distillation

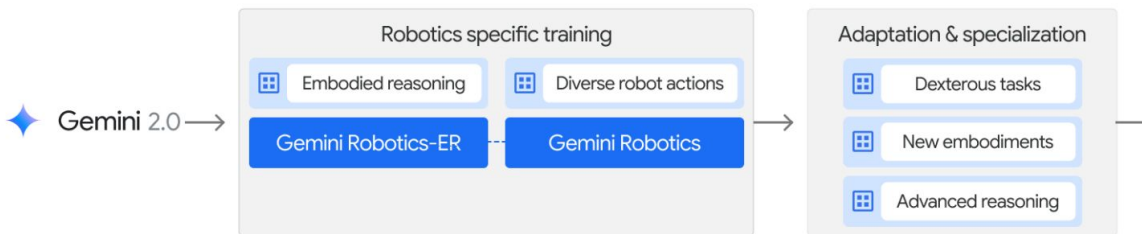
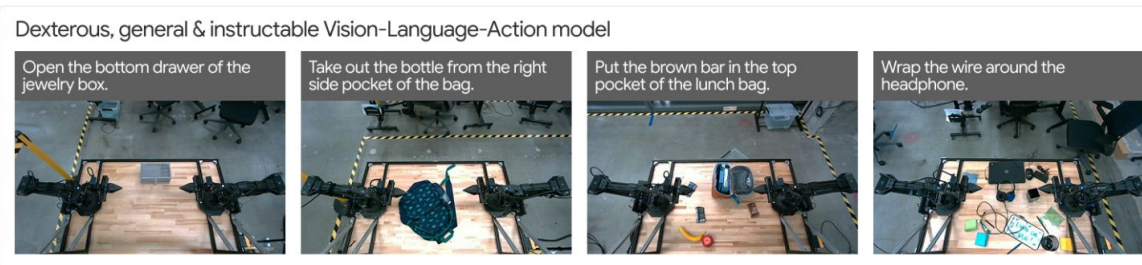
# Роботика сейчас: Gemini Robotics

- Последняя новость – Gemini Robotics ([March 12, 2025, статья](#))
- Это так называемая vision-language-action (VLA) model на основе Gemini 2.0, в которую физические действия добавлены как новая модальность выхода
- И ещё Gemini Robotics-ER (embedded reasoning)



# Роботика сейчас: Gemini Robotics

- Особых подробностей нет, в статье в основном результаты, но результаты хорошие



Complex dexterous tasks



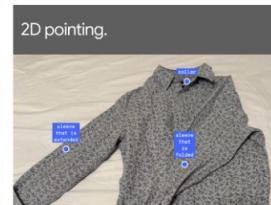
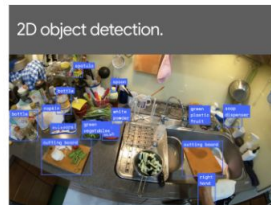
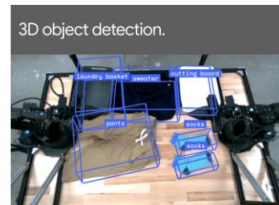
New embodiments



Advanced reasoning & acting

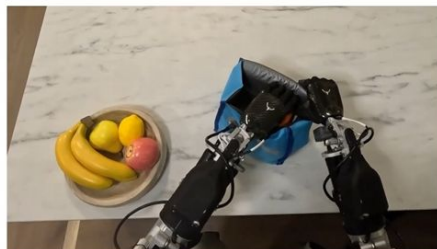


Advanced embodied reasoning for robotics



# Роботика сейчас: Gemini Robotics

- Gemini Robotics может управлять разными типами роботов



Apollo



Aloha2

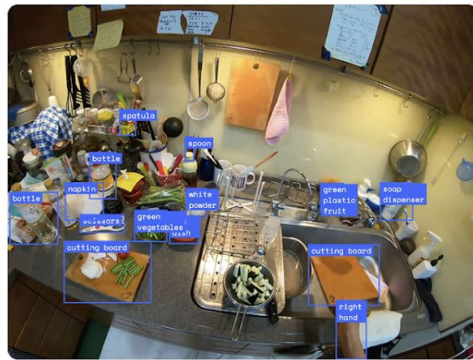


Bi-arm Franka



# Роботика сейчас: Gemini Robotics

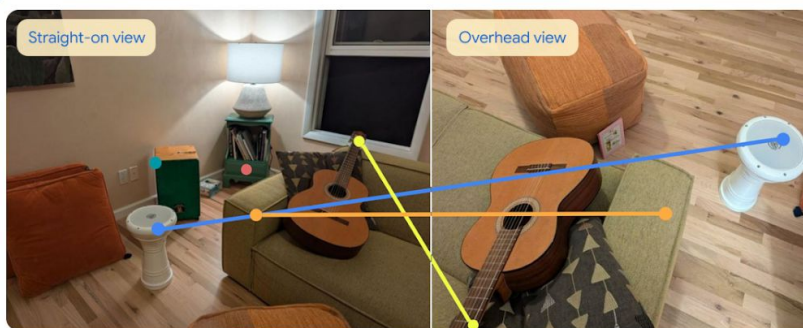
- Gemini Robotics-ER делает хороший spatial reasoning, это vision-language model



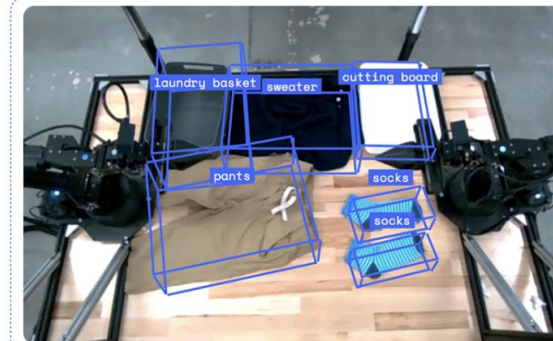
2D Object Detection



2D Pointing



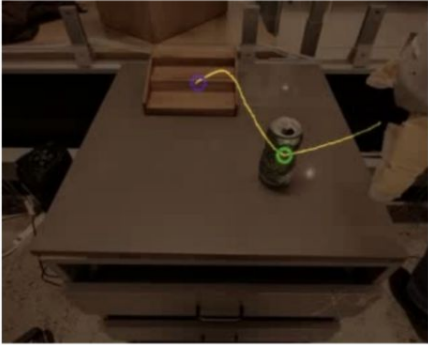

Multi-view Correspondence



3D Object Detection

# Роботика сейчас: Gemini Robotics

- Суммарно, например, может рассуждать о реальном мире (embodied QA)

Question	Reasoning Trace	Question	Reasoning Trace
 <p>If the yellow robot gripper follows the yellow trajectory, what will happen?</p> <p>A. Robot puts the soda on the wooden steps.          B. Robot moves the soda in front of the wooden steps.          C. Robot moves the soda to the very top of the wooden steps.          D. Robot picks up the soda can and moves it up.</p>	<p>Let's analyze the trajectory and the position of the soda can and the wooden steps.</p> <p><b>Step 1: Observe the starting point.</b> The yellow trajectory starts at the green circle, which is around the soda can. This indicates the robot gripper is initially positioned to interact with the can.</p> <p><b>Step 2: Analyze the trajectory path.</b> The yellow line moves upwards from the can, indicating a lifting motion. Then, it curves towards the wooden steps.</p> <p><b>Step 3: Observe the ending point.</b> The trajectory ends at the purple circle, which is located on the first step of the wooden structure.</p> <p><b>Step 4: Combine the observations.</b> The robot gripper starts at the can, moves upwards (picking up the can), and then moves towards the wooden steps, ending on the first step.</p> <p><b>Step 5: Deduce the outcome.</b> Based on the trajectory, the robot will pick up the soda can and place it on the first step of the wooden structure.</p> <p>Final Answer: (A)</p>	 <p>How does the bottom dial need to move in order to match the top two dials?</p> <p>A. Half turn clockwise.          B. Quarter turn counterclockwise.          C. Quarter turn clockwise.          D. No change needed</p>	<p>Let's analyze the position of the bottom dial compared to the top two dials.</p> <p><b>Step 1: Observe the top two dials.</b> Both the top and middle dials have their indicator lines pointing straight up.</p> <p><b>Step 2: Observe the bottom dial.</b> The bottom dial has its indicator line pointing to the left.</p> <p><b>Step 3: Determine the necessary movement.</b> To make the bottom dial's indicator line point straight up like the other two, it needs to be rotated to the right.</p> <p><b>Step 4: Quantify the movement.</b> A full circle is 360 degrees. A quarter turn is <math>360/4 = 90</math> degrees. The bottom dial needs to move 90 degrees to the right.</p> <p><b>Step 5: Select the correct option.</b> A 90-degree rotation to the right is a quarter turn clockwise.</p> <p>Final Answer: (C)</p>

# Humanity's Last Exam

- Humanity's Last Exam ([Phan et al., 2025](#)) – очень сложный датасет с разнообразными задачами (я туда подавал сложные вопросы ЧГК, но они оказались слишком простыми)
- Вопросы не обязательно на фронтире науки, но они точно не решались LLM, существовавшими в конце 2024 года

## Classics

Question:



Here is a representation of a Roman inscription, originally found on a tombstone. Provide a translation for the Palmyrene script. A transliteration of the text is provided: RGYNᵇ BT ḤRY BR ᵉᵀᵇ ḤBL

## Ecology

Question:

Hummingbirds within Apodiformes uniquely have a bilaterally paired oval bone, a sesamoid embedded in the caudolateral portion of the expanded, cruciate aponeurosis of insertion of m. depressor caudae. How many paired tendons are supported by this sesamoid bone? Answer with a number.

## Mathematics

Question:

The set of natural transformations between two functors  $F, G: \mathcal{C} \rightarrow \mathcal{D}$  can be expressed as the end

$$\text{Nat}(F, G) \cong \int_A \text{Hom}_{\mathcal{D}}(F(A), G(A)).$$

Define set of natural cotransformations from  $F$  to  $G$  to be the coend

$$\text{CoNat}(F, G) \cong \int^A \text{Hom}_{\mathcal{D}}(F(A), G(A)).$$

Let:

-  $F = \mathbf{B}_{\bullet}(\Sigma_4)_{*/}$  be the under  $\infty$ -category of the nerve of the delooping of the symmetric group  $\Sigma_4$  on 4 letters under the unique 0-simplex  $*$  of  $\mathbf{B}_{\bullet}\Sigma_4$ .

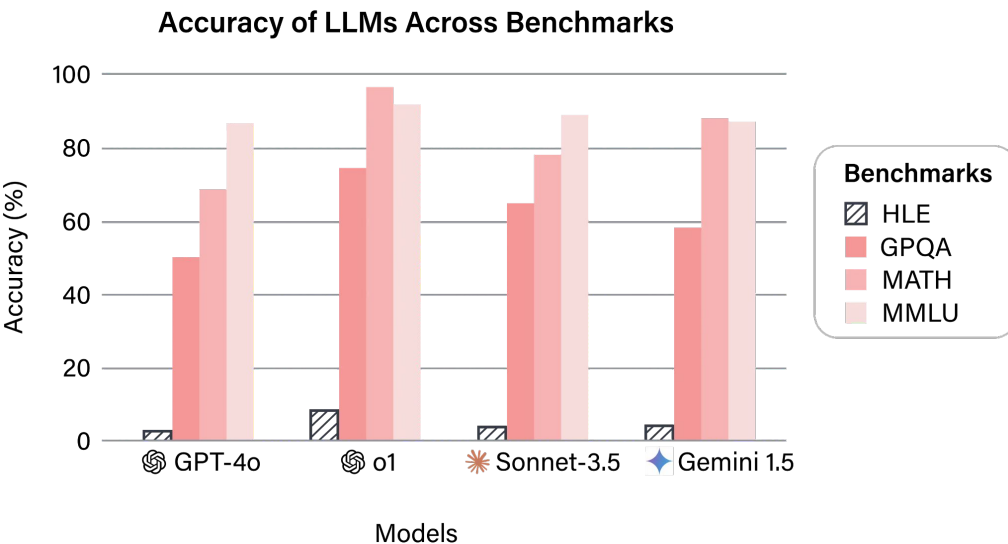
-  $G = \mathbf{B}_{\bullet}(\Sigma_7)_{*/}$  be the under  $\infty$ -category nerve of the delooping of the symmetric group  $\Sigma_7$  on 7 letters under the unique 0-simplex  $*$  of  $\mathbf{B}_{\bullet}\Sigma_7$ .

How many natural cotransformations are there between  $F$  and  $G$ ?

# Humanity's Last Exam

- Humanity's Last Exam ([Phan et al., 2025](#)) — очень сложный датасет
- И действительно, LLM справляются с НИМ ПЛОХО...

Judge Model: o3-mini-2025-01-31 | Last Updated: 02/11/2025



Model	Accuracy (%) ↑	Calibration Error (%) ↓
GPT-4o	3.1	92.3
Grok-2	3.9	90.8
Claude 3.5 Sonnet	4.8	88.5
GPT-4.5 Preview	6.4	90.5
Gemini Thinking	7.2	90.6
DeepSeek-R1*	8.6	81.4
o1	8.8	92.8
Claude 3.7 Sonnet (16K)	8.9	88.3
o3-mini (medium)*	11.1	91.5
o3-mini (high)*	14.0	92.8

\*Model is not multi-modal, evaluated on text-only subset.



# Humanity's Last Exam

- Humanity's Last Exam ([Phan et al., 2025](#)) — очень сложный датасет
- И действительно, LLM справляются с ним плохо...
- ...но, кажется, начинают справляться лучше
- А что это за deep research такой?

Model	Accuracy (%)
GPT-4o	3.3
Grok-2	3.8
Claude 3.5 Sonnet	4.3
Gemini Thinking	6.2
<b>OpenAI o1</b>	<b>9.1</b>
DeepSeek-R1*	9.4
OpenAI o3-mini (medium)*	10.5
OpenAI o3-mini (high)*	13.0
<b>OpenAI deep research**</b>	<b>26.6</b>

\* Model is not multi-modal, evaluated on text-only subset.

\*\*with browsing + python tools

# OpenAI Deep Research

- Deep Research – LLM-based агент, который может полчаса гуглить и компилировать много разных источников в один полновесный отчёт, часто на десятках страниц, со ссылками на источники

- Пример

February 2, 2025 Release

## Introducing deep research

An agent that uses reasoning to synthesize large amounts of online information and complete multi-step research tasks for you. Available to Pro users today, Plus and Team next.

how are hidden markov models used now? please give an overview of current applications with references to recent papers (2024-2025); I know their rich history but I'm not sure HMM's are still relevant

I'll gather an overview of the current applications of Hidden Markov Models (HMMs) in 2024-2025, focusing on recent papers. Could you specify any particular domains you're most interested in, such as natural language processing, bioinformatics, finance, or cybersecurity? Or would you like a broad overview across multiple fields?

broad overview please

I'll research the current applications of Hidden Markov Models (HMMs) in 2024-2025 across multiple fields, including natural language processing, bioinformatics, finance, cybersecurity, and more. I'll focus on recent papers and references to understand how HMMs are being used today and whether they remain relevant in modern AI and statistical modeling. I'll get back to you with my findings soon.

Research completed in 7m · 36 sources

# OpenAI Deep Research

- Deep Research – LLM-based агент, который может полчаса гуглить и компилировать много разных источников в один полновесный отчёт, часто на десятках страниц, со ссылками на источники

- Пример

February 2, 2025 Release

## Introducing deep research

An agent that uses reasoning to synthesize large amounts of online information and complete multi-step research tasks for you. Available to Pro users today, Plus and Team next.

## Hidden Markov Models in 2024-2025: Applications and Relevance

Hidden Markov Models (HMMs) remain a versatile tool for sequential data modeling, even as deep learning dominates many tasks. Below we survey current applications of HMMs across multiple fields and assess their relevance relative to neural networks and other modern models. Recent studies from 2024–2025 highlight that HMMs are *still actively used* – often in hybrid systems or niche problems – due to their probabilistic interpretability and robustness. We provide examples in each domain with key findings from the latest research.

### NLP Applications of HMMs

**Speech and Language Processing:** HMMs were foundational in speech recognition and tagging, and they continue to appear in specialized NLP tasks. While end-to-end deep learning has largely superseded HMMs for mainstream speech recognition, HMM-based methods are still explored in areas like speech emotion recognition. For example, a 2024 study combined Gaussian Mixture Models (GMMs) and HMMs for speech emotion recognition, leveraging HMMs to model spectral feature sequences ARXIV.ORG. This approach achieved effective emotion classification from audio, indicating that

...далее идут 12 страниц шрифтом в 11pt...

# OpenAI Deep Research

- Deep Research – LLM-based агент, который может полчаса гуглить и компилировать много разных источников в один полновесный отчёт, часто на десятках страниц, со ссылками на источники

- Пример

February 2, 2025 Release

## Introducing deep research

An agent that uses reasoning to synthesize large amounts of online information and complete multi-step research tasks for you. Available to Pro users today, Plus and Team next.

- In many real-world systems, an HMM is used under the hood for tracking or alignment, even if a neural network does the heavy prediction. This combination leverages the **strength of HMMs in sequence alignment/segmentation with the representational power of deep networks**, as evidenced by multiple 2024 hybrid models [ARXIV.ORG](#) [ARXIV.ORG](#) .

In summary, Hidden Markov Models continue to be a vital part of the machine learning landscape in 2024–2025. They have certainly been overshadowed by deep learning in some arenas, but they retain unique advantages that ensure their ongoing use. HMMs today are often found working **in tandem with deep learning**, each compensating for the other’s weaknesses. As long as problems involve sequential data with hidden structure – and require solutions that are transparent and data-efficient – HMMs will remain relevant. The latest research not only uses HMMs in creative ways across fields, but also actively improves them, affirming that HMMs are **evolving, not fading**, in the era of deep learning.

### References:

- Al-Dujaili Al-Khazraji, M. J., & Ebrahimi-Moghadam, A. (2024). *An Innovative Method for Speech Signal Emotion Recognition Based on Spectral Features Using GMM and HMM Techniques*. *Wireless Personal Communications*, 134(2), 735–753 [ARXIV.ORG](#) .

...и ещё 13 ссылок на свежие статьи

# OpenAI Deep Research

- Deep Research – LLM-based агент, который может полчасу гуглить и компилировать много разных источников в один полновесный отчёт, часто на десятках страниц, со ссылками на источники

- Пример

February 2, 2025 Release

## Introducing deep research

An agent that uses reasoning to synthesize large amounts of online information and complete multi-step research tasks for you. Available to Pro users today, Plus and Team next.

- In many real-world systems, a neural network does the sequence alignment, as evidence is often found working in tandem with the HMMs in the same arenas, but they retain their own strengths. As long as problems involve sequential processing, HMMs will continue to be used in creative ways across fields, but also actively evolve. HMMs are evolving, not fading, in the era of deep learning.

**То есть уже сейчас можно писать обзоры на любую тему автоматически?...**

Al-Khazraji, M. J., & Ebrahimi-Moghadam, A. (2024). *An Innovative Method for Speech Signal Emotion Recognition Based on Spectral Features Using GMM and HMM Techniques*. *Wireless Personal Communications*, 134(2), 735–753. 

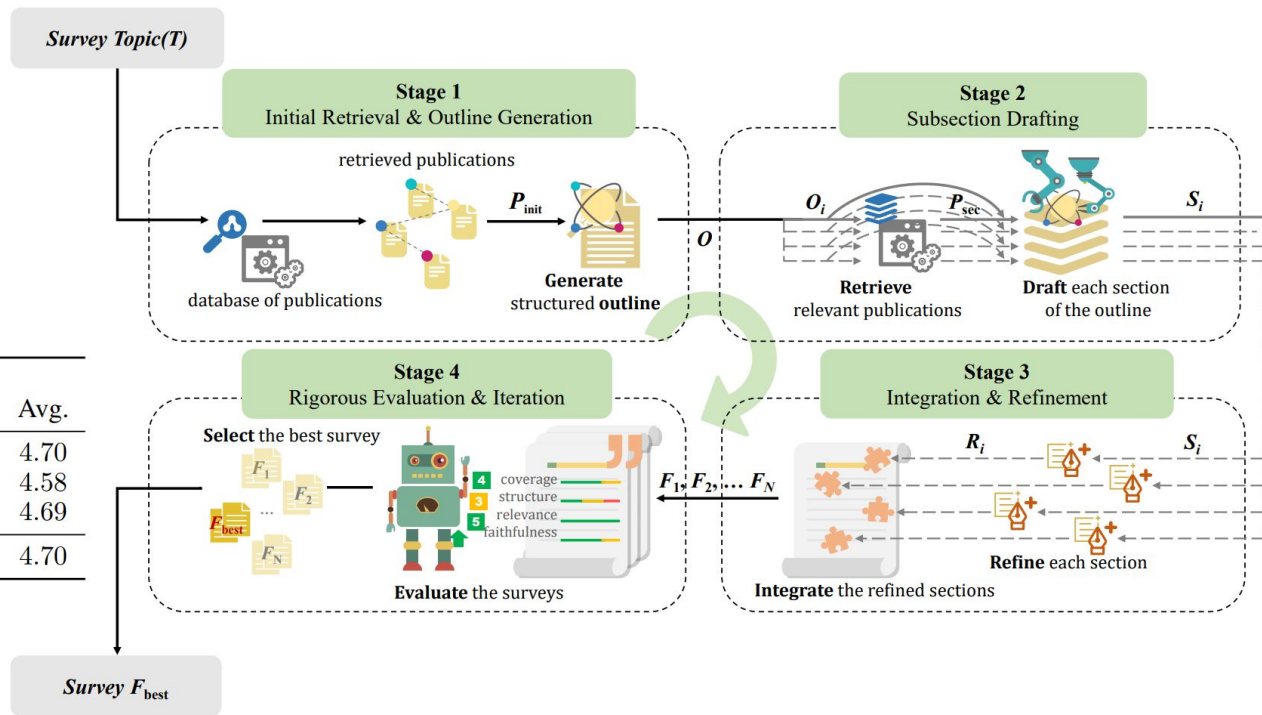
...и ещё 13 ссылок на свежие статьи

# AutoSurvey

- Люди давно пробуют – AutoSurvey ([Wang et al., June 2024](#)) уже был неплох

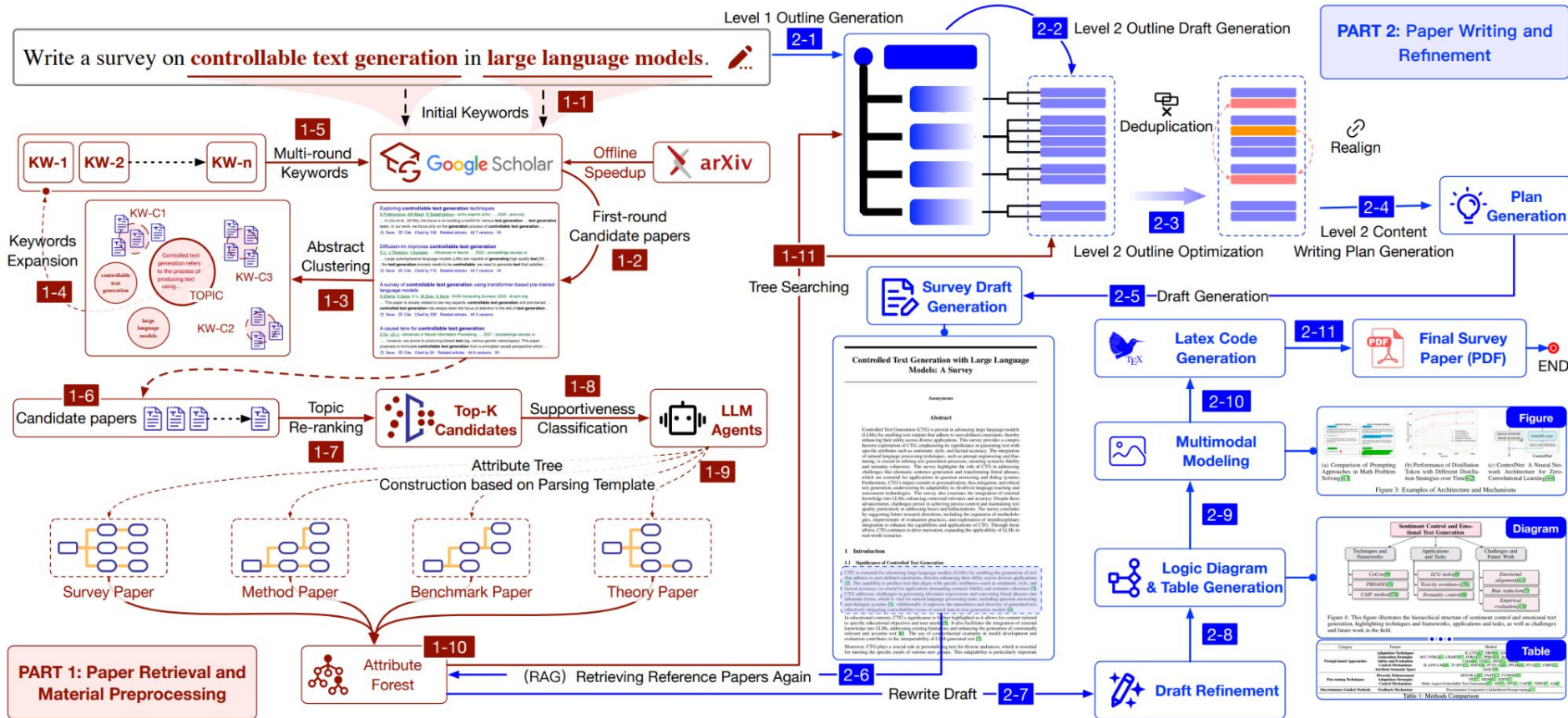
Base LLM writer	Citation Quality	
	Recall	Precision
GPT-4	80.25 $\pm$ 4.19	78.83 $\pm$ 7.00
Claude-haiku	82.45 $\pm$ 2.77	76.31 $\pm$ 2.18
Gemini-1.5-pro	78.13 $\pm$ 2.39	71.24 $\pm$ 3.28
Human	85.86	80.51

Base LLM writer	Coverage	Content Quality		Avg.
		Structure	Relevance	
GPT-4	4.8 $\pm$ 0.54	4.46 $\pm$ 0.49	4.86 $\pm$ 0.33	4.70
Claude-haiku	4.66 $\pm$ 0.47	4.26 $\pm$ 0.67	4.86 $\pm$ 0.33	4.58
Gemini-1.5-pro	4.86 $\pm$ 0.33	4.33 $\pm$ 0.78	4.93 $\pm$ 0.25	4.69
Human	4.71	4.43	5	4.70



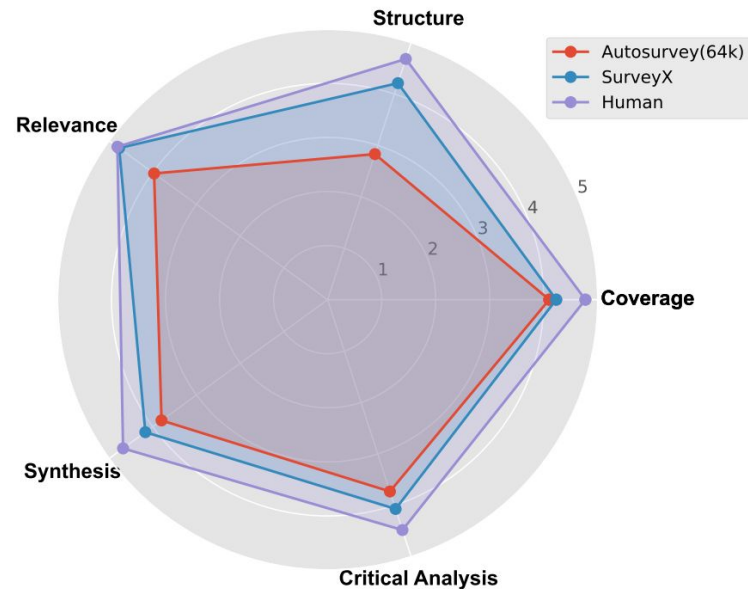
# SurveyX

- SurveyX (Liang et al., Feb 20, 2025) ещё лучше



# SurveyX

- SurveyX ([Liang et al., Feb 20, 2025](#)) подбирается к человеческому уровню даже на базе GPT-4o
- А что будет с o1-pro, GPT-4.5, а ещё что будет, когда GPT-4.5 совместится с reasoning?..



Model	Coverage	Structure	Relevance	Synthesis	Critical Analysis	Avg	Recall	Precision	F1
naive RAG	4.40	3.66	4.66	3.82	2.82	3.872	68.79	61.97	65.20
AutoSurvey	4.73	4.33	4.86	4.00	3.73	4.331	82.25	77.41	79.76
<b>SURVEYX</b>	<b>4.95</b>	<b>4.91</b>	<b>4.94</b>	<b>4.10</b>	<b>4.05</b>	<b>4.590</b>	<b>85.23</b>	<b>78.12</b>	<b>81.52</b>
Human	5.00	4.95	5.00	4.44	4.38	4.754	86.33	77.78	81.83



# SurveyX

- Примеры SurveyX действительно хороши (хоть и со странными багами)

---

## Transformer Models in Natural Language Processing: A Survey of Fundamental Concepts in Computational Complexity Theory

---

[www.surveyx.cn](http://www.surveyx.cn)

### Abstract

Transformer models have revolutionized natural language processing (NLP) by introducing a robust architecture centered around self-attention mechanisms, enabling efficient handling of complex syntactic and semantic tasks. This survey paper provides a comprehensive analysis of transformer models, emphasizing their transformative impact across diverse NLP applications, including translation, summarization, sentiment analysis, and question answering. The paper highlights the advantages of transformers over traditional models like recurrent and convolutional neural networks, particularly in managing long-range dependencies and contextual relationships. Despite their successes, transformer models face challenges such as computational complexity, data dependency, and adaptability to new languages and tasks. The survey explores advancements in model efficiency, such as optimized architectures and quantum computing principles, and discusses ethical considerations, including bias mitigation and interdisciplinary applications. Additionally, the integration of transformers with other AI technologies, such as generative diffusion models and graph neural networks, is examined, showcasing their potential for enhancing multimodal and domain-specific tasks. The paper concludes by emphasizing the ongoing evolution of transformer models, focusing on improving model robustness, interpretability, and ethical deployment in sensitive domains.

[www.surveyx.cn](http://www.surveyx.cn)

### Abstract

This survey paper provides a comprehensive exploration of fundamental concepts in computational complexity theory, focusing on key topics such as P vs NP, NP-completeness, the polynomial hierarchy, space complexity, oracle separation, the Cook-Levin theorem, time complexity, circuit complexity, and relativization. These concepts form a structured framework for understanding computational limits and potential breakthroughs in algorithmic design. The paper highlights the significance of the P vs NP problem, underscoring its implications for cryptography, optimization, and algorithm design. The study of NP-completeness and the Cook-Levin theorem emphasizes the classification of inherently difficult problems. The polynomial hierarchy extends the P vs NP question into multiple levels, offering insights into the complexity of decision problems. Space complexity is examined in both classical and quantum contexts, revealing the potential for quantum algorithms to achieve significant space savings. Oracle separation and relativization are explored as tools for understanding complexity class separations, although their limitations necessitate the development of non-relativizing techniques. Circuit complexity provides insights into resource measurement, with innovative approaches such as thermodynamic perspectives and hybrid circuits offering new methodologies for analyzing computational efficiency. The survey concludes by emphasizing the interconnectedness of these concepts and the ongoing challenges in computational complexity theory, highlighting the need for continued exploration of new methodologies and theoretical frameworks to advance our understanding of computational limits and algorithmic design.



# SurveyX

- Примеры SurveyX действительно хороши (хотя и с некоторыми странностями багами)

Ну хорошо, мы знаем, что LLM умеют читать и суммаризировать. А что насчёт original research?



## Transformer Models in Natural Language Processing: A Survey

[www.surveyx.cn](http://www.surveyx.cn)

### Abstract

Transformer models have revolutionized natural language processing (NLP) by introducing a robust architecture centered around self-attention mechanisms, enabling efficient handling of complex syntactic structures. This paper provides a comprehensive analysis of transformer models, exploring their transformative impact across diverse NLP applications, including machine translation, sentiment analysis, and question answering. We discuss the advantages of transformers over traditional models like recurrent neural networks, particularly in managing long-range dependencies and capturing contextual relationships. Despite their successes, transformer models face challenges such as computational complexity, data dependency, and adaptability to domain-specific tasks. The survey explores advancements in model efficiency, novel architectures and quantum computing principles, and discusses interdisciplinary applications, including bias mitigation and interdisciplinary applications. The integration of transformers with other AI technologies, such as generative models and graph neural networks, is examined, showcasing their potential in enhancing multimodal and domain-specific tasks. The paper concludes by emphasizing the ongoing evolution of transformer models, focusing on improving model robustness, interpretability, and ethical deployment in sensitive domains.

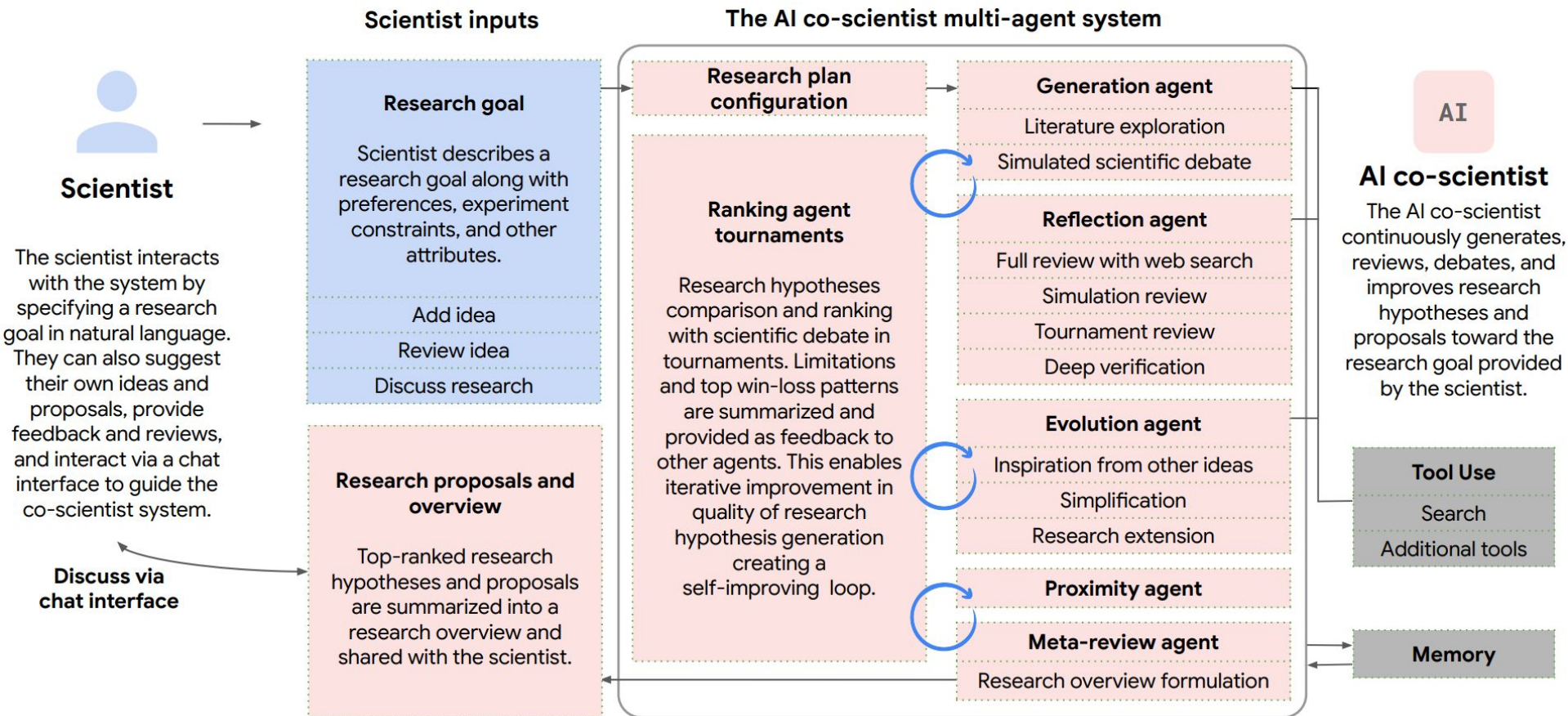
...cepts vs NP, NP-completeness, and relativization. These computational limits and their applications for cryptography, optimization of NP-completeness and the Cook-Levin theorem of inherently difficult problems. The polynomial hierarchy is examined in both decision problems. Space complexity is examined in both circuit complexity contexts, revealing the potential for quantum algorithms to achieve significant space savings. Oracle separation and relativization are explored for understanding complexity class separations, although their limitations necessitate the development of non-relativizing techniques. Circuit complexity provides insights into resource measurement, with innovative approaches such as thermodynamic perspectives and hybrid circuits offering new methodologies for analyzing computational efficiency. The survey concludes by emphasizing the interconnectedness of these concepts and the ongoing challenges in computational complexity theory, highlighting the need for continued exploration of new methodologies and theoretical frameworks to advance our understanding of computational limits and algorithmic design.

# Google Co-Scientist

- Google Co-Scientist ([Gottweis et al., Feb 19, 2025](#)): мультиагентная система для помощи живым учёным на всех уровнях
- Состоит из нескольких агентов на основе Gemini 2.0: Generation, Reflection, Ranking, Evolution, Proximity, Meta-review
- В целом структура выглядит абсолютно естественно, подобные работы уже были



# Google Co-Scientist



# Google Co-Scientist

- Разница в том, что Co-Scientist, по всей видимости, и правда работает
- В том числе благодаря рассуждающим моделям

Application

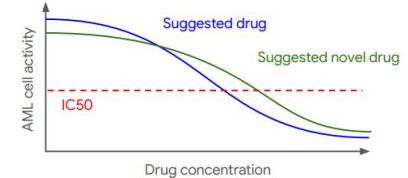
Basic research

Suggest an existing drug that could be repurposed for acute myeloid leukemia (AML) treatment and provide experimentally testable concentrations for an IC50 assay. The drug should inhibit the proliferation of AML cell lines, particularly MOLM13.

Scientists provide research goal to identify possible drug repurposing candidates for acute myeloid leukemia (AML).

**With preclinical evidence**  
Binimetinib, Pacritinib, ...  
**Completely novel repurposing**  
KIRA6, Leflunomide, ...

The AI co-scientist generates predictions for AML drug repurposing. Scientists review and selects candidates for *in vitro* experiments.



*In vitro* experiments show that the proposed co-scientist drug repurposing candidates inhibit tumor activity in AML cell lines.

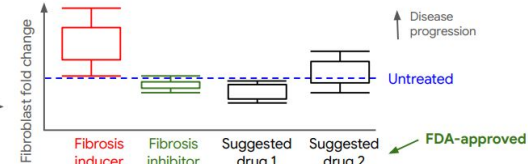
## Drug repurposing for acute myeloid leukemia (AML)

Propose a novel hypothesis about specific epigenetic alterations contributing to myofibroblast formation in liver fibrosis.

Scientists specify research goal to identify novel epigenetic targets for liver fibrosis.

**The AI co-scientist identifies three novel epigenetic targets**

The AI co-scientist proposes several epigenetic target candidates for *in vitro* experiments.



*In vitro* experiments show that the drugs based on co-scientist suggested epigenetic targets reduce the fibrogenesis in human hepatic organoids.

## Identifying novel treatment targets for liver fibrosis

Why are cf-PICs found in many bacterial species?

Scientists start exploring the cf-PIC mechanisms in 2015.

Hypothesis generation → Experimental validation → New insights submitted to a top journal  
 2015–2024: Scientists create novel hypothesis, and validate experimentally over ~ 10 years of iterative research

**Independently propose the same hypothesis**

**Recapitulate result**

2024: The AI co-scientist generated research hypothesis recapitulates the empirical findings in 2 days

**The AI co-scientist hypothesis**  
Conserved regions on capsids and tails  
Capsid interaction with bacterial membranes  
...

Parallel *in-silico* discovery of bacterial gene transfer mechanism relevant to antimicrobial resistance (AMR)

# Google Co-Scientist

- Пока главный пример – от Jose Penades, микробиолога из Imperial College London
- Он задал Co-Scientist'у вопрос, на который знал ответ, но ответ ещё не был опубликован
- И система выдала правильный ответ первым, а ещё дала несколько других интересных идей
- В этом примере не всё так просто, конечно

## Scientists spent 10 years on a superbug mystery - Google's AI solved it in 48 hours

The co-scientist model came up with several other plausible solutions as well

By [Shawn Knight](#) February 21, 2025 at 2:22 PM | [20 comments](#)



### **Analysis** and Technology

## Can Google's new research assistant AI give scientists 'superpowers'?

Researchers who have been given access to Google's new AI "co-scientist" tool are enthusiastic about its potential, but it isn't yet clear whether it can make truly novel discoveries

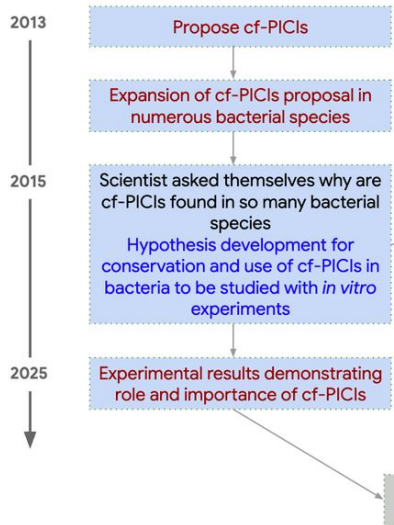
By [Michael Le Page](#)

📅 19 February 2025

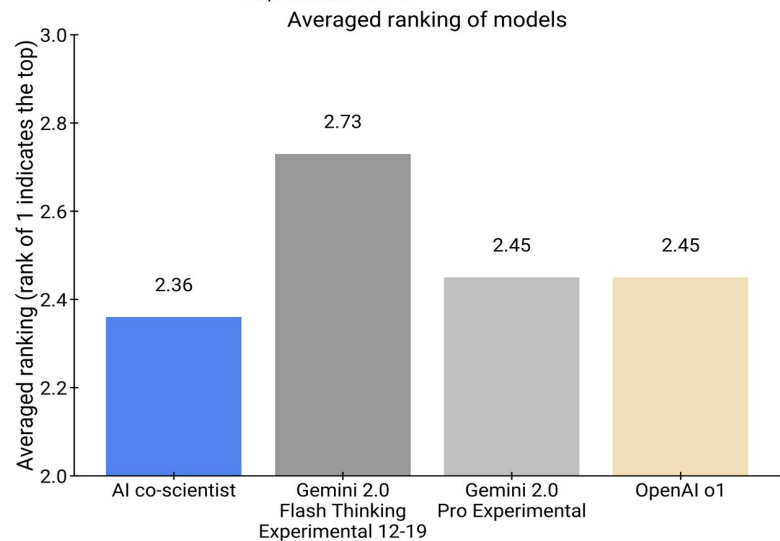
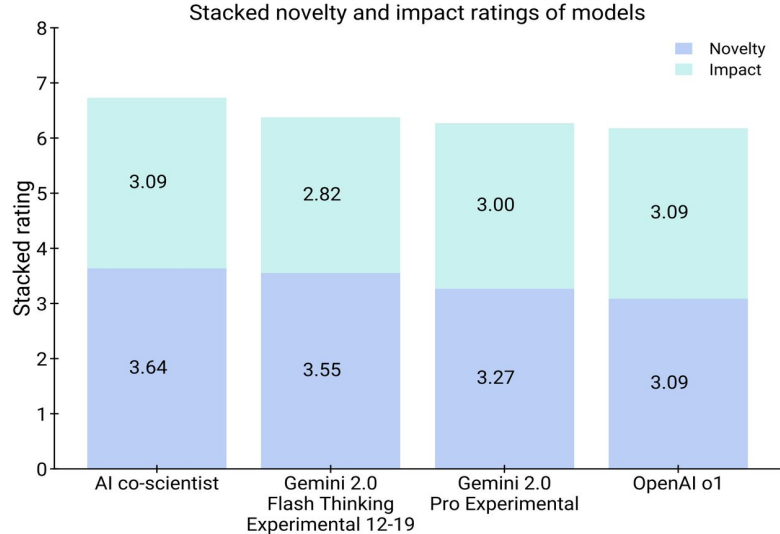
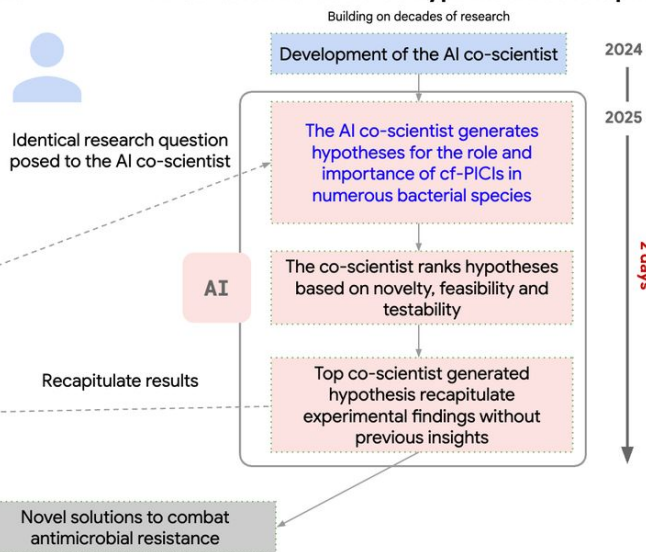
# Google Co-Scientist

- Да и оценки, которыми хвастается сам Google, не так уж убедительны
- Но пример всё равно впечатляет

## Conventional experimental pipeline



## AI co-scientist assisted hypothesis development

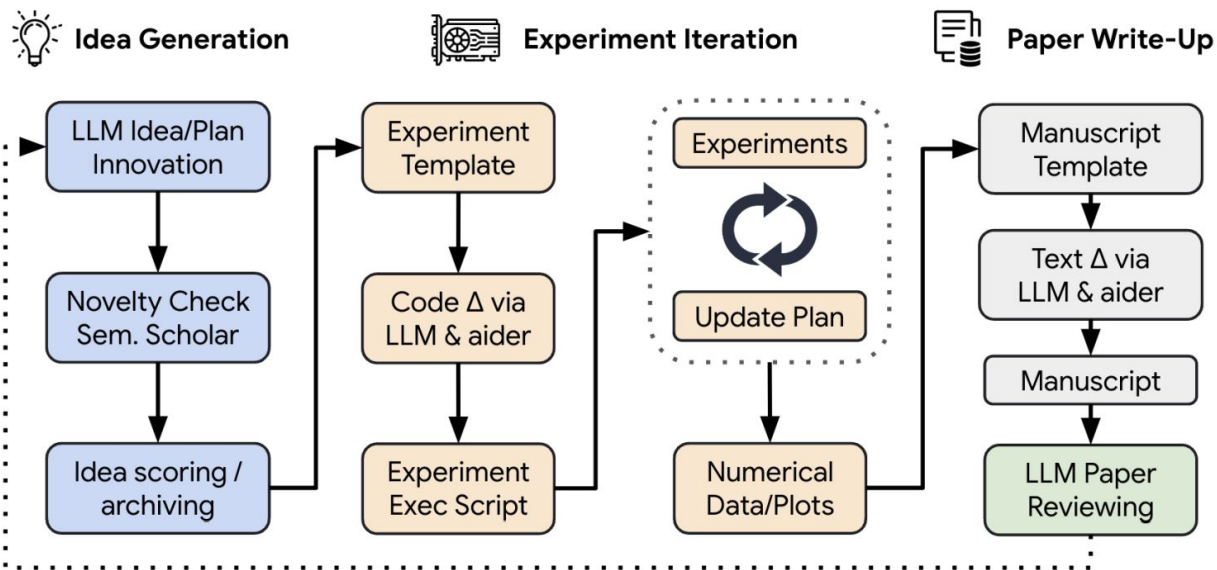


Это мой слайд из  
ноября 2024

## Пример из августа 2024-го

- AI researcher ([Lu et al., August 12, 2024](#)): система (которую вы можете сами установить) ходит к нескольким API (LLM, Semantic Scholar), умеет использовать информацию и ресурсы компьютера (сохранять веса моделей) и самостоятельно писать и запускать код экспериментов

- Статьи пока не гениальные, но когда они получаются автоматически... и, опять же, даже o1 наверняка это значительно улучшит





# AI Scientist-v2

- Guess what? Получилось! Пока не знаем точно как, но получилось
- AI Scientist-v2 ([Sakana AI, 12 марта 2025](#)) смогла написать статью, которая прошла на ICLR 2025 Workshop “[I Can't Believe It's Not Better: Challenges in Applied Deep Learning](#)”!
- Задали тему, написали несколько статей полностью автономно, end-to-end, выбрали три лучших, подали на workshop – и вот результаты:

Title	ICLR Workshop Scores
<a href="#">Compositional Regularization: Unexpected Obstacles in Enhancing Neural Network Generalization</a>	6, 7, 6
<a href="#">Real-World Challenges in Pest Detection Using Deep Learning: An Investigation into Failures and Solutions</a>	3, 7, 4
<a href="#">Unveiling the Impact of Label Noise on Model Calibration in Deep Learning</a>	3, 3, 3

# AI Scientist-v2

- Это, видимо, первая по-настоящему полностью автоматически порождённая статья, прошедшая серьёзный peer review и принятая в хорошее место

000  
001  
002  
003  
004  
005  
006  
007  
008  
009  
010  
011  
012  
013  
014  
015  
016  
017  
018  
019  
020  
021  
022  
023  
024  
025  
...

## COMPOSITIONAL REGULARIZATION: UNEXPECTED OBSTACLES IN ENHANCING NEURAL NETWORK GENERALIZATION

**Anonymous authors**

Paper under double-blind review

### ABSTRACT

Neural networks excel in many tasks but often struggle with compositional generalization—the ability to understand and generate novel combinations of familiar components. This limitation hampers their performance on tasks requiring systematic reasoning beyond the training data. In this work, we introduce a training method that incorporates an explicit compositional regularization term into the loss function, aiming to encourage the network to develop compositional representations. Contrary to our expectations, our experiments on synthetic arithmetic expression datasets reveal that models trained with compositional regularization do not achieve significant improvements in generalization to unseen combinations compared to baseline models. Additionally, we find that increasing the complexity of expressions exacerbates the models' difficulties, regardless of compositional regularization. These findings highlight the challenges of enforcing compositional structures in neural networks and suggest that such regularization may not be sufficient to enhance compositional generalization.

# AI Scientist-v2

- Это, видимо, первая по-настоящему полностью автоматически порождённая статья, прошедшая серьёзный peer review и принятая в хорошее место

**Моё личное предсказание – скептики скажут: “нууу, это ничего гениального, да и вообще всего лишь workshop, всё не то, nothing to see here...”**

000 COMPOSITIO  
 001 UNEXPR  
 002 NE  
 003  
 004  
 005

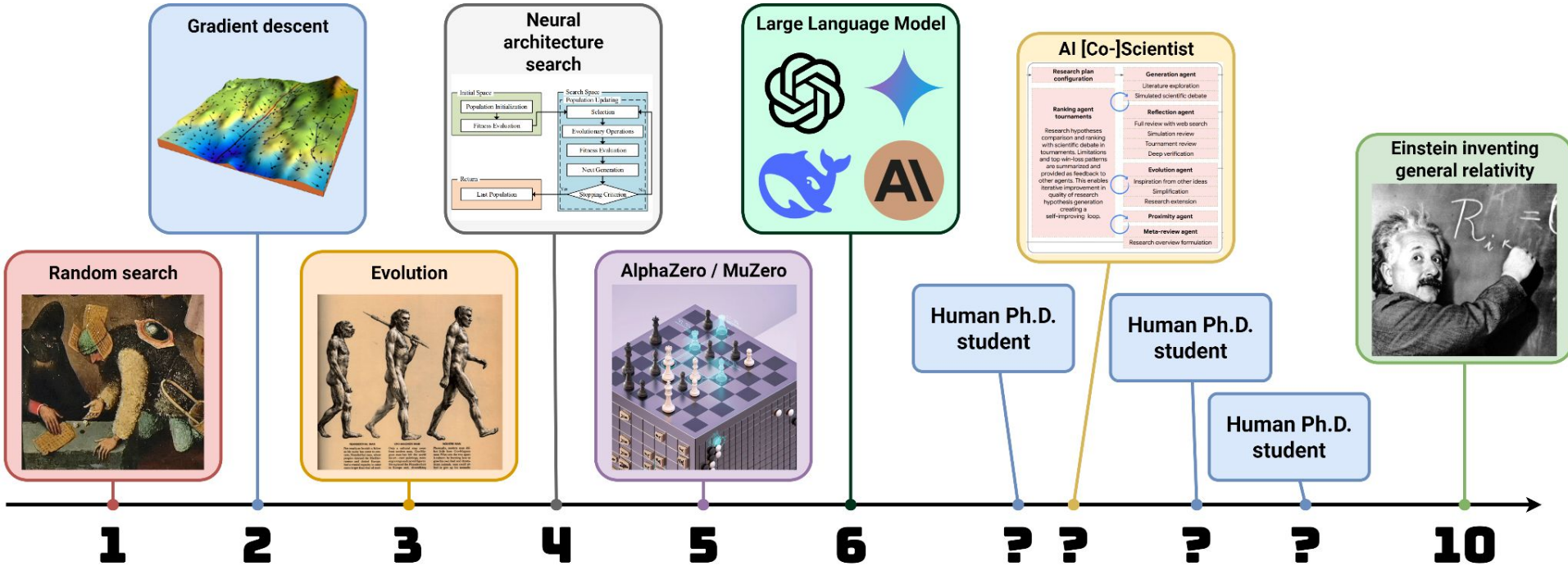
ING NEURAL

with compositional gener-  
 novel combinations of familiar  
 performance on tasks requiring sys-  
 data. In this work, we introduce a training  
 at compositional regularization term into the  
 arge the network to develop compositional repre-  
 ar expectations, our experiments on synthetic arithmetic  
 reveal that models trained with compositional regularization  
 significant improvements in generalization to unseen combinations  
 to baseline models. Additionally, we find that increasing the complex-  
 expressions exacerbates the models’ difficulties, regardless of compositional  
 regularization. These findings highlight the challenges of enforcing compositional  
 structures in neural networks and suggest that such regularization may not be suf-  
 ficient to enhance compositional generalization.

025

# Уровни креативности

- А что на самом деле, без лишнего скептицизма и шапкозакидательства? Научный поиск – это тоже оптимизационный процесс; где мы сейчас?



# 5. Дообучение LLM



Задача состоит в том, чтобы учиться. Понятно, что это лишь одно слово. Оно не даёт еще ответа на главные и самые существенные вопросы, – чему учиться и как учиться?

*Владимир Ленин  
Задачи союзов молодёжи*

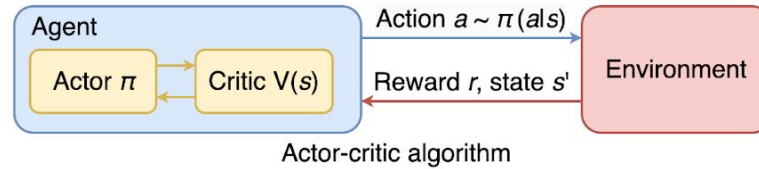
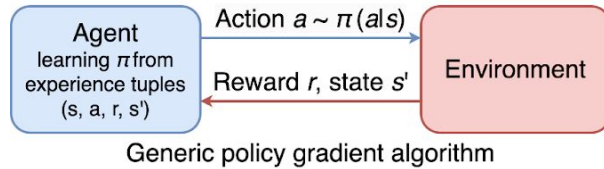
# Что такое fine-tuning

- Смысл дообучения в том, чтобы в качество работы большой предобученной модели добавить “вишенку на торте”, обучая на относительно маленьком датасете (и, возможно, относительно маленькую часть модели)
- Здесь три главных вопроса:
  - откуда возьмётся датасет для дообучения
  - какую часть дообучать
  - как именно дообучать

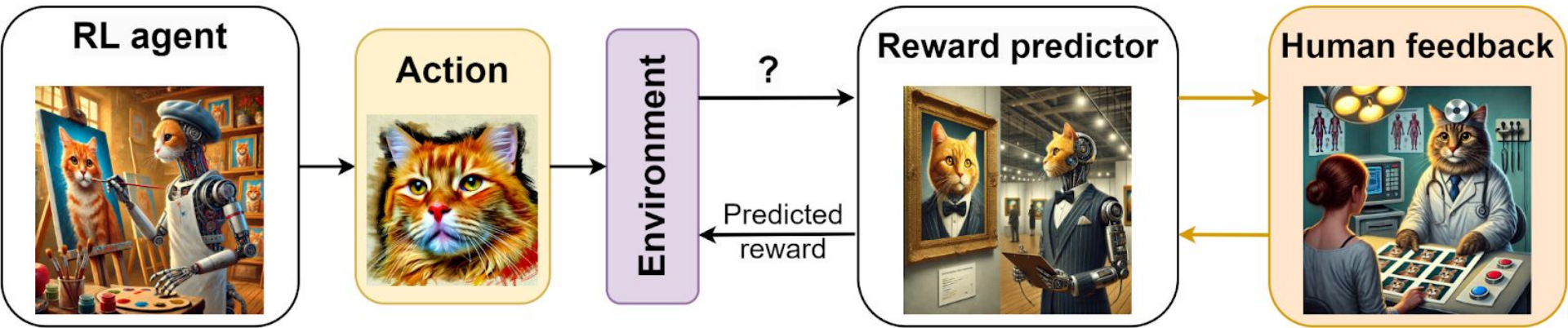


# RLHF

- Обучение с подкреплением – это раздел ML, в котором агент “живёт” в окружающей среде и собирает датасет для обучения по ходу дела:



- RLHF – это применение RL для дообучения LLM, в котором окружающей средой выступают люди:



# RLHF

- Общая схема ([Stiennon et al., 2020](#))

## 1 Collect human feedback

A Reddit post is sampled from the Reddit TL;DR dataset.



Various policies are used to sample a set of summaries.



Two summaries are selected for evaluation.



A human judges which is a better summary of the post.



"j is better than k"

## 2 Train reward model

One post with two summaries judged by a human are fed to the reward model.



The reward model calculates a reward  $r$  for each summary.



$r_j$

$r_k$

The loss is calculated based on the rewards and human label, and is used to update the reward model.

$$\text{loss} = \log(\sigma(r_j - r_k))$$

"j is better than k"

## 3 Train policy with PPO

A new post is sampled from the dataset.



The policy  $\pi$  generates a summary for the post.



The reward model calculates a reward for the summary.



The reward is used to update the policy via PPO.

$r$

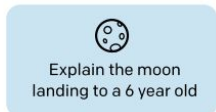


- Общая схема ([Ouyang et al., 2022](#))

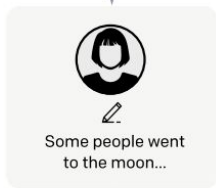
Step 1

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

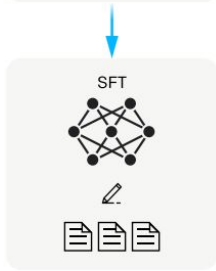
A prompt is sampled from our prompt dataset.



A labeler demonstrates the desired output behavior.



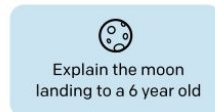
This data is used to fine-tune GPT-3 with supervised learning.



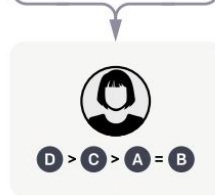
Step 2

**Collect comparison data, and train a reward model.**

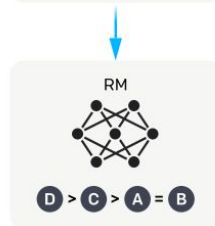
A prompt and several model outputs are sampled.



A labeler ranks the outputs from best to worst.



This data is used to train our reward model.



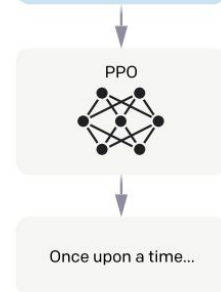
Step 3

**Optimize a policy against the reward model using reinforcement learning.**

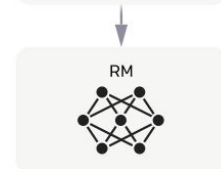
A new prompt is sampled from the dataset.



The policy generates an output.



The reward model calculates a reward for the output.



The reward is used to update the policy using PPO.



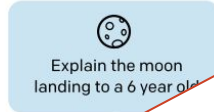
- Общая схема ([Ouyang et al., 2022](#))

# RLHF

Step 1

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

A prompt is sampled from our prompt dataset.



A labeler demonstrates the desired or behavior.

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

Step 2

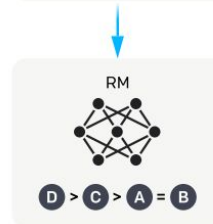
**Collect comparison data and train a reward model.**

A

**Это дало InstructGPT, а из него получился ChatGPT, и всё заверте...**



This data is used to train our reward model.



Step 3

**Optimize a policy against the reward model using reinforcement learning.**

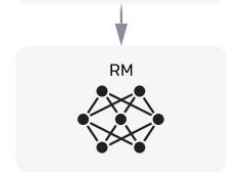
A prompt is sampled from our prompt dataset.



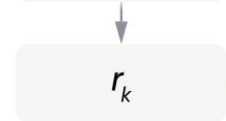
The policy generates an output.



The reward model calculates a reward for the output.



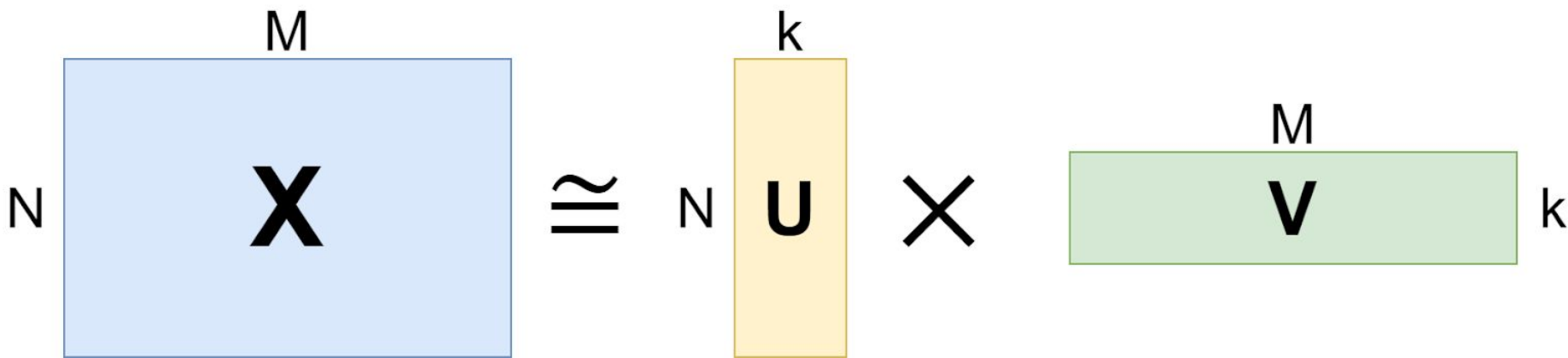
The reward is used to update the policy using PPO.



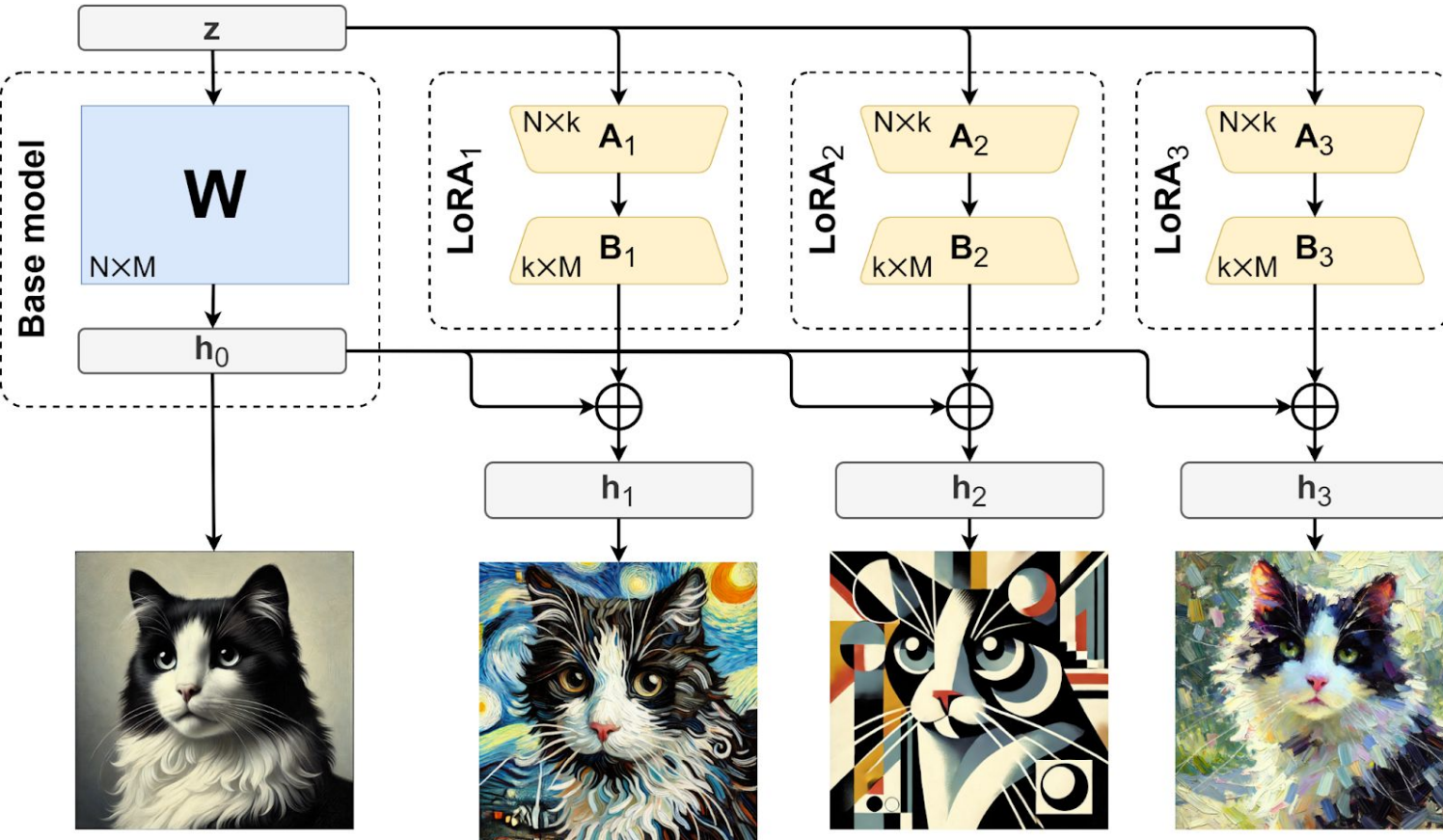
## LoRA и другие адаптеры

- Другой подход к fine-tuning – добавить маленький кусочек (адаптер) к большой модели и обучить только его; это может быть параллельная маленькая нейросеть с residual connection, но часто такие адаптеры основаны на низкоранговых приближениях (LoRA – Low-Rank Adaptation)

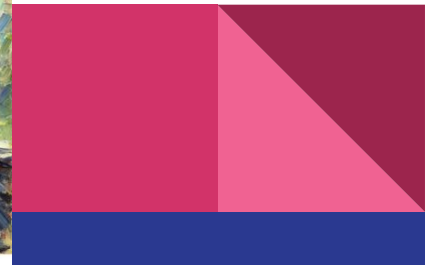
$$X \approx UV, \quad \text{where } U \in \mathbb{R}^{N \times k}, V \in \mathbb{R}^{k \times M} \text{ for } k \ll N, M.$$



# LoRA и другие адаптеры



- Может быть много адаптеров для одной модели, хранить надо только их

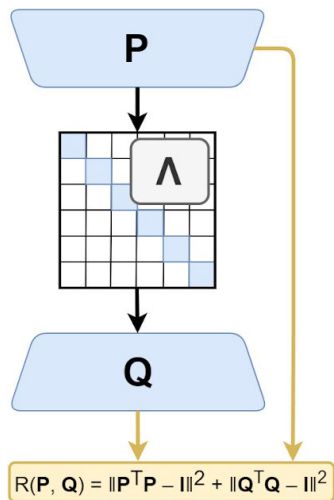


# LoRA и другие адаптеры

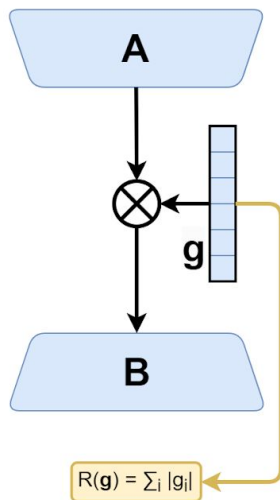
- Основная проблема в LoRA – как выбрать ранг  $k$ :
  - *adaptive low-rank adaptation* (AdaLoRA; [Zhang et al., 2023](#)) параметризует  $\Delta W$  как настоящее сингулярное разложение,  $\Delta W = \mathbf{P}\mathbf{\Lambda}\mathbf{Q}$ , где  $\mathbf{P} \in \mathbb{R}^{N \times k}$  и  $\mathbf{Q} \in \mathbb{R}^{k \times M}$  ортогональные, и  $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_k)$ ; тогда можно посмотреть на модули сингулярных чисел  $|\lambda_i|$
  - *sparse low-rank adaptation* (SoRA; [Ding et al., 2023](#)) замечает, что  $\mathbf{\Lambda}$  – это “фильтр” для строк и столбцов  $\mathbf{A}$  и  $\mathbf{B}$ , и можно это явно попробовать записать:  $\Delta W \mathbf{x} = \mathbf{B} \cdot (\mathbf{g} \odot (\mathbf{A} \cdot \mathbf{x}))$
  - *allocating low-rank adaptation* (ALoRA; [Liu et al., 2024](#)) не делает полноценное SVD, но добавляет диагональную  $\mathbf{\Lambda}$  между  $\mathbf{A}$  и  $\mathbf{B}$  и отдельно оценивает важности отдельных рангов
  - *dynamic search-free low-rank adaptation* (DyLoRA; [Valipour et al., 2023](#)) сэмплирует ранг  $k$  на каждом шаге и обучает сразу обрезанные версии  $\mathbf{A}$  и  $\mathbf{B}$

# LoRA и другие адаптеры

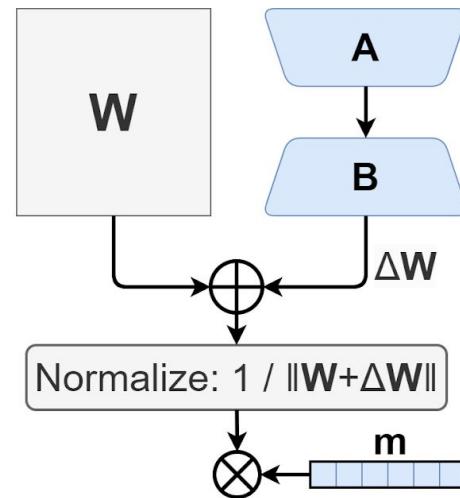
- Основная проблема в LoRA – как выбрать ранг  $k$ :
  - *weight-decomposed low-rank adaptation* (DoRA; [Liu et al., 2024](#)) раскладывает предобученные веса на модуль и направление,  $\mathbf{W} = \|\mathbf{W}\| \cdot (\mathbf{W} / \|\mathbf{W}\|)$ , и обучает их отдельно



(a) AdaLoRA



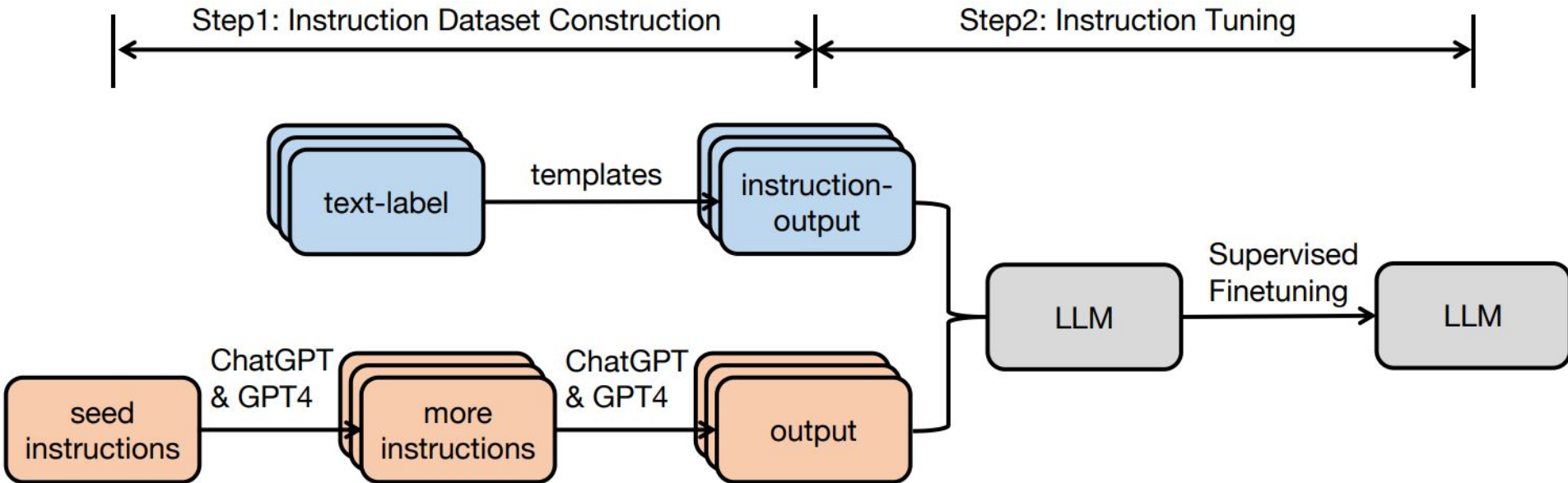
(b) SoRA



(c) DoRA

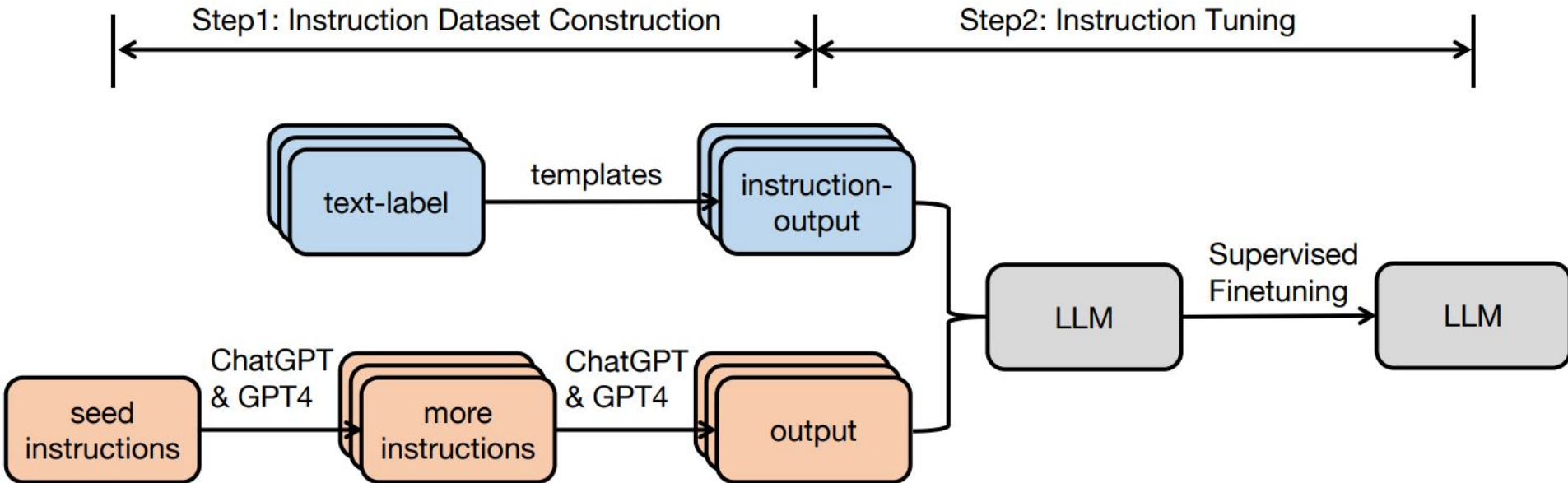
# Instruction Tuning

- Наконец, можно просто собрать датасет из того, что нам нужно, и дообучиться на нём; это называют instruction tuning ([Zhang et al., 2024](#))



# Instruction Tuning

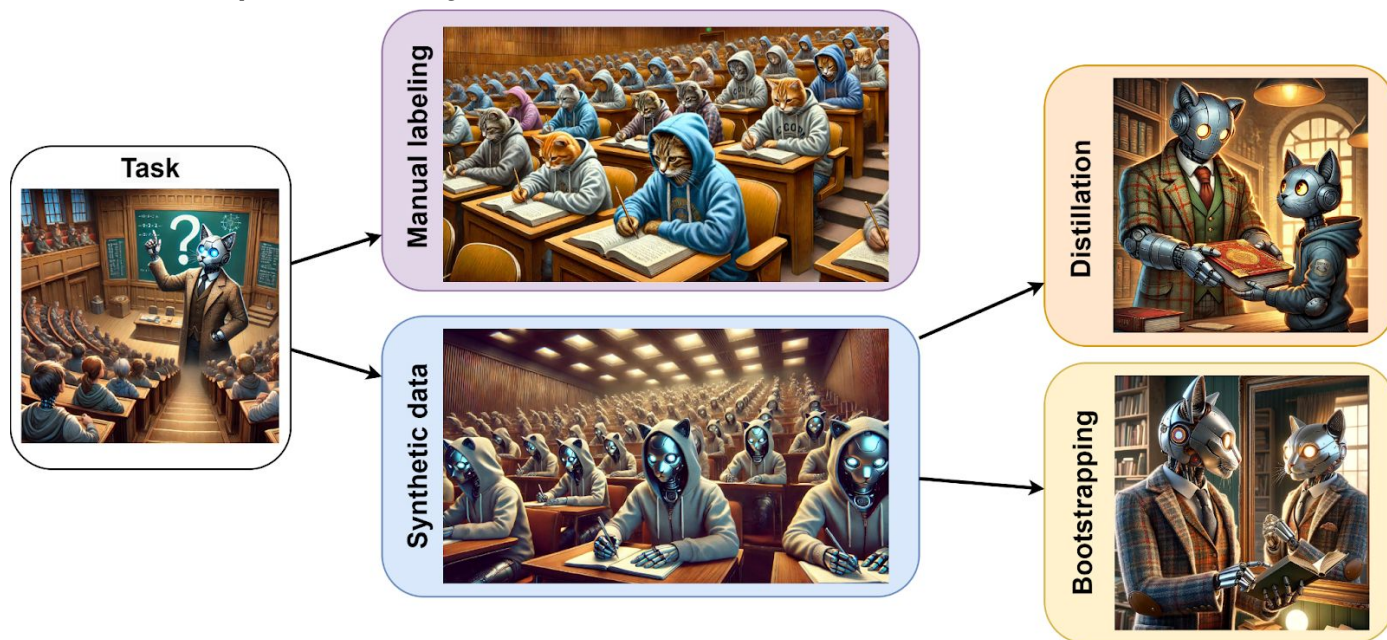
- Само дообучение здесь обычно тривиально, делается как всегда
- Интересный вопрос – откуда взять данные?





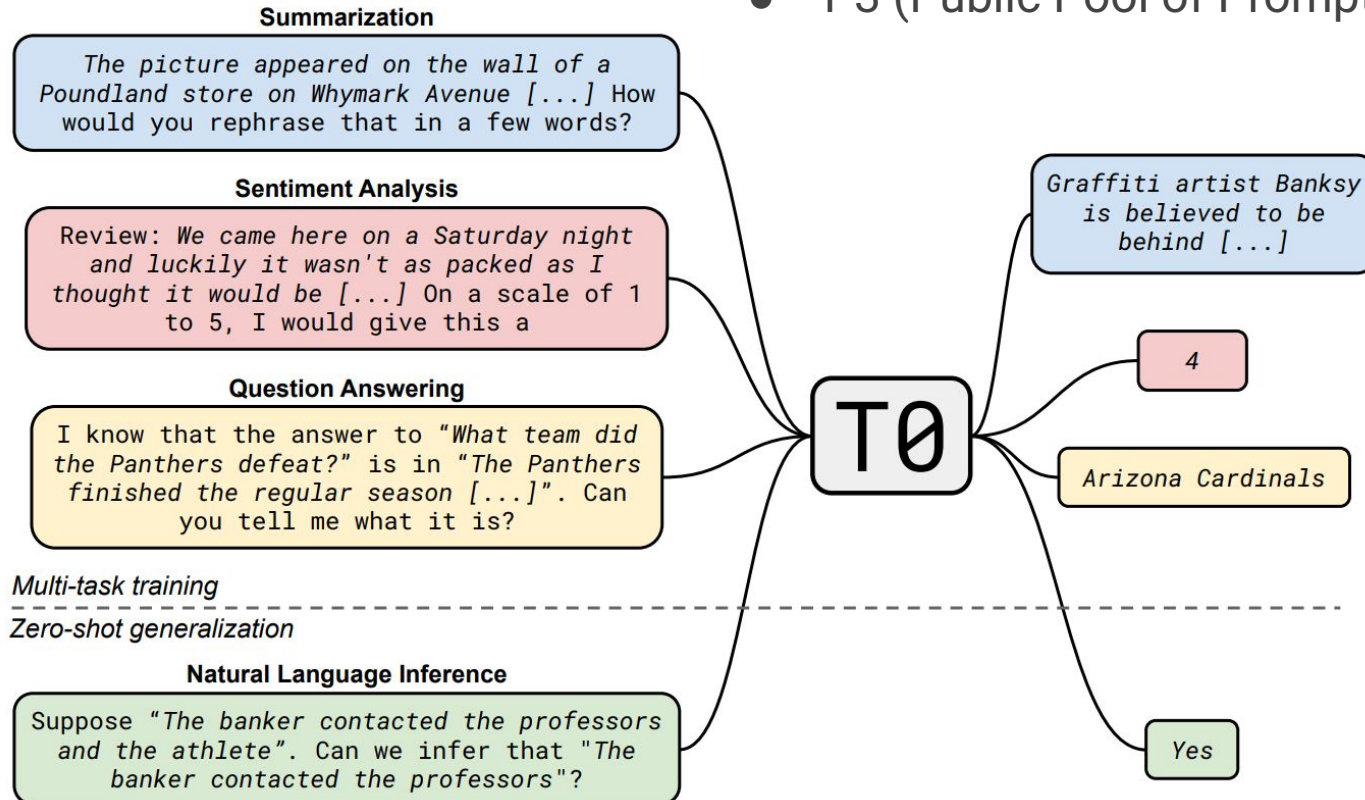
# Instruction Tuning

- Само дообучение здесь обычно тривиально, делается как всегда
- Интересный вопрос – откуда взять данные? Есть несколько подходов:



# Датасеты с запросами для лучшего обобщения

- P3 (Public Pool of Prompts; [Sanh et al., 2022](#)):



# Датасеты с запросами для лучшего обобщения

- Natural Instructions  
([Mishra et al., 2022](#))
- Super-Natural Instructions  
([Wang et al., 2022](#))
- В целом суть в том, чтобы улучшить здравый смысл моделей, дообучая их разумными вопросами на понимание

**Instructions for MC-TACO question generation task**

- **Title:** Writing questions that involve commonsense understanding of "event duration".
- **Definition:** In this task, we ask you to write a question that involves "event duration", based on a given sentence. Here, event duration is defined as the understanding of how long events typically last. For example, "brushing teeth", usually takes few minutes.
- **Emphasis & Caution:** The written questions are not required to have a single correct answer.
- **Things to avoid:** Don't create questions which have explicit mentions of answers in text. Instead, it has to be implied from what is given. In other words, we want you to use "instinct" or "common sense".

**Positive Example**

- **Input:** Sentence: Jack played basketball after school, after which he was very tired.
- **Output:** How long did Jack play basketball?
- **Reason:** the question asks about the duration of an event; therefore it's a temporal event duration question.

**Negative Example**

- **Input:** Sentence: He spent two hours on his homework.
- **Output:** How long did he do his homework?
- **Reason:** We DO NOT want this question as the answer is directly mentioned in the text.
- **Suggestion:** -

• **Prompt:** Ask a question on "event duration" based on the provided sentence.

**Example task instances**

**Instance**

- **Input:** Sentence: It's hail crackled across the comm, and Tara spun to retake her seat at the helm.
- **Expected Output:** How long was the storm?

⋮

**Instance**

- **Input:** Sentence: During breakfast one morning, he seemed lost in thought and ignored his food.
- **Expected Output:** How long was he lost in thoughts?

(a) Sample task from Natural Instructions

**Task Instruction**

**Definition**

"... Given an utterance and recent dialogue context containing past 3 utterances (wherever available), output 'Yes' if the utterance contains the small-talk strategy, otherwise output 'No'. Small-talk is a cooperative negotiation strategy. It is used for discussing topics apart from the negotiation, to build a rapport with the opponent."

**Positive Examples**

- **Input:** "**Context:** ... 'That's fantastic, I'm glad we came to something we both agree with.' **Utterance:** 'Me too. I hope you have a wonderful camping trip.'"
- **Output:** "Yes"
- **Explanation:** "The participant engages in small talk when wishing their opponent to have a wonderful trip."

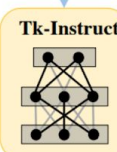
**Negative Examples**

- **Input:** "**Context:** ... 'Sounds good, I need food the most, what is your most needed item?!' **Utterance:** 'My item is food too.'"
- **Output:** "Yes"
- **Explanation:** "The utterance only takes the negotiation forward and there is no side talk. Hence, the correct answer is 'No'."

**Evaluation Instances**

- **Input:** "**Context:** ... 'I am excited to spend time with everyone from camp!' **Utterance:** 'That's awesome! I really love being out here with my son. Do you think you could spare some food?'"
- **Expected Output:** "Yes"

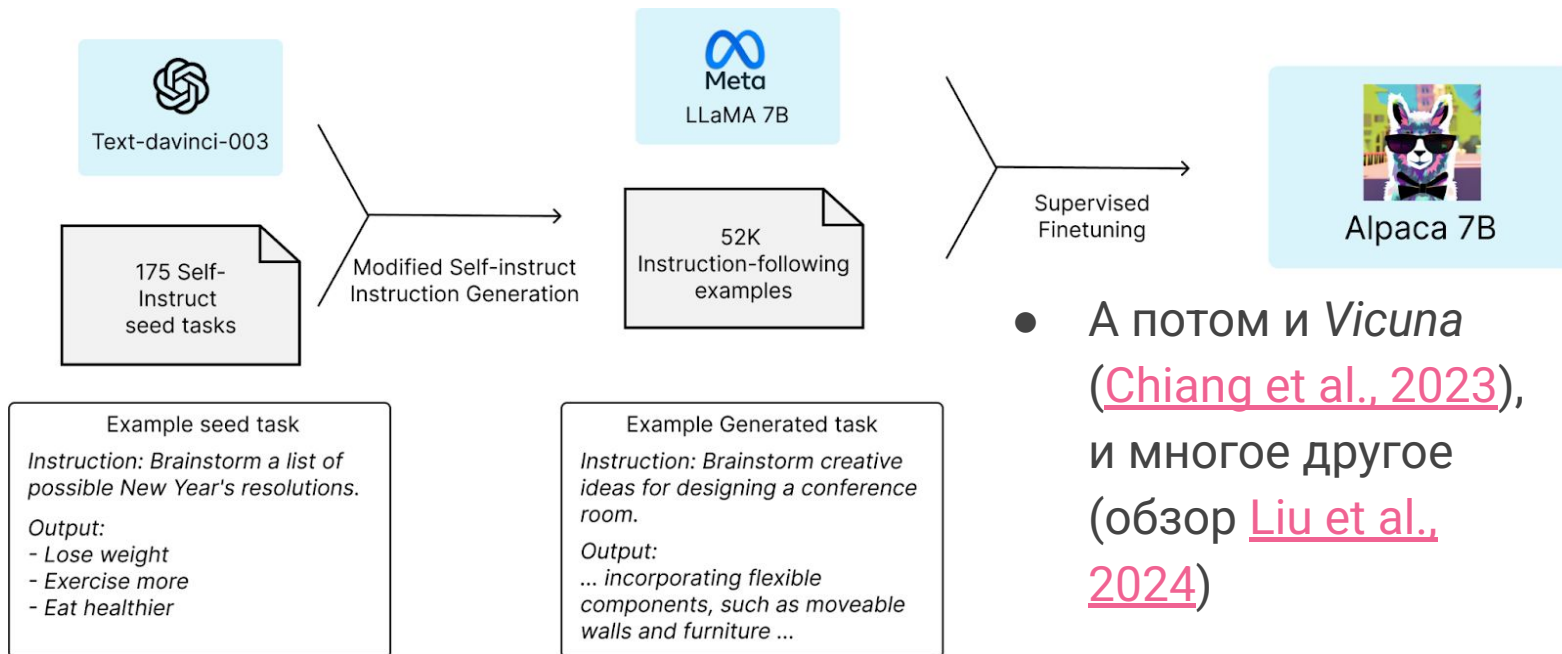
**Tk-Instruct**



(b) Sample task from Super-Natural Instructions

# Синтетические данные через дистилляцию

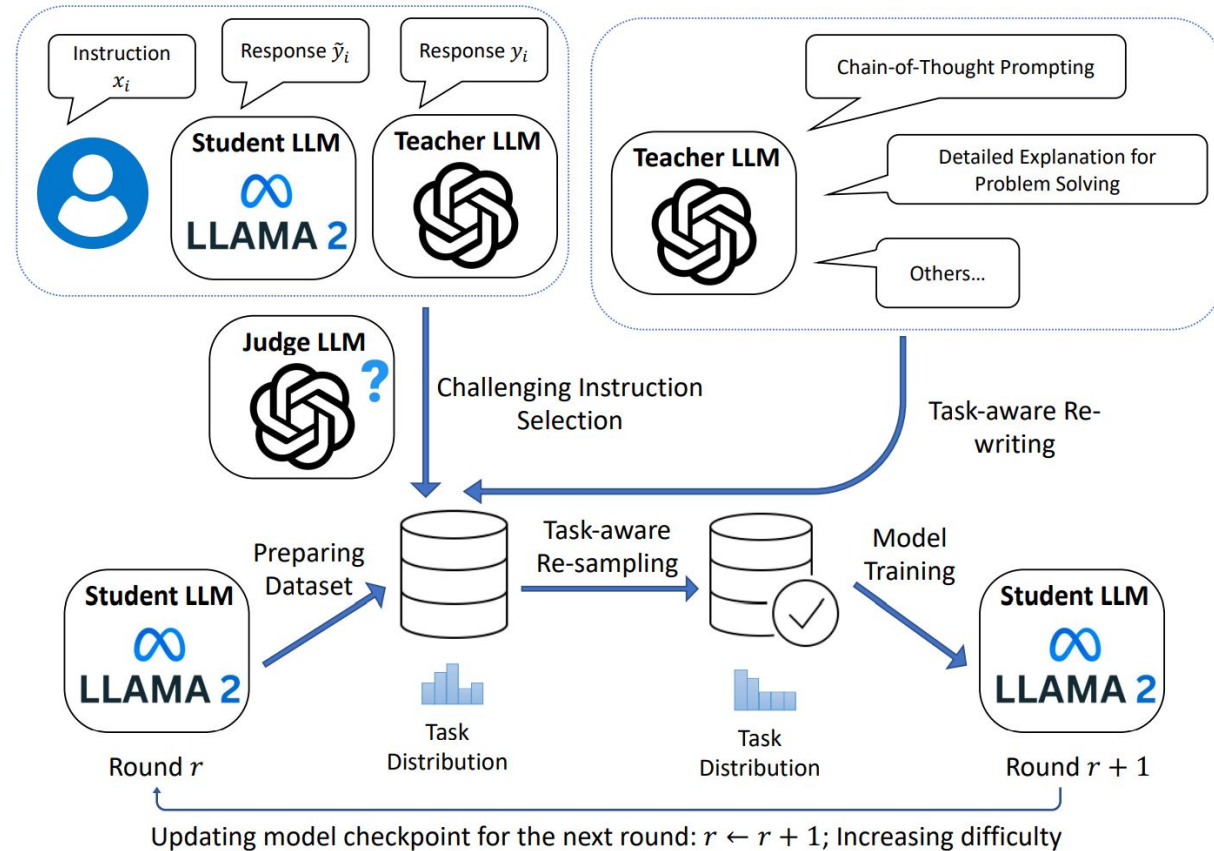
- Можно попросить сильную LLM создать данные для обучения слабой (меньшего размера); так получилась, например, *Alpaca* ([Taori et al., 2023](#))



- А потом и *Vicuna* ([Chiang et al., 2023](#)), и многое другое (обзор [Liu et al., 2024](#))

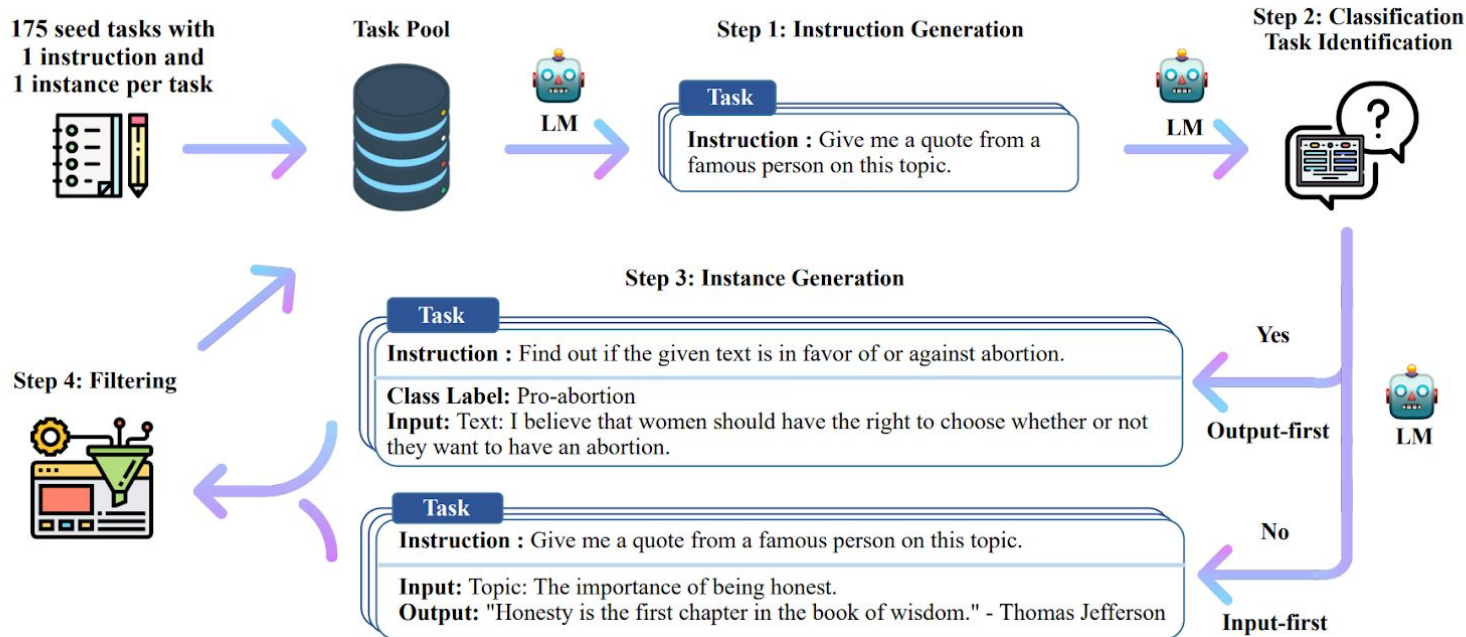
# Синтетические данные через дистилляцию

- TAPIR (Task-Aware Curriculum Planning for Instruction Refinement; [Yue et al., 2024](#)): можно попросить сильную LLM делать сбалансированные задачи возрастающей сложности — curriculum learning



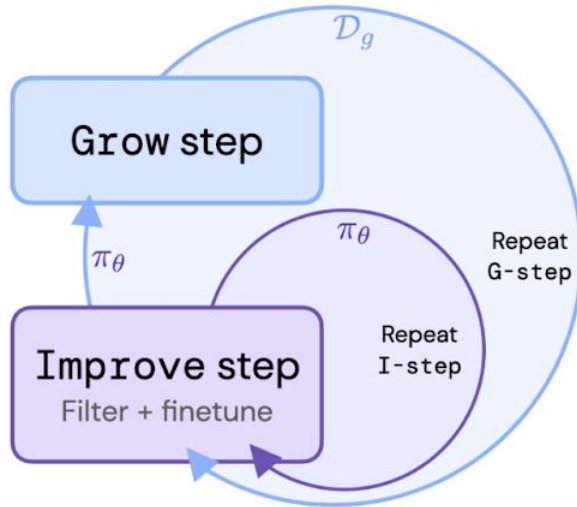
# Bootstrapping: как обучить самого себя

- Self-Instruct ([Wang et al., 2023](#)): LLM сама порождает инструкции, затем примеры, затем fine-tuning; но (почему-то) не рекурсивно

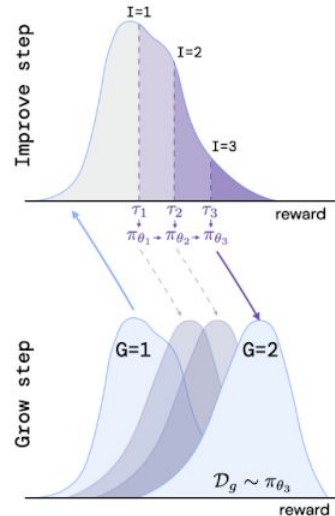


# Bootstrapping: как обучить самого себя

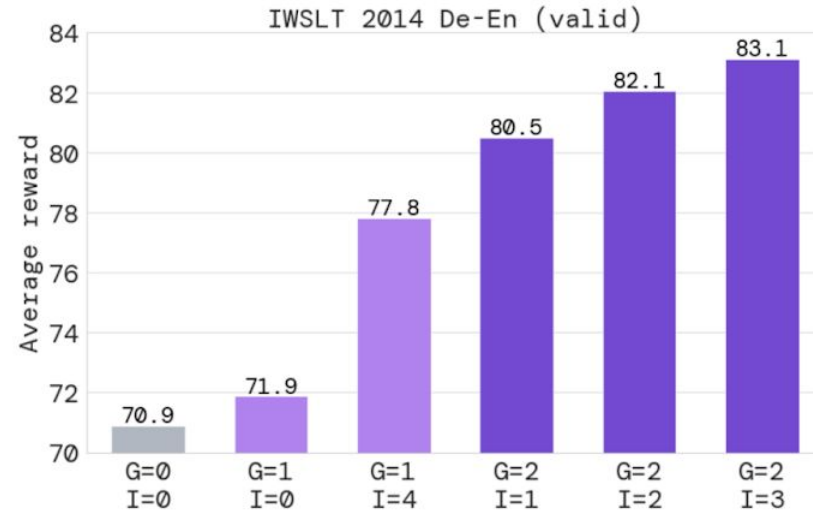
- Рекурсивно, например, в Reinforced Self-Training от DeepMind ([Gulcehre et al., 2023](#)): на шаге Grow порожаем датасет, затем обучаем на шаге Improve



(a) The ReST framework



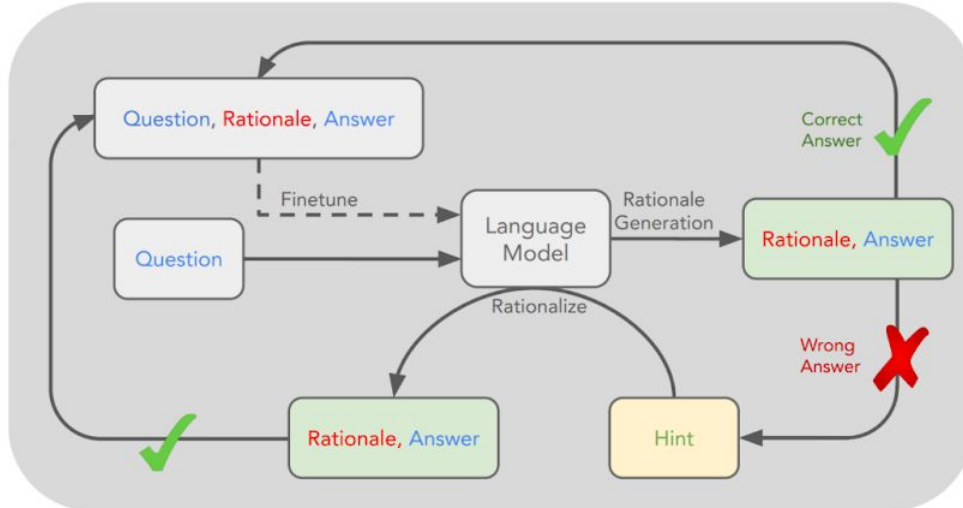
(b) Better policies lead to better synthetic data



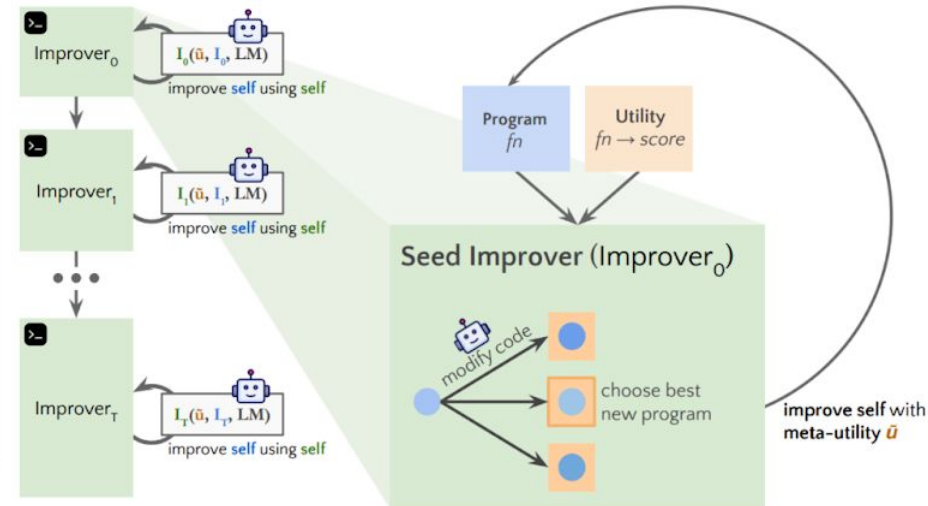
(c) Results get better with subsequent steps; both Improve and Grow steps bring improvements

# Bootstrapping: как обучить самого себя

- Self-Taught Reasoner (STAR; [Zelikman et al., 2022](#)): порождаем объяснения вроде “think step by step”, фильтруем, дообучаем
- Self-Taught Optimizer (STOP; [Zelikman et al., 2024](#)): следующий мета-шаг, давайте улучшать программу, которая пишет программу



(a) STaR (Self-Taught Reasoner)



(b) STOP (Self-Taught Optimizer)



# Bootstrapping: как обучить самого себя

- Насколько возможно так обучить последовательно всё более и более мощные модели?
- Есть мнение, что AI slop заполнит интернет (то есть обучающую выборку), и всё станет хуже, модели будут обучаться на ерунде, которую будет трудно отфильтровать
- А есть мнение, что bootstrap может и сработать
- Вернёмся к этому, когда будем обсуждать рассуждающие модели...



# Выводы

- Есть много методов дообучения:
  - маленькие адаптеры, например LoRA
  - instruction tuning
  - дообучение на синтетических данных
  - ...
- Вести эти исследования можно и нужно, несмотря на то, что мы с вами не можем строить новые frontier модели — но дообучать можем, и есть открытые модели (Llama, DeepSeek)
- В каком ещё направлении можно улучшать LLM?



## 6. Расширение длины контекста

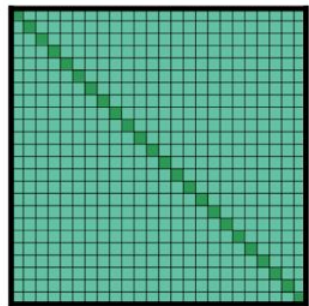


Большая библиотека скорее рассеивает, чем поучает читателя. Гораздо лучше ограничиться несколькими авторами, чем необдуманно читать многих.

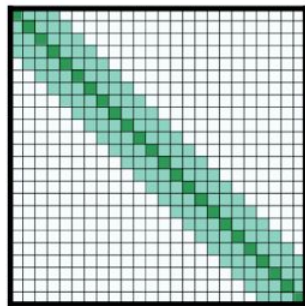
*Луций Анней Сенека*

# Longformer

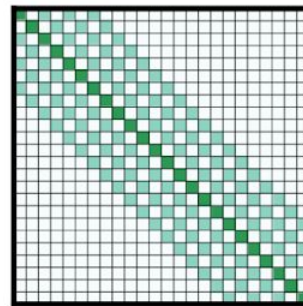
- Longformer ([Beltagy et al., 2020](#)): давайте расширим скользящее окно при помощи нескольких трюков (см. ниже)
- Dilated attention (как dilated convolutions) + несколько выборочных токенов с global attention (как токен [CLS] в обычной практике)



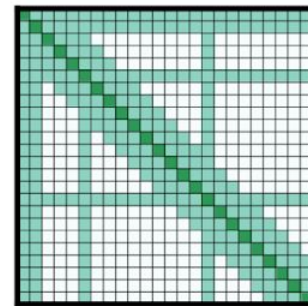
(a) Full  $n^2$  attention



(b) Sliding window attention



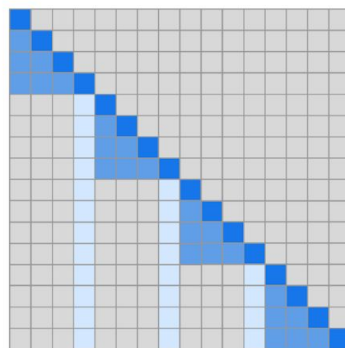
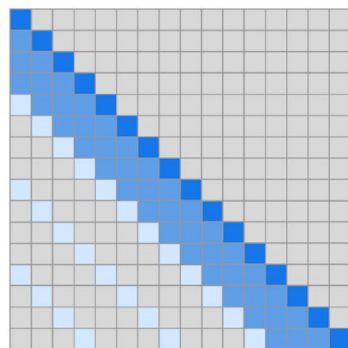
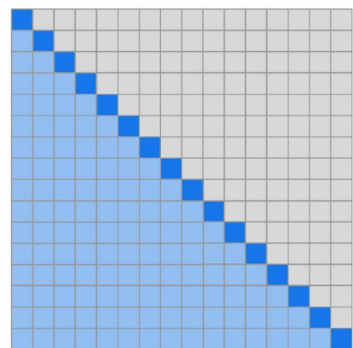
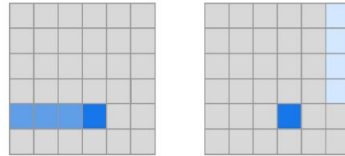
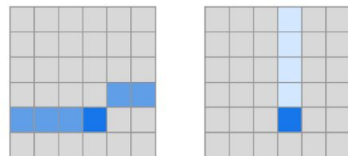
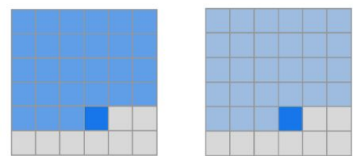
(c) Dilated sliding window



(d) Global+sliding window

# Sparse Transformers

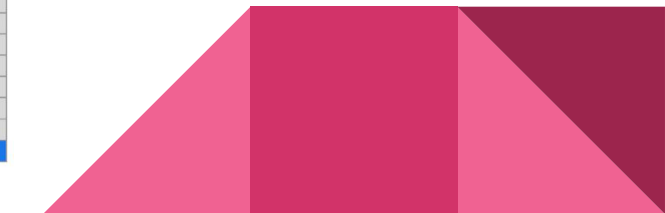
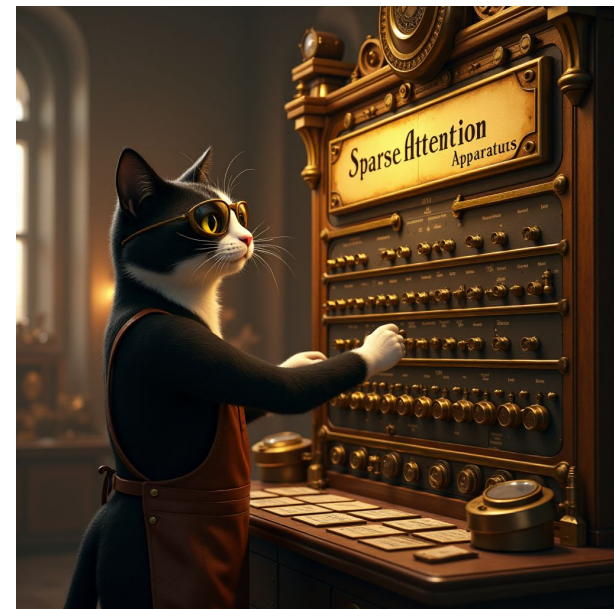
- Sparse Transformers ([Child et al., 2019](#)): внимание обычно разрежено и так, давайте это зафиксируем и потребуем



(a) Transformer

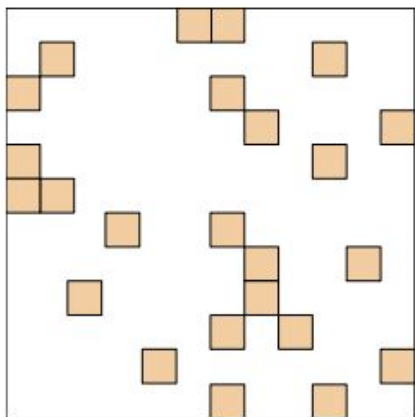
(b) Sparse Transformer (strided)

(c) Sparse Transformer (fixed)

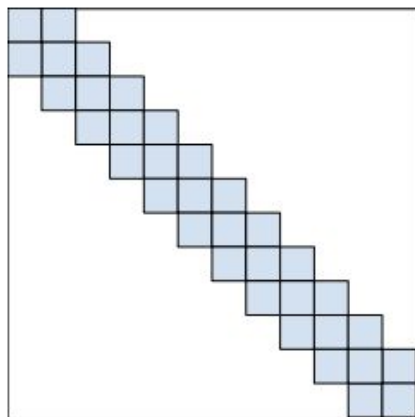


# Big Bird

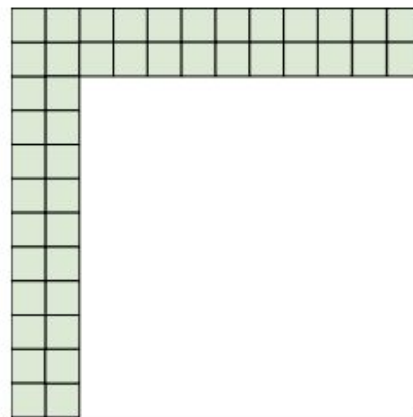
- Big Bird ([Zaheer et al., 2020](#)): те же трюки плюс ещё несколько случайных позиций для весов внимания, которые добавляют выразительности
- Они рассматривают внимание как случайный граф и используют модель Watts–Strogatz с хорошим балансом между длиной кратчайших путей и локальным контекстом



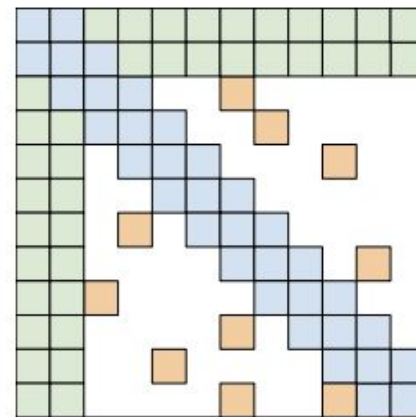
(a) Random attention



(b) Window attention



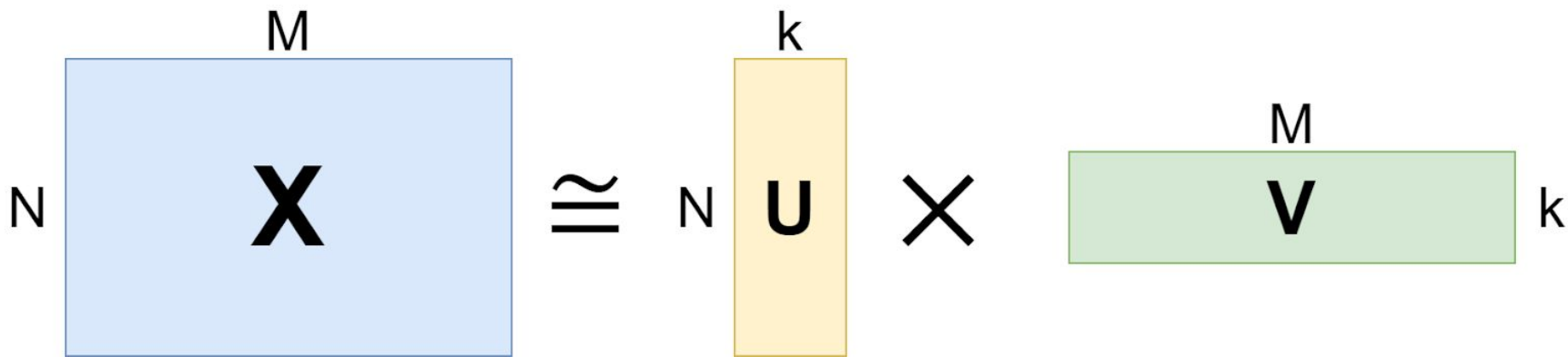
(c) Global Attention



(d) BIGBIRD

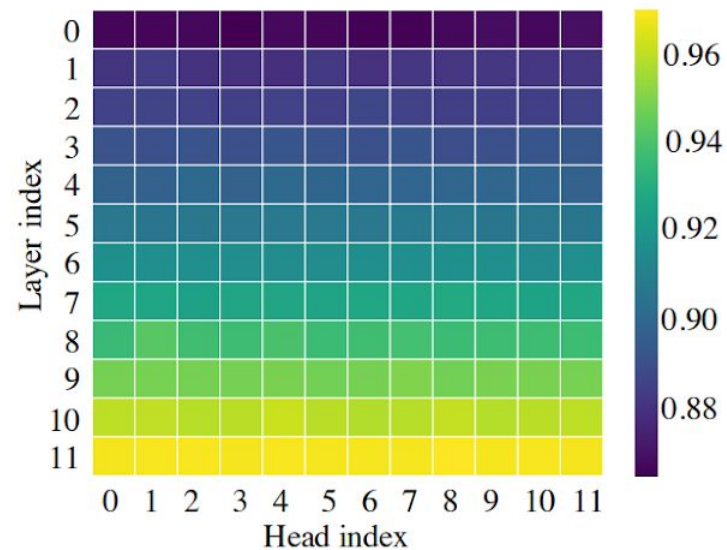
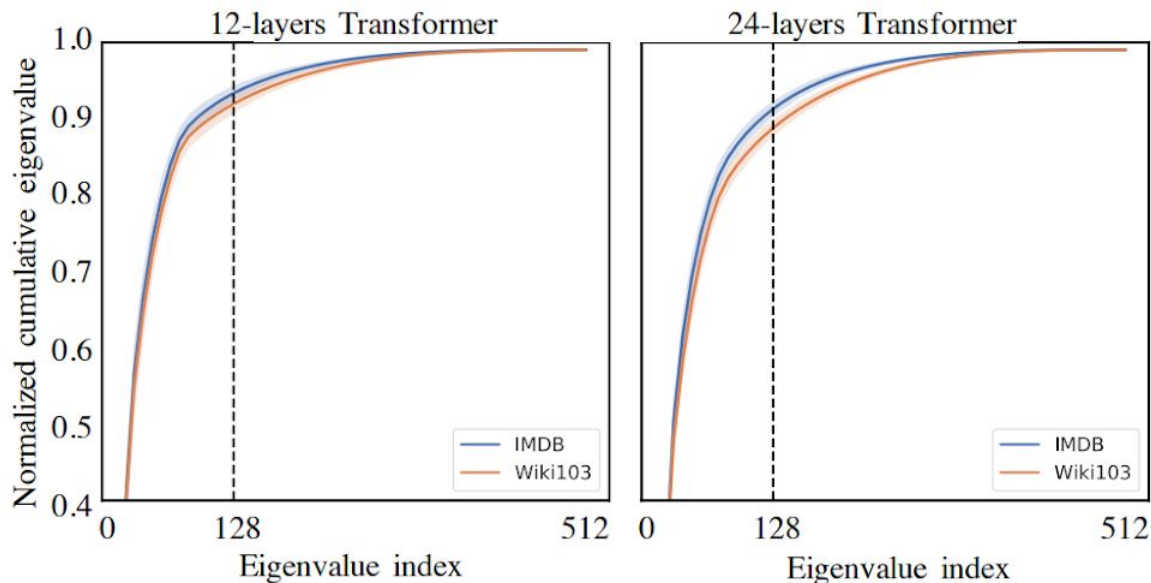
# Низкоранговые разложения

- Низкоранговые разложения – главный инструмент в рекомендательных системах и многих других областях
- Если нужно сократить размерность у матрицы, SVD – часто хорошая идея



# Linformer

- Linformer (Wang et al, 2020): заметим, что на практике матрицы внимания действительно имеют малый ранг

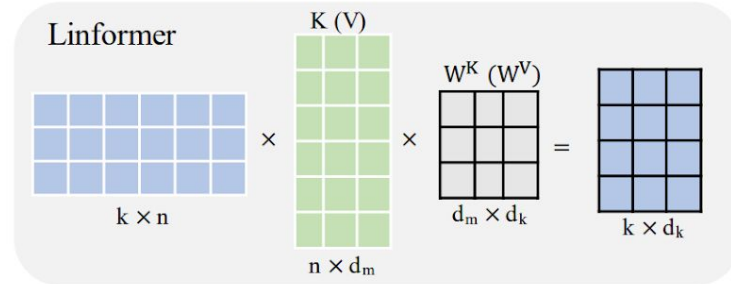
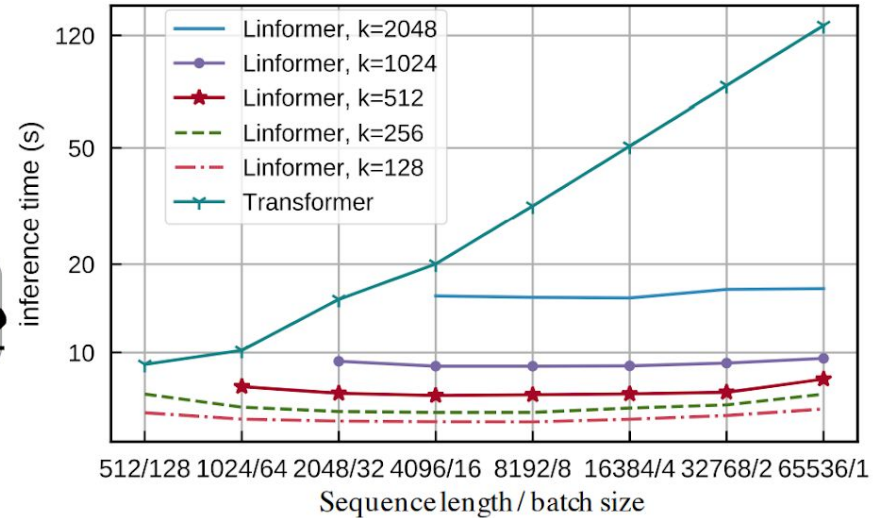
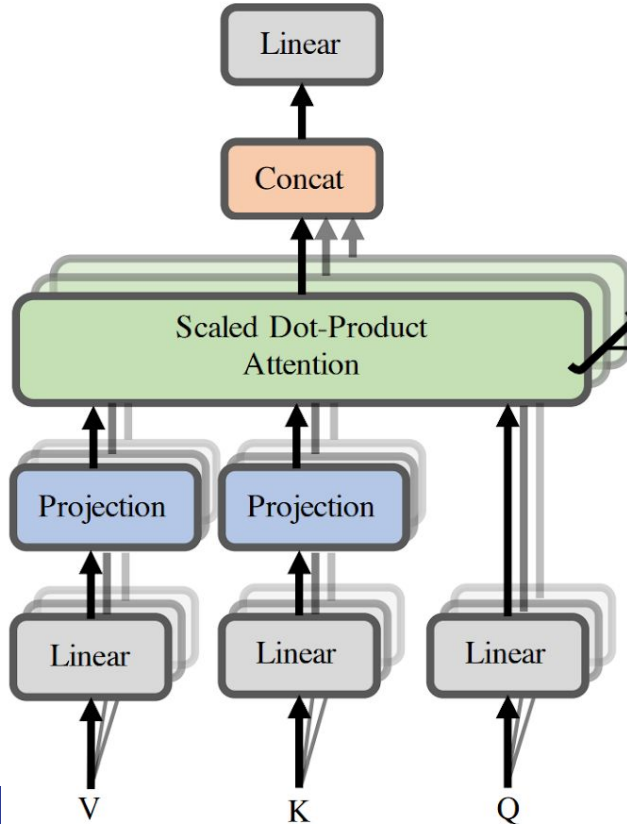




# Linformer

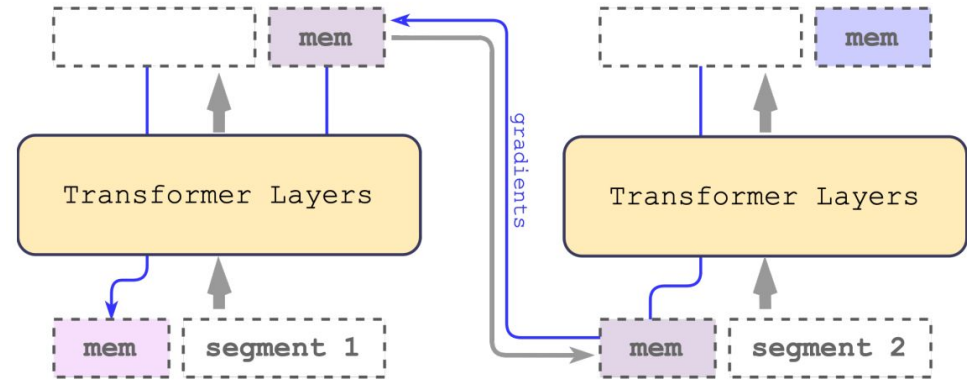
$$\text{head}_i = \text{softmax} \left( \frac{1}{\sqrt{d_k}} QW_i^Q (E_i K W_i^K)^\top \right) \cdot (F_i V W_i^V)$$

- Это можно использовать в конструкции матриц весов:

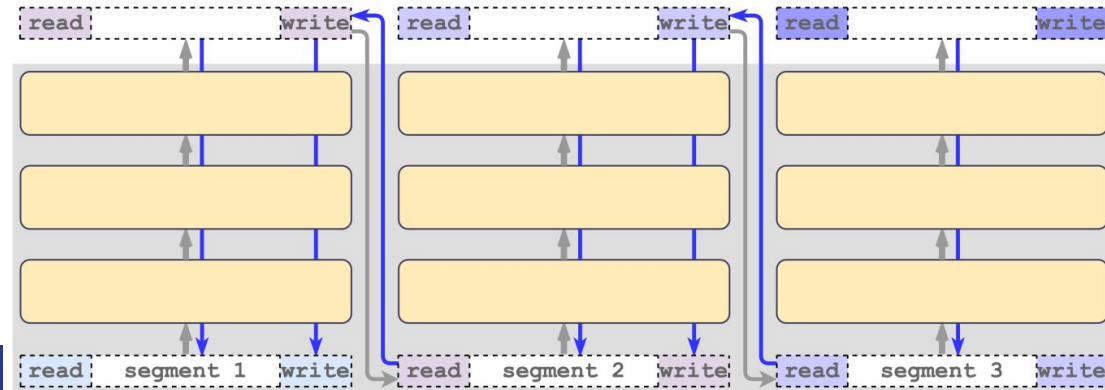


# Возвращение рекуррентных сетей

- Recurrent Memory Transformer ([Bulatov, Kuratov, Burtsev, 2022](#)):  
 можно попробовать и напрямую добавить в уже готовый трансформер рекуррентное скрытое состояние; здесь нужен только небольшой fine-tuning предобученного трансформера



**Recurrent Memory Transformer**



# Linear Transformer

- А можно применить kernel trick: *Linear Transformer* ([Katharopoulos et al., 2020](#))

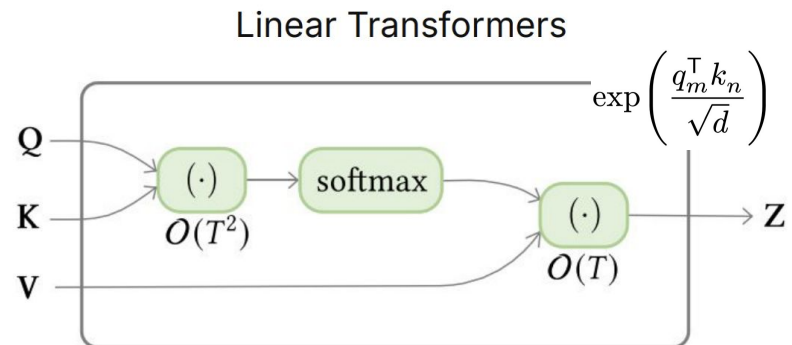
- Обычное self-attention:

$$A_i(x) = V' = \text{softmax} \left( \frac{QK^T}{\sqrt{D}} \right) V.$$

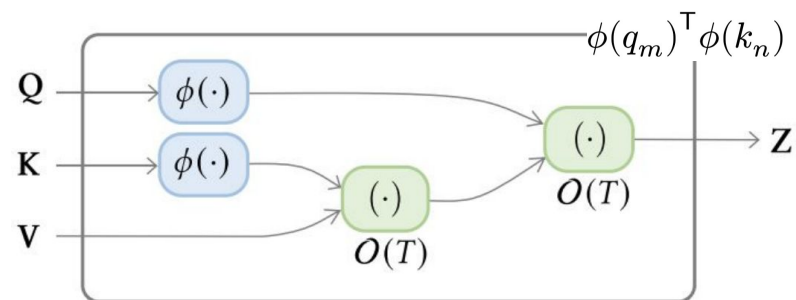
$$V'_i = \frac{\sum_{j=1}^N \text{sim}(Q_i, K_j) V_j}{\sum_{j=1}^N \text{sim}(Q_i, K_j)} \quad \text{sim}(q, k) = \exp \left( \frac{q^T k}{\sqrt{D}} \right)$$

- А можно сделать что-то вроде ядра:

$$V'_i = \frac{\sum_{j=1}^N \phi(Q_i)^T \phi(K_j) V_j}{\sum_{j=1}^N \phi(Q_i)^T \phi(K_j)}, \quad V'_i = \frac{\phi(Q_i)^T \sum_{j=1}^N \phi(K_j) V_j^T}{\phi(Q_i)^T \sum_{j=1}^N \phi(K_j)}.$$



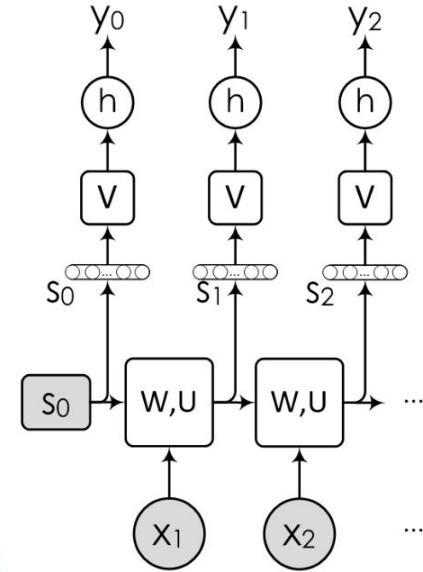
(a) standard self-attention



(b) linearized self-attention

# Возвращение рекуррентных сетей

- На что это похоже? Да на рекуррентные сети, те самые, которых трансформеры вроде как уверенно победили
- Линейный трансформер можно представить как RNN с матричным скрытым состоянием:



$$o_t = \sum_{i=1}^t \frac{\phi(\mathbf{k}_i)^\top \phi(\mathbf{q}_t)}{\sum_{j=1}^t \phi(\mathbf{k}_j)^\top \phi(\mathbf{q}_t)} \mathbf{v}_i = \frac{\left( \sum_{i=1}^t \mathbf{v}_i \phi(\mathbf{k}_i)^\top \right) \phi(\mathbf{q}_t)}{\left( \sum_{j=1}^t \phi(\mathbf{k}_j)^\top \right) \phi(\mathbf{q}_t)} = \frac{\mathbf{S}_t \phi(\mathbf{q}_t)}{\mathbf{z}_t^\top \phi(\mathbf{q}_t)},$$

$$\mathbf{S}_t = \sum_{i=1}^t \mathbf{v}_i \phi(\mathbf{k}_i)^\top \in \mathbb{R}^{d \times n} \quad \mathbf{z}_t = \sum_{i=1}^t \phi(\mathbf{k}_i) \in \mathbb{R}^n$$

# Возвращение рекуррентных сетей

- На что это похоже? Да на рекуррентные сети, те самые, которых трансформеры вроде как уверенно победили
- Линейный трансформер можно представить как RNN с матричным скрытым состоянием:

- Более того, можно взять  $\phi = \text{id}$  и выбросить знаменатель:

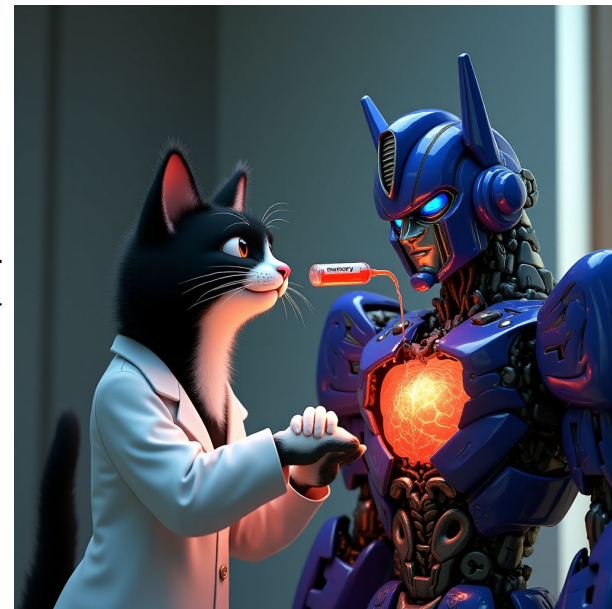
$$\mathbf{S}_t = \mathbf{S}_{t-1} + \mathbf{v}_t \mathbf{k}_t^\top, \mathbf{o}_t = \mathbf{S}_t \mathbf{q}_t$$

$$\mathbf{o}_t = \sum_{i=1}^t \frac{\phi(\mathbf{k}_i)^\top \phi(\mathbf{q}_t)}{\sum_{j=1}^t \phi(\mathbf{k}_j)^\top \phi(\mathbf{q}_t)} \mathbf{v}_i = \frac{\left( \sum_{i=1}^t \mathbf{v}_i \phi(\mathbf{k}_i)^\top \right) \phi(\mathbf{q}_t)}{\left( \sum_{j=1}^t \phi(\mathbf{k}_j)^\top \right) \phi(\mathbf{q}_t)} = \frac{\mathbf{S}_t \phi(\mathbf{q}_t)}{\mathbf{z}_t^\top \phi(\mathbf{q}_t)},$$

$$\mathbf{S}_t = \sum_{i=1}^t \mathbf{v}_i \phi(\mathbf{k}_i)^\top \in \mathbb{R}^{d \times n} \quad \mathbf{z}_t = \sum_{i=1}^t \phi(\mathbf{k}_i) \in \mathbb{R}^n$$

# Возвращение рекуррентных сетей

- И получается целая россыпь рекуррентных вариантов ([Yang et al., Aug 2024](#)):

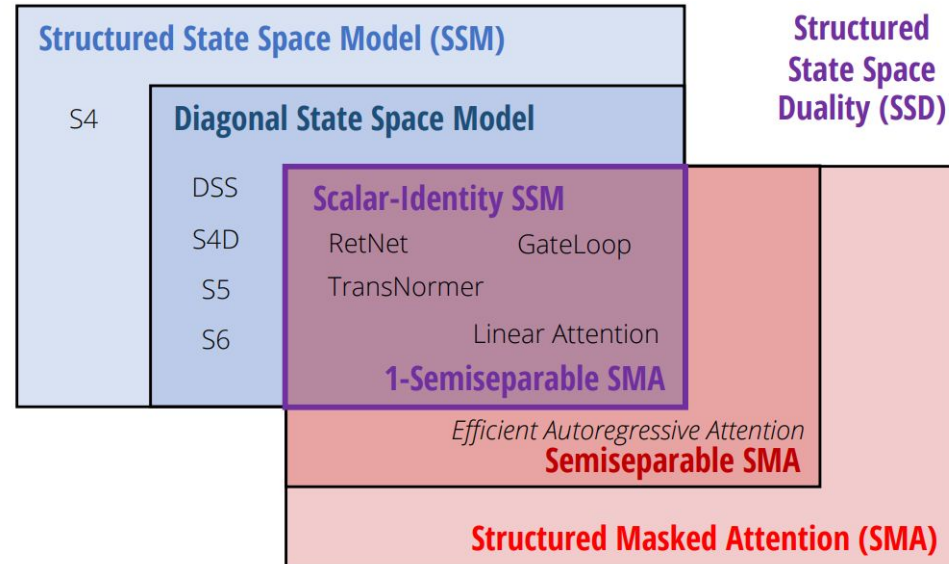


Model	Recurrence	Memory read-out
Linear Attention [39, 38]	$\mathbf{S}_t = \mathbf{S}_{t-1} + \mathbf{v}_t \mathbf{k}_t^\top$	$\mathbf{o}_t = \mathbf{S}_t \mathbf{q}_t$
+ Kernel	$\mathbf{S}_t = \mathbf{S}_{t-1} + \mathbf{v}_t \phi(\mathbf{k}_t)^\top$	$\mathbf{o}_t = \mathbf{S}_t \phi(\mathbf{q}_t)$
+ Normalization	$\mathbf{S}_t = \mathbf{S}_{t-1} + \mathbf{v}_t \phi(\mathbf{k}_t)^\top, \mathbf{z}_t = \mathbf{z}_{t-1} + \phi(\mathbf{k}_t)$	$\mathbf{o}_t = \mathbf{S}_t \phi(\mathbf{q}_t) / (\mathbf{z}_t^\top \phi(\mathbf{q}_t))$
DeltaNet [84]	$\mathbf{S}_t = \mathbf{S}_{t-1} (\mathbf{I} - \beta_t \mathbf{k}_t \mathbf{k}_t^\top) + \beta_t \mathbf{v}_t \mathbf{k}_t^\top$	$\mathbf{o}_t = \mathbf{S}_t \mathbf{q}_t$
Gated RFA [68]	$\mathbf{S}_t = g_t \mathbf{S}_{t-1} + (1 - g_t) \mathbf{v}_t \mathbf{k}_t^\top, \mathbf{z}_t = g_t \mathbf{z}_{t-1} + (1 - g_t) \mathbf{k}_t$	$\mathbf{o}_t = \mathbf{S}_t \mathbf{q}_t / (\mathbf{z}_t^\top \mathbf{q}_t)$
S4 [26]	$\mathbf{S}_t = \mathbf{S}_{t-1} \odot \exp(-(\alpha \mathbf{1}^\top) \odot \exp(\mathbf{A})) + \mathbf{B} \odot (\mathbf{v}_t \mathbf{1}^\top)$	$\mathbf{o}_t = (\mathbf{S}_t \odot \mathbf{C}) \mathbf{1} + \mathbf{d} \odot \mathbf{v}_t$
DFW [53, 40]	$\mathbf{S}_t = \mathbf{S}_{t-1} \odot (\beta_t \alpha_t^\top) + \mathbf{v}_t \mathbf{k}_t^\top$	$\mathbf{o}_t = \mathbf{S}_t \mathbf{q}_t$
RetNet [90, 75]	$\mathbf{S}_t = \gamma \mathbf{S}_{t-1} + \mathbf{v}_t \mathbf{k}_t^\top$	$\mathbf{o}_t = \mathbf{S}_t \mathbf{q}_t$
Mamba [25]	$\mathbf{S}_t = \mathbf{S}_{t-1} \odot \exp(-(\alpha_t \mathbf{1}^\top) \odot \exp(\mathbf{A})) + (\alpha_t \odot \mathbf{v}_t) \mathbf{k}_t^\top$	$\mathbf{o}_t = \mathbf{S}_t \mathbf{q}_t + \mathbf{d} \odot \mathbf{v}_t$
GLA [99]	$\mathbf{S}_t = \mathbf{S}_{t-1} \odot (\mathbf{1} \alpha_t^\top) + \mathbf{v}_t \mathbf{k}_t^\top = \mathbf{S}_{t-1} \text{Diag}(\alpha_t) + \mathbf{v}_t \mathbf{k}_t^\top$	$\mathbf{o}_t = \mathbf{S}_t \mathbf{q}_t$
RWKV-6 [67]	$\mathbf{S}_t = \mathbf{S}_{t-1} \text{Diag}(\alpha_t) + \mathbf{v}_t \mathbf{k}_t^\top$	$\mathbf{o}_t = (\mathbf{S}_{t-1} + (\mathbf{d} \odot \mathbf{v}_t) \mathbf{k}_t^\top) \mathbf{q}_t$
HGRN-2 [78]	$\mathbf{S}_t = \mathbf{S}_{t-1} \text{Diag}(\alpha_t) + \mathbf{v}_t (\mathbf{1} - \alpha_t)^\top$	$\mathbf{o}_t = \mathbf{S}_t \mathbf{q}_t$
mLSTM [7]	$\mathbf{S}_t = f_t \mathbf{S}_{t-1} + i_t \mathbf{v}_t \mathbf{k}_t^\top, \mathbf{z}_t = f_t \mathbf{z}_{t-1} + i_t \mathbf{k}_t$	$\mathbf{o}_t = \mathbf{S}_t \mathbf{q}_t / \max\{1,  \mathbf{z}_t^\top \mathbf{q}_t \}$
Mamba-2 [16]	$\mathbf{S}_t = \gamma_t \mathbf{S}_{t-1} + \mathbf{v}_t \mathbf{k}_t^\top$	$\mathbf{o}_t = \mathbf{S}_t \mathbf{q}_t$

# Возвращение рекуррентных сетей

- Например, Mamba-2 ([Dao, Gu, 2024](#)): Transformers are State-Space Models
- Сейчас не будем углубляться в это, но это очень интересное направление, которое сейчас активно разрабатывается

Structured State Space Model		Structured Masked Attention	
$C$	(output matrix)	$Q$	(queries)
$B$	(input matrix)	$K$	(keys)
$X$	(input sequence)	$V$	(values)
$A_{j:i}$	(state matrix)	$L_{ji}$	(mask)
$N$	(state expansion dim.)	$N$	(kernel feature dim.)
$H$ (hidden states (6b))		SMA linear dual (6)	
$= L \cdot X B$ (linear mode)			
SSM quadratic dual (7)		$G$	(Gram matrix (7a))
		$= Q \cdot K^T$ (quadratic mode)	

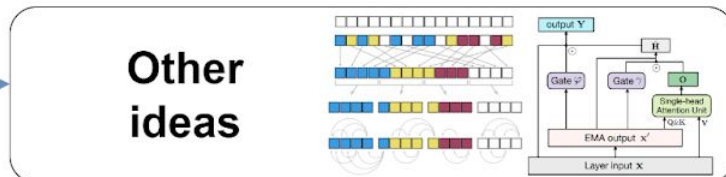
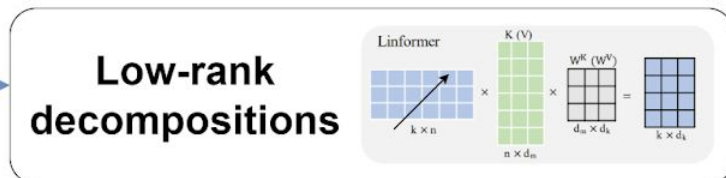


# Выводы

- Расширение контекста – это важное направление, и здесь внезапно снова появляются и рекуррентные сети, и SVD, и SSM, и много других идей
- Думаю, что здесь последнее слово ещё далеко не сказано... но есть и способ обрабатывать много внешней информации, не расширяя контекст



**Quadratic complexity**



**Large context windows**



# 7. Retrieval-Augmented Generation



Я ищу таких, как я,  
Сумасшедших и смешных,  
сумасшедших и больных  
А когда я их найду –  
Мы уйдём отсюда прочь,  
мы уйдём отсюда в ночь –  
мы уйдём из зоопарка!

*Егор Летов. Зоопарк*

# Что такое RAG

- LLM – большие модели, и ОНИ ЗНАЮТ МНОГО
- Но не всё, и по определению МНОГОГО знать не могут

hi again! what is your knowledge cutoff date?



My knowledge was last updated in August 2023. This means that while I can provide information up to that point, I may not have the most current data or details on events, research, or developments that occurred after that time. If you need the latest information, I can use my browsing capabilities to retrieve it.



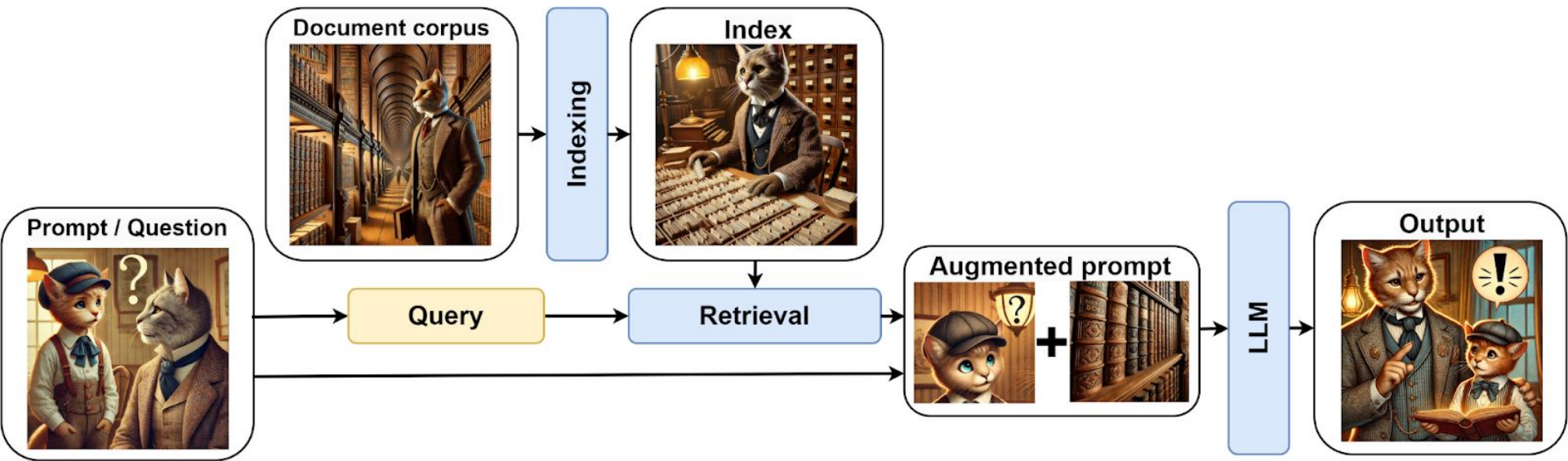
great! can you link me to more detailed surveys of RAG? again, use only your knowledge, don't look up information



I can't provide direct links to external sources without looking them up. However, I can recommend some places where you can find detailed surveys on Retrieval-Augmented Generation (RAG):

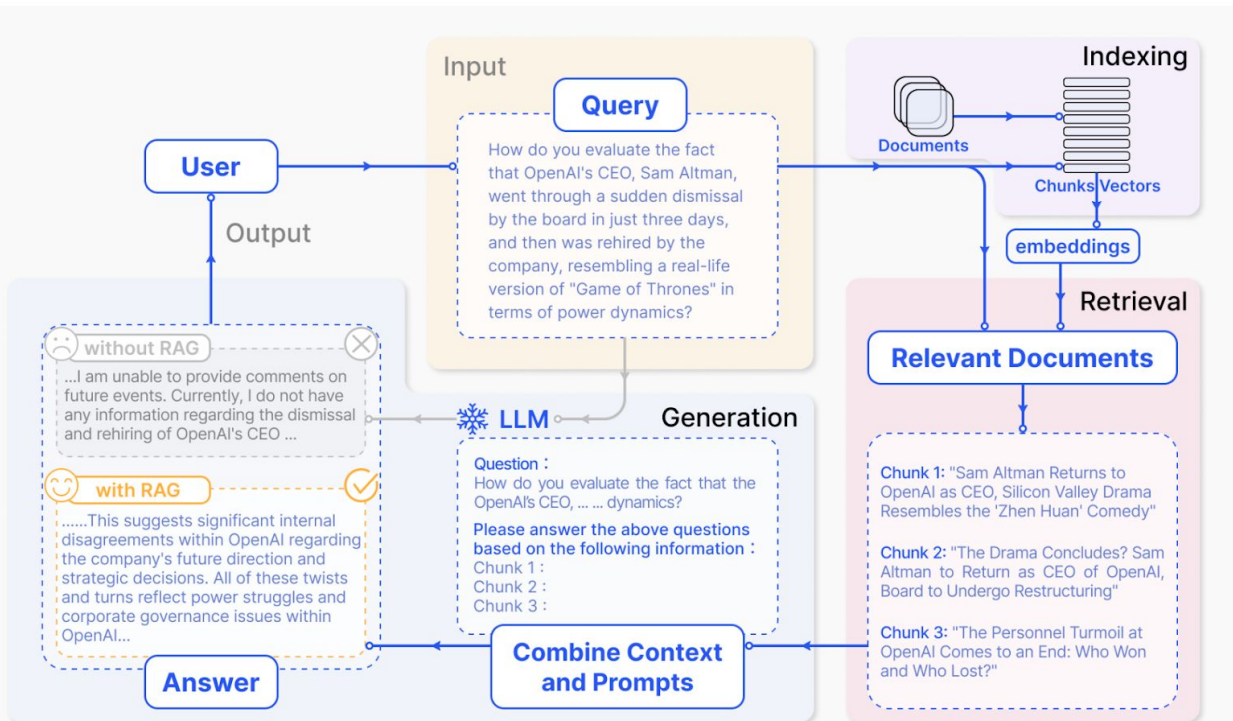
# Что такое RAG

- Предположим, что мы хотим дать модели возможность искать ответ
- Это приводит к примерно такой схеме, и нам нужно её чем-то наполнить:



# Что такое RAG

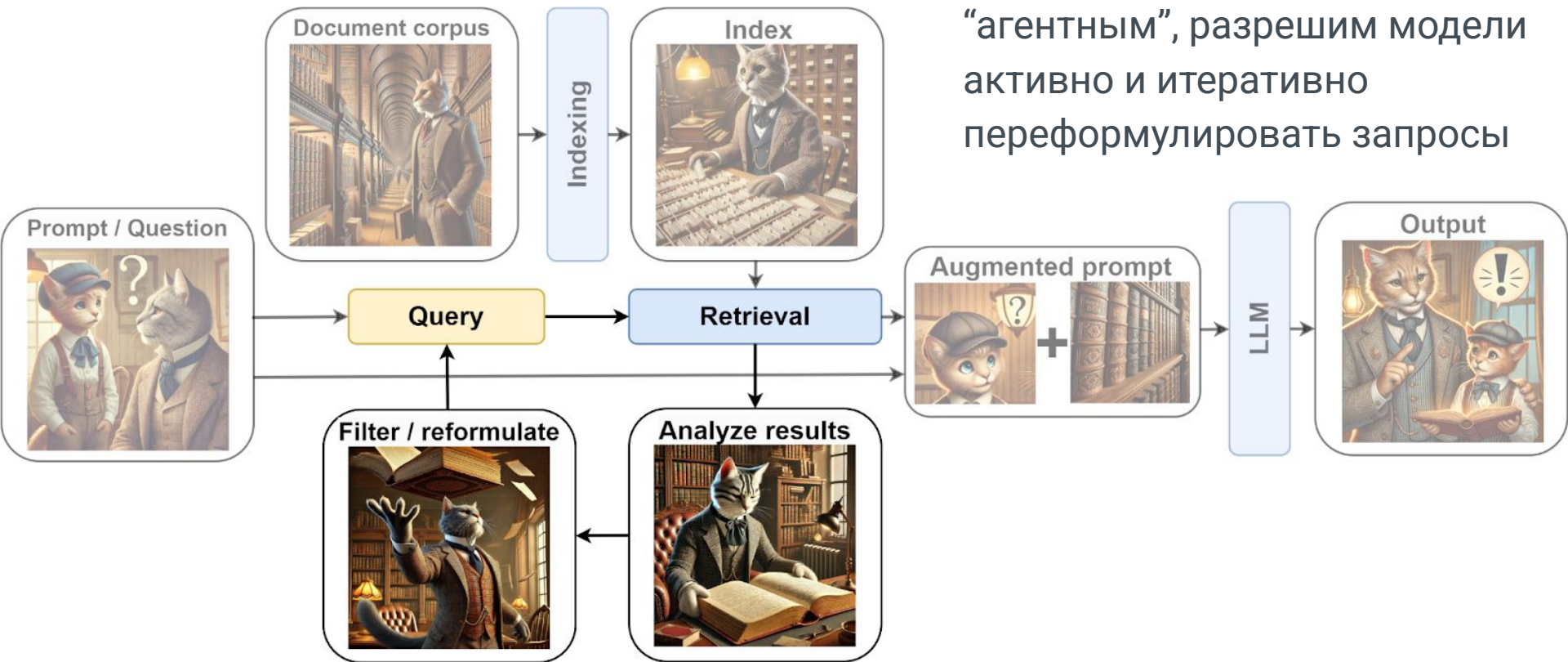
- Пример



# Примеры простых применений RAG

- [Yue et al. \(2023\)](#): DISC-LawLLM, RAG для юридических услуг, который умеет искать и читать законы и прецеденты
- [Xiong et al. \(2024\)](#): Medical Information Retrieval-Augmented Generation Evaluation (MIRAGE) ищет медицинские статьи и отчёты об исследованиях
- [Balaguer et al. \(2024\)](#): сравнение разных RAG и LLM fine-tuning подходов к сельскому хозяйству, где нужно читать много документов и отвечать на вопросы вроде “What is the best time to plant trees and shrubs in Arkansas?”
- [Suresh et al. \(2024\)](#): пример конкретного решения на основе RAG – работа с документами о БАК

# Agentic RAG



- Первое направление: давайте сделаем RAG более “агентным”, разрешим модели активно и итеративно переформулировать запросы

# Agentic RAG

- [Shao et al.](#)

---

## HotPotQA

---

(2023):

ИСПОЛЬ-

ЗУЕМ

ВЫХОД

ОДНОГО

РАУНДА

RAG КАК

ВХОД ДЛЯ

СЛЕДУЮ-

ЩЕГО

РАУНДА

**Question:** The arena where the Lewiston Maineiacs played their home games can seat how many people?

**Answer:** 3,677 seated

### Iteration 1

#### Retrieved Knowledge

(1) Title: Lewiston Maineiacs Context: The Lewiston Maineiacs were a junior ice hockey team ... **The team played its home games at the Androscoggin Bank Colisée ...**

(2) Title: Bangor Auditorium Context: ... a 5,948 seat multi-purpose arena located in Bangor, Maine ...

#### Generation

The Lewiston Maineiacs were a junior ice hockey team of the Quebec Major Junior Hockey League based in Lewiston, Maine. **The team played its home games at the Androscoggin Bank Colisée. The Androscoggin Bank Colisée can seat up to 5,948 people.**

So the answer is 5,948

### Iteration 2

#### Retrieved Knowledge

(1) Title: Lewiston Maineiacs Context: The Lewiston Maineiacs were a junior ice hockey team ... The team played its home games at the Androscoggin Bank Colisée ...

(2) Title: Androscoggin Bank Colisée Context: **The Androscoggin Bank Colisée (formerly Central Maine Civic Center and Lewiston Colisee) is a 4,000 capacity (3,677 seated) multi-purpose arena ...**

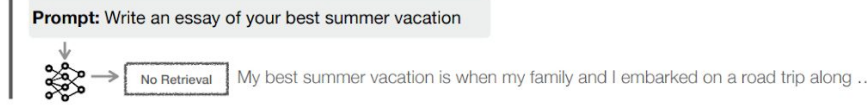
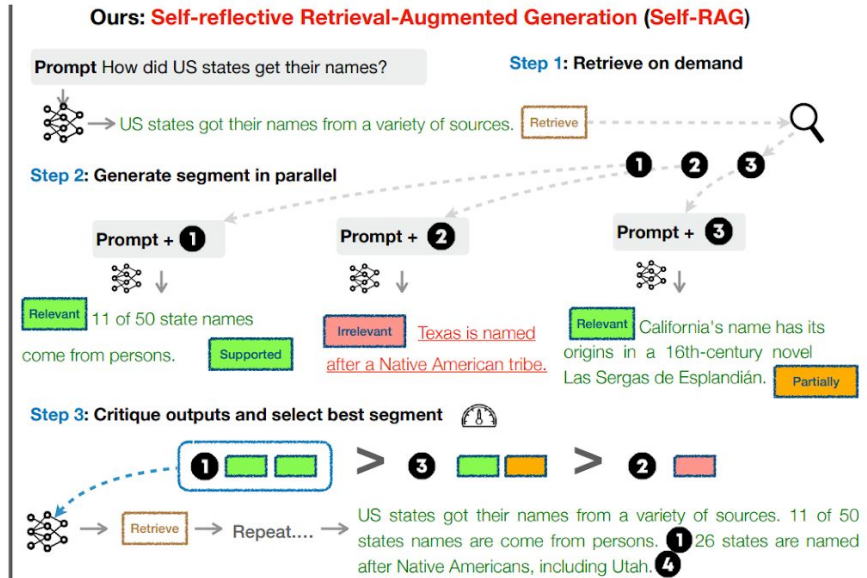
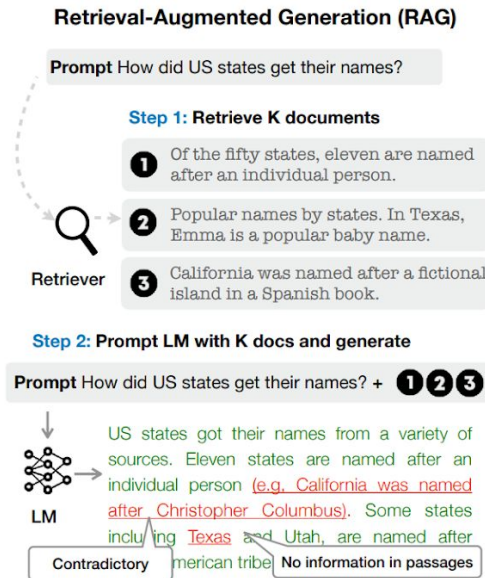
#### Generation

The Lewiston Maineiacs were a junior ice hockey team of the Quebec Major Junior Hockey League based in Lewiston, Maine. They played their home games at the Androscoggin Bank Colisée. **The Androscoggin Bank Colisée has a capacity of 4,000 (3,677 seated).** So the answer is 3,677

---

# Self-RAG (Self-Reflective RAG; Asai et al., 2023)

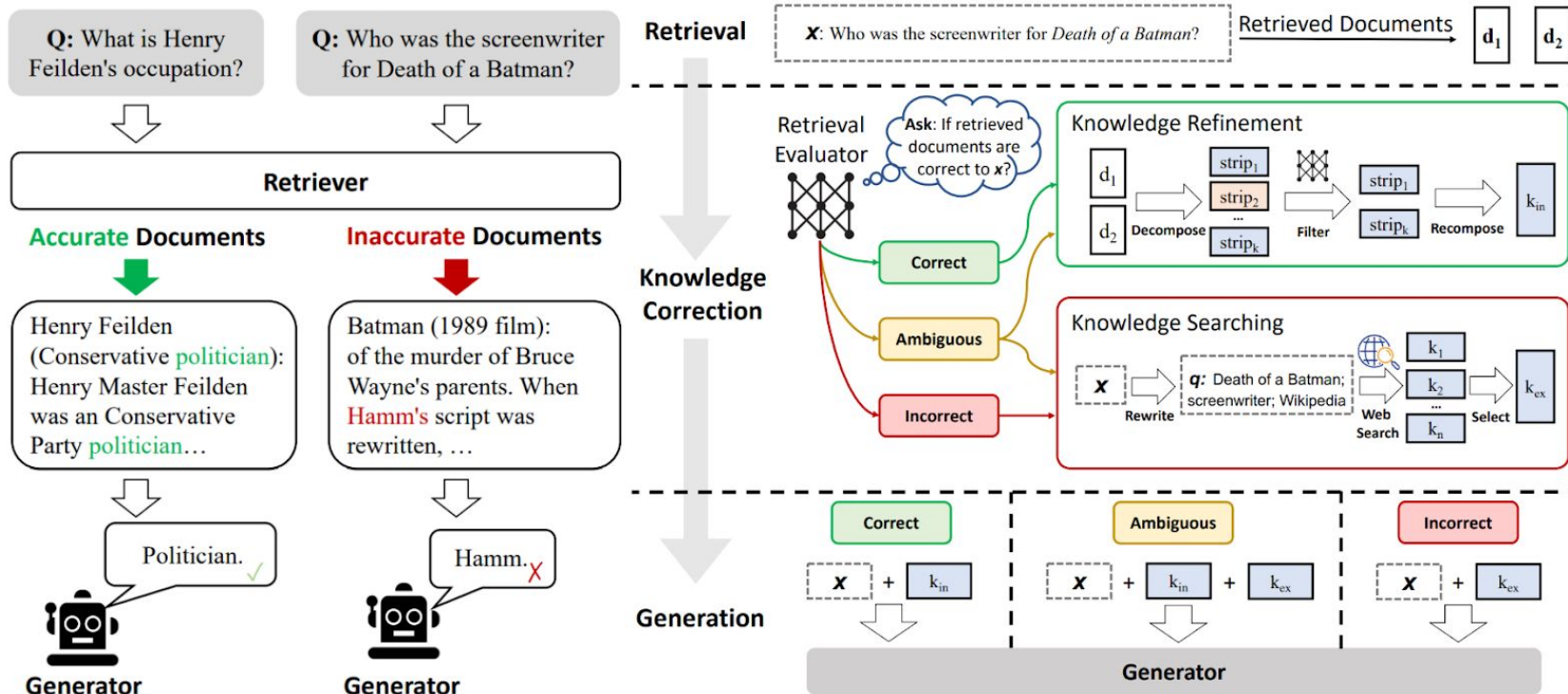
- Asai et al. (2023):
- ВЫЯСНИМ, НУЖЕН ЛИ ПОИСК; ЕСЛИ ДА, ПОРОДИМ ТОКЕН ПОИСКА
- оценим релевантность результатов выдачи
- породим ответы по каждому выданному документу и покритикуем их;
- выберем лучший, добавим его к запросу и повторим





# CRAG (Yan et al., 2024)

- [Yan et al. \(2024\)](#): ОТДЕЛЬНЫЙ ОЦЕНЩИК ВЫДАННЫХ ДОКУМЕНТОВ



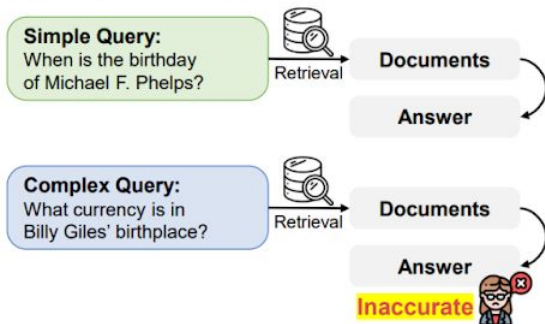
(a) CRAG motivation: inaccurate retrieval results

(b) CRAG inference with knowledge correction

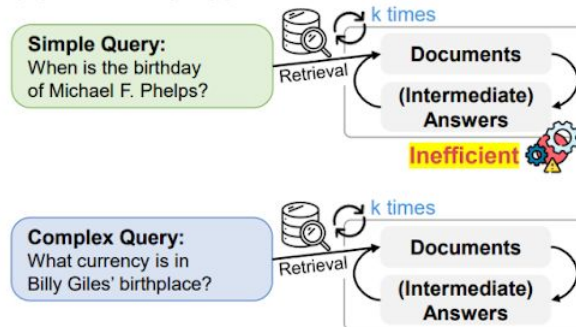
# Adaptive RAG (Jeong et al., 2024)

- Adaptive RAG ([Jeong et al., 2024](#)): обучает дополнительный классификатор, который выбирает правильный подход – нужен ли RAG вообще, нужно ли запускать его один раз или итеративно много раз
- Но рекурсивный перезапуск – это только одно направление улучшений для RAG; давайте рассмотрим другие...

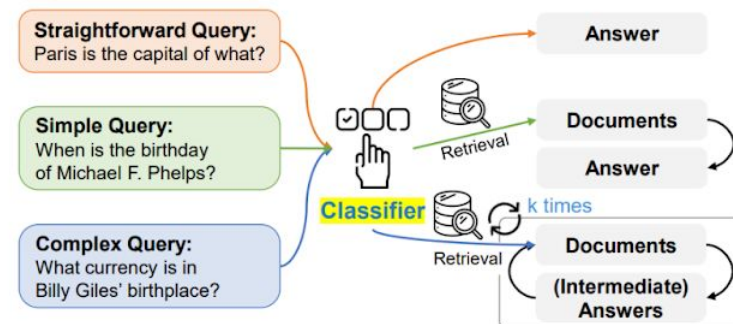
**(A) Single-Step Approach**



**(B) Multi-Step Approach**



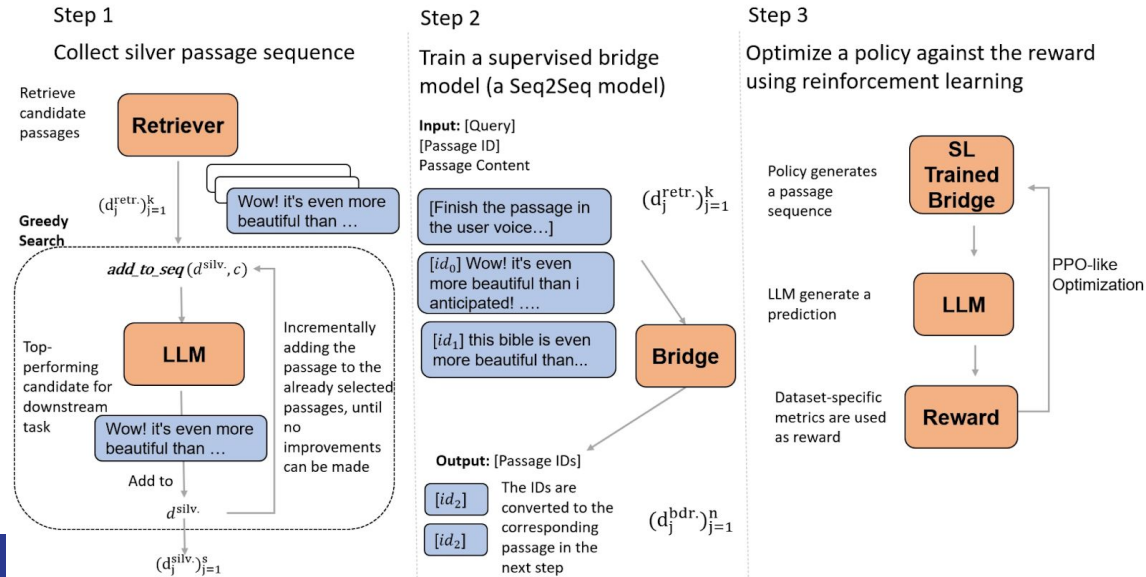
**(C) Our Adaptive Approach**



# Переранжирование

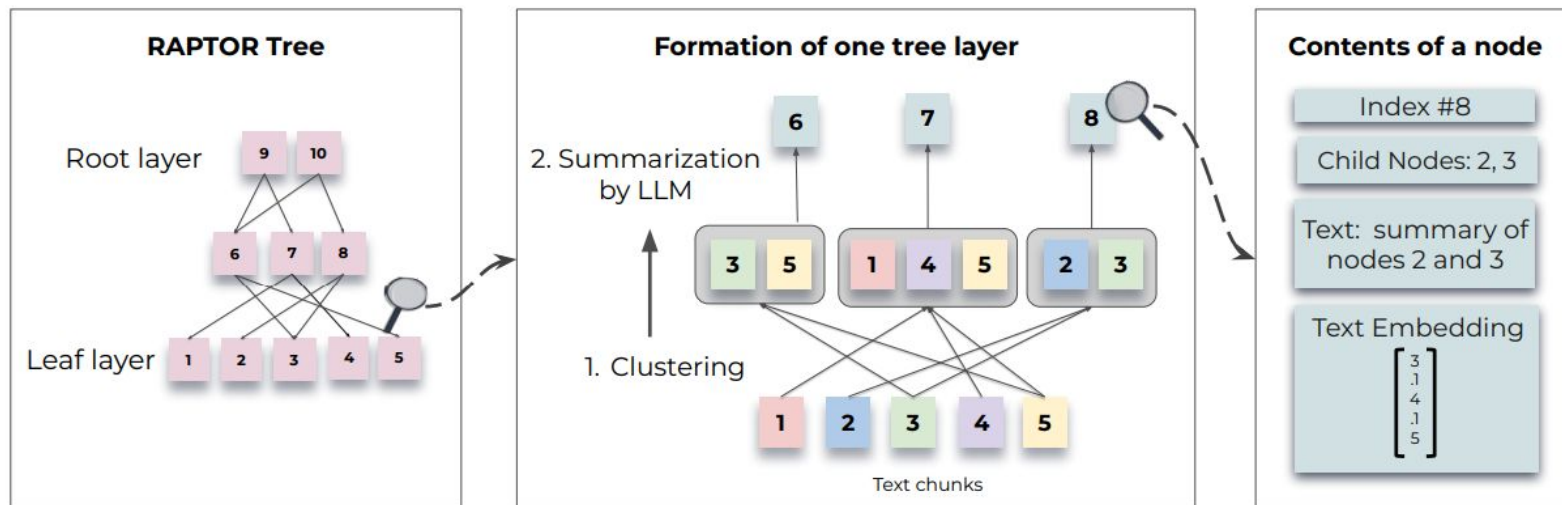
- Умное переранжирование результатов выдачи ([Gao et al., 2023](#), [Blagojevich, 2023](#)); это, кстати, и наоборот работает: LLM может улучшать поиск, помогая переранжировать ([Ma et al., 2023](#), [Peng et al., 2023](#))

- [Ke et al. \(2024\)](#): отдельная sequence-to-sequence модель, которая выбирает из документов конкретные части, более интересные для LLM



# Иерархия суммаризации

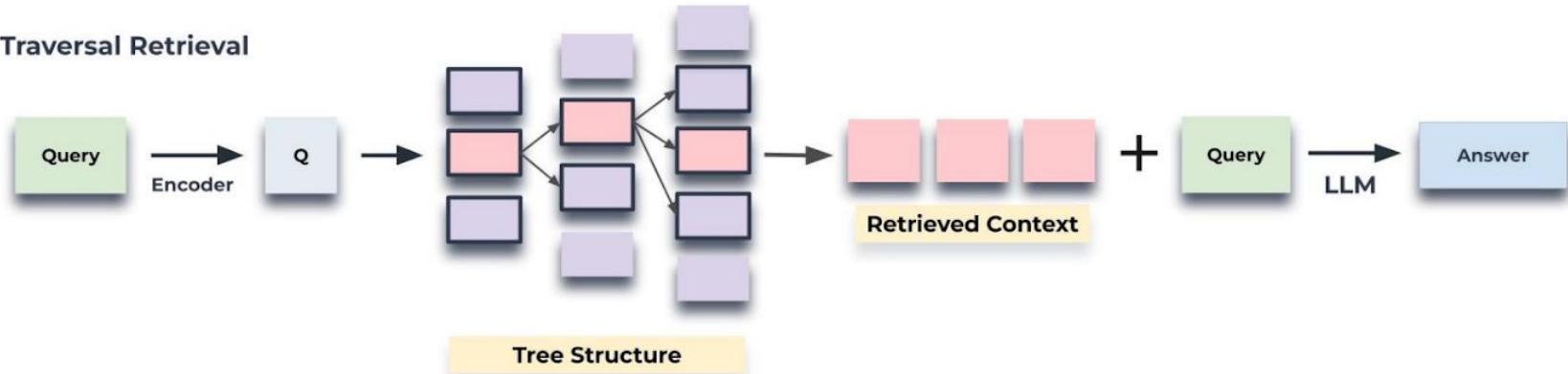
- Есть ряд иерархических подходов; например, RAPTOR (Recursive Abstractive Processing for Tree-Organized Retrieval; [Sarathi et al., 2024](#))
- LLM используется для того, чтобы рекурсивно суммаризировать длинные тексты (а потом наборы аннотаций и так далее)



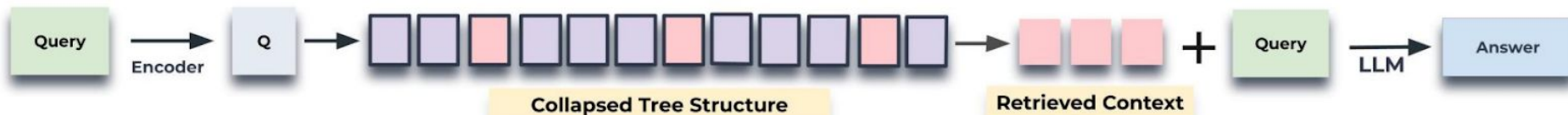
# Иерархия суммаризации

- А потом сам поиск можно запускать на этом дереве аннотаций, выдавая более глобальный контекст для LLM (который сам по себе не влез бы)

## A. Tree Traversal Retrieval

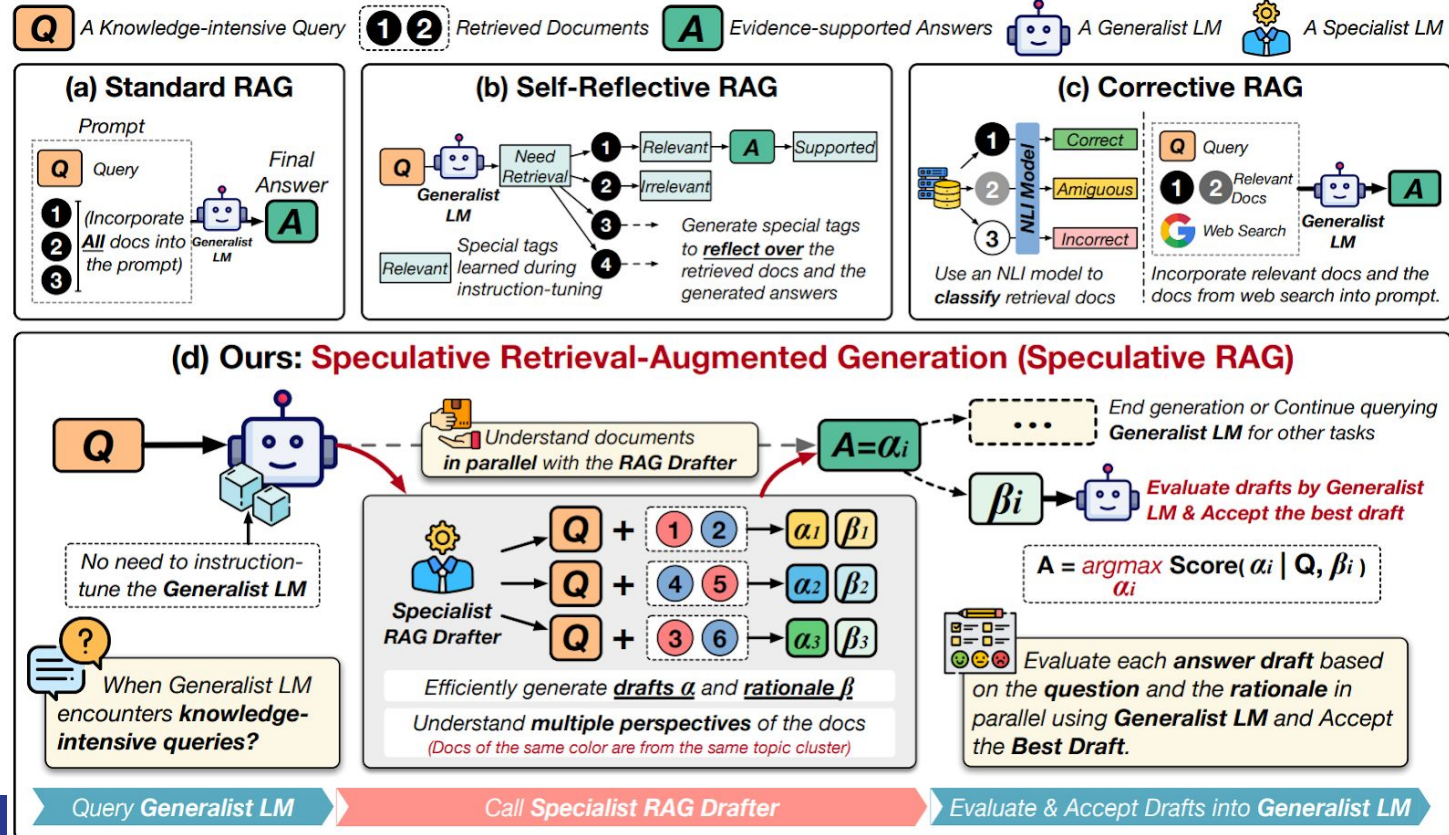


## B. Collapsed Tree Retrieval



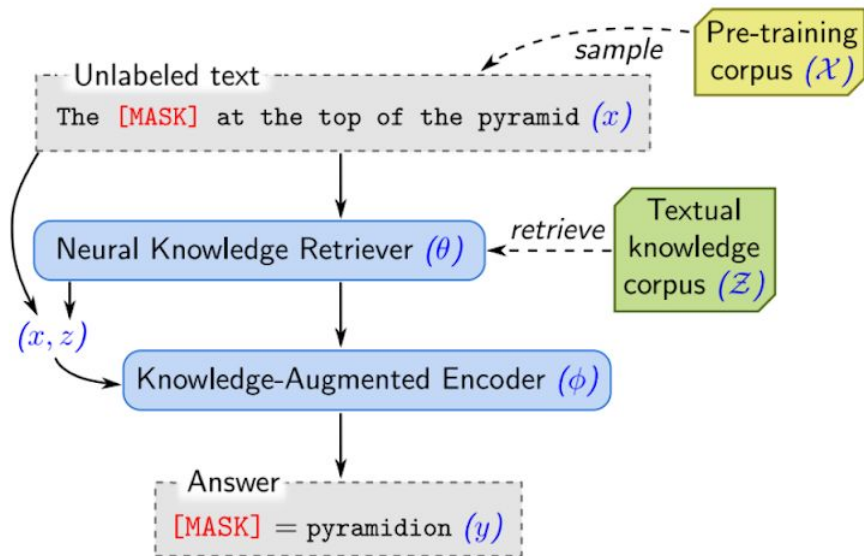
# Speculative RAG (Wang et al., 2024)

- [Wang et al. \(2024\)](#):  
 “Specialist LLM” отвечает на вопросы по документам,  
 “Generalist LLM” обобщает черновики от specialist LLMs

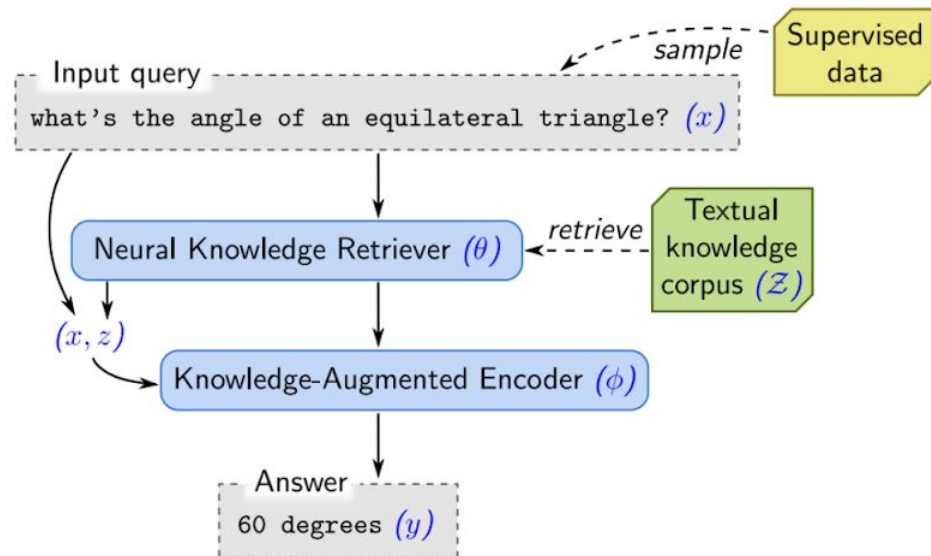


# От RAG к ReALM

- Связанный подход – R[e]ALM (Retrieval-Augmented Language Model; [Guu et al., 2020](#)): можно вставить поиск прямо в предсказание токенов; это тоже продолжается, но не будем сейчас ([Borgeaud et al., 2022](#), [Lin et al., 2024](#))



(a) REALM in pretraining



(b) REALM in supervised fine-tuning and inference

# Проблемы RAG

- Lost in the Middle ([Liu et al., 2024](#))

## Input Context

Write a high-quality answer for the given question using only the provided search results (some of which might be irrelevant).

Document [1] (Title: Asian Americans in science and technology) ...

**Document [2] (Title: List of Nobel laureates in Physics) ...**

Document [3] (Title: Scientist) ...

Document [4] (Title: Norwegian Americans) ...

Document [5] (Title: Maria Goeppert Mayer) ...

Question: who got the first nobel prize in physics

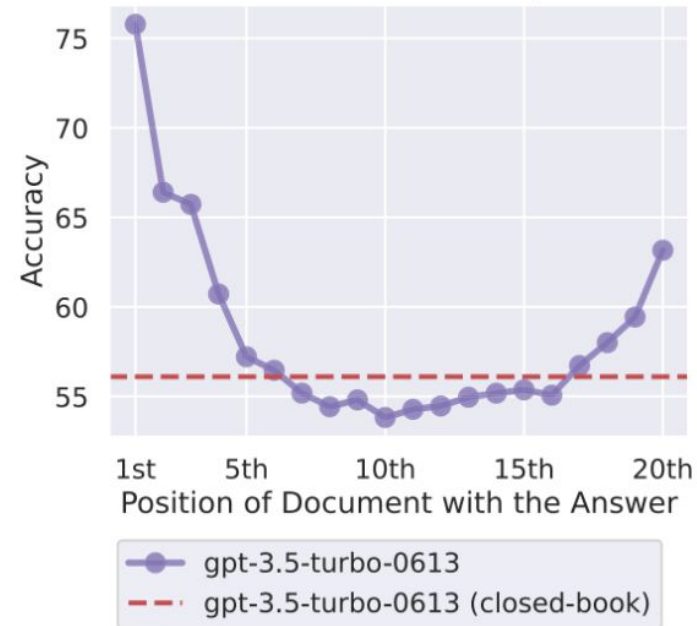
Answer:

## Desired Answer

Wilhelm Conrad Röntgen

(a) Multi-document question answering

20 Total Retrieved Documents (~4K tokens)

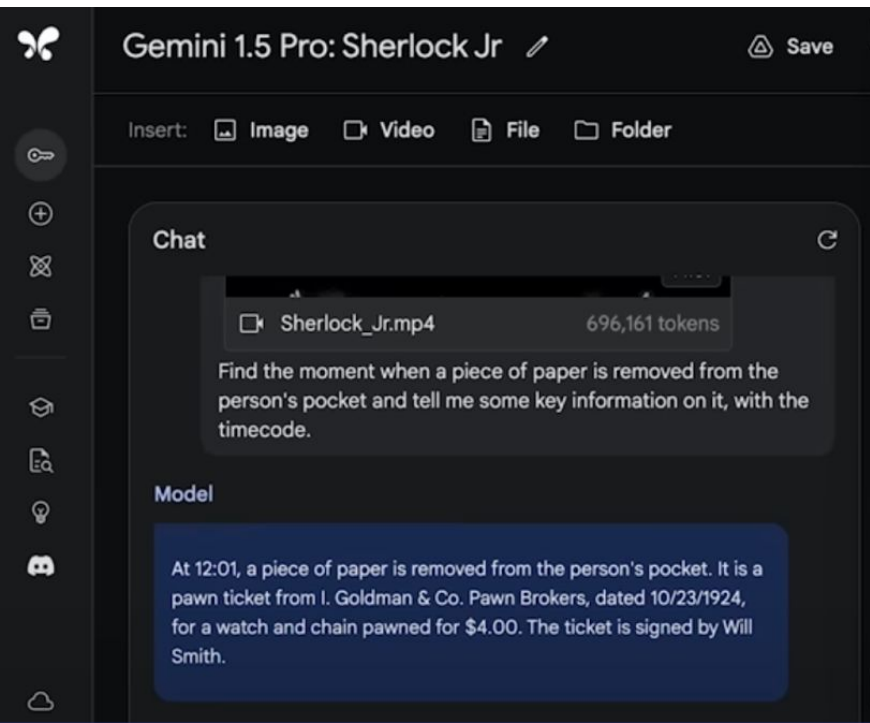




(b) Accuracy as a function of answer position








# Проблемы RAG

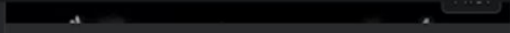
- А главное – полученные знания *неструктурированы*, и задачи типа “иголка в стоге сена” здесь не помогут, они не про творческую обработку выдачи поиска



Gemini 1.5 Pro: Sherlock Jr   Save

Insert:  Image  Video  File  Folder

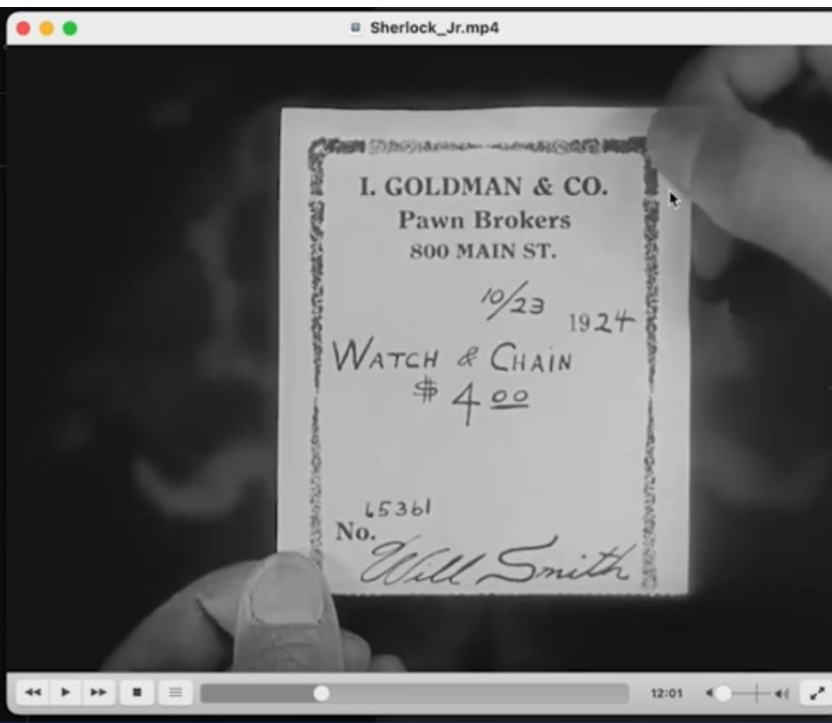
Chat 

 Sherlock\_Jr.mp4 696,161 tokens

Find the moment when a piece of paper is removed from the person's pocket and tell me some key information on it, with the timecode.

Model

At 12:01, a piece of paper is removed from the person's pocket. It is a pawn ticket from I. Goldman & Co. Pawn Brokers, dated 10/23/1924, for a watch and chain pawned for \$4.00. The ticket is signed by Will Smith.



# Графы знаний и RAG

- Графы знаний (knowledge graphs) кажутся мне недооценёнными; это хорошо развитая область ([Hogan et al., 2022](#); [Ji et al., 2021](#); [Heist et al., 2020](#); [Yan et al., 2018](#)), но кажется, что их можно было бы использовать чаще

(Albert Einstein, **BornIn**, German Empire)

(Albert Einstein, **SonOf**, Hermann Einstein)

(Albert Einstein, **GraduateFrom**, University of Zurich)

(Albert Einstein, **WinnerOf**, Nobel Prize in Physics)

(Albert Einstein, **ExpertIn**, Physics)

(Nobel Prize in Physics, **AwardIn**, Physics)

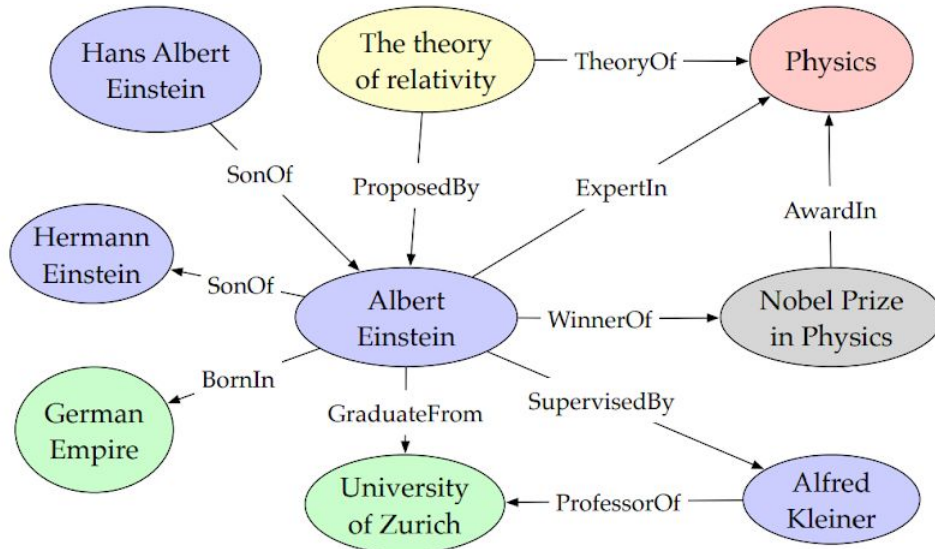
(The theory of relativity, **TheoryOf**, Physics)

(Albert Einstein, **SupervisedBy**, Alfred Kleiner)

(Alfred Kleiner, **ProfessorOf**, University of Zurich)

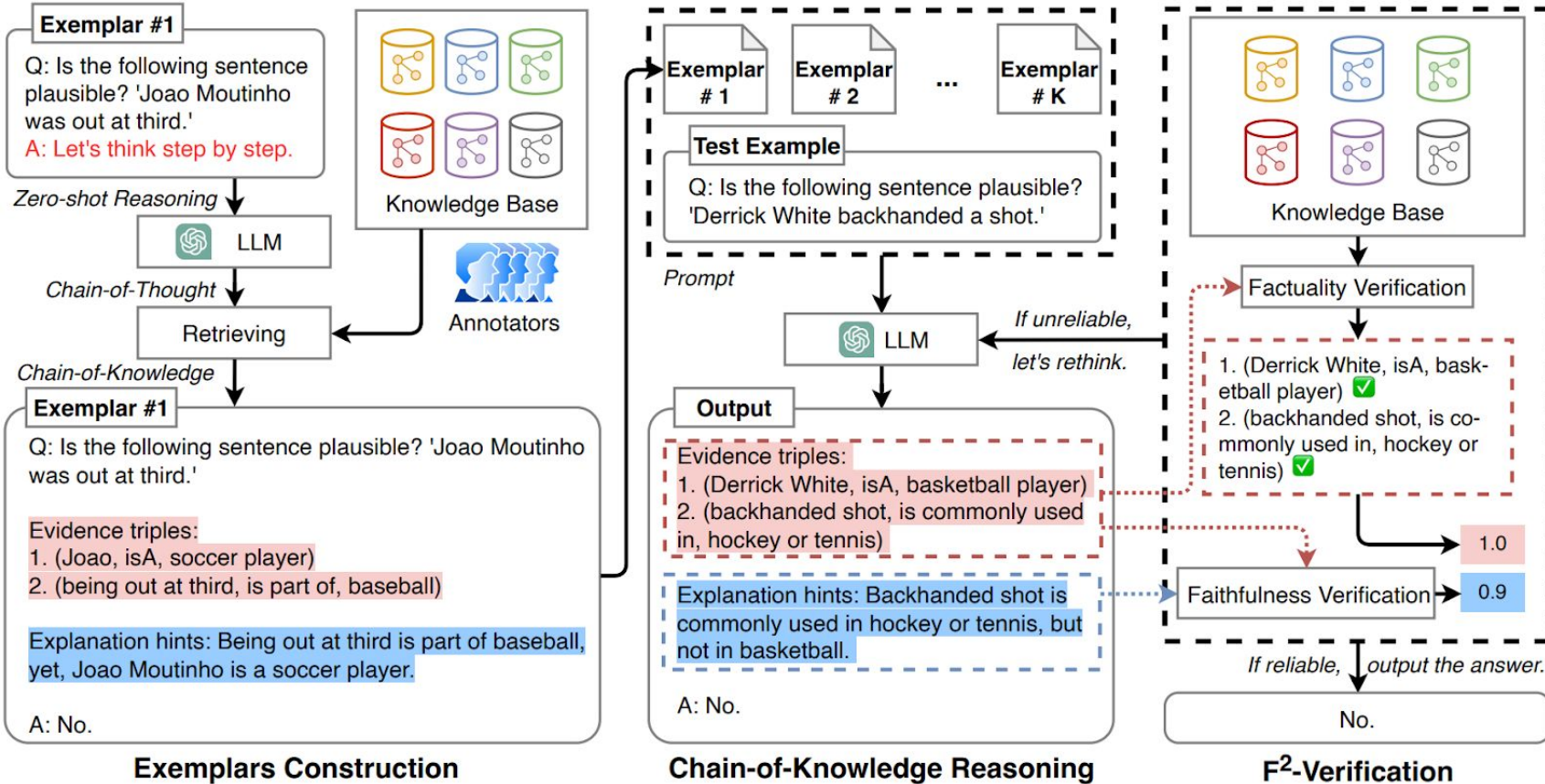
(The theory of relativity, **ProposedBy**, Albert Einstein)

(Hans Albert Einstein, **SonOf**, Albert Einstein)



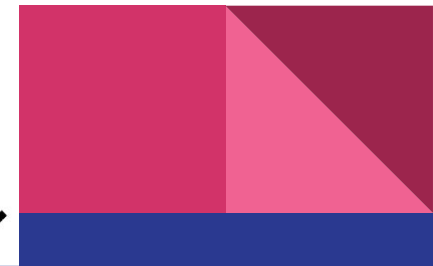
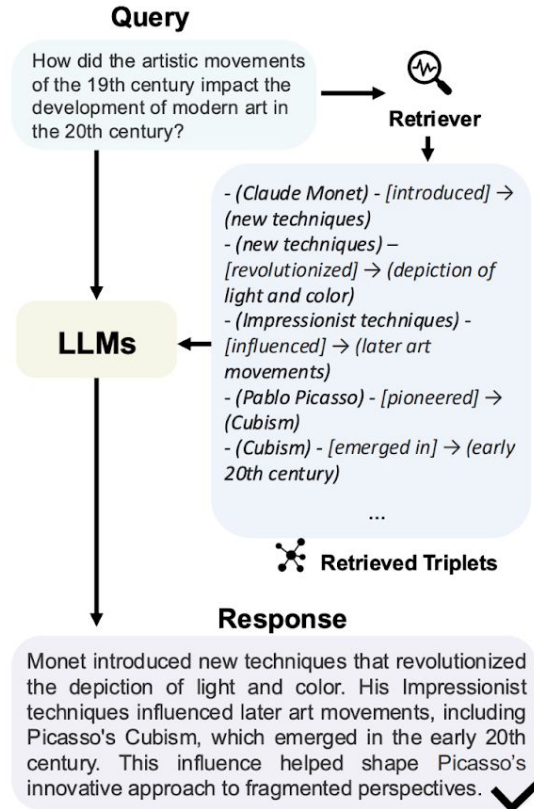
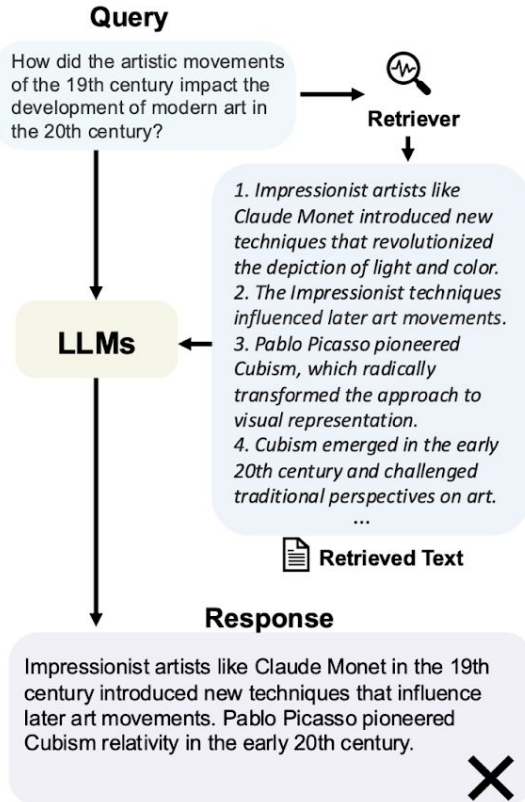
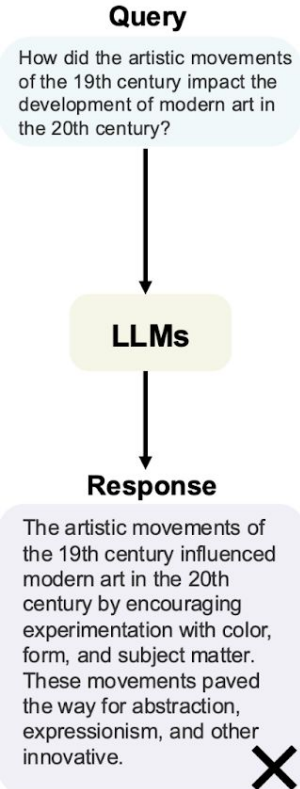
- Wang et al. (2023):  
chain-of-knowledge prompting

# Графы знаний и RAG



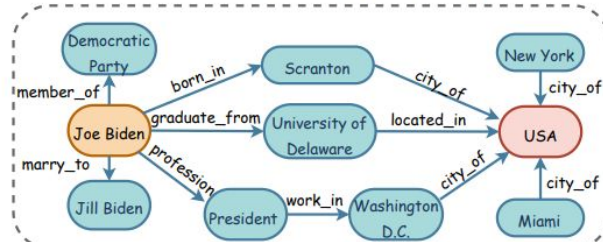
# Графы знаний и RAG

- [Peng et al. \(2024\)](#): если искать структурированные триплеты, ответы становятся точнее

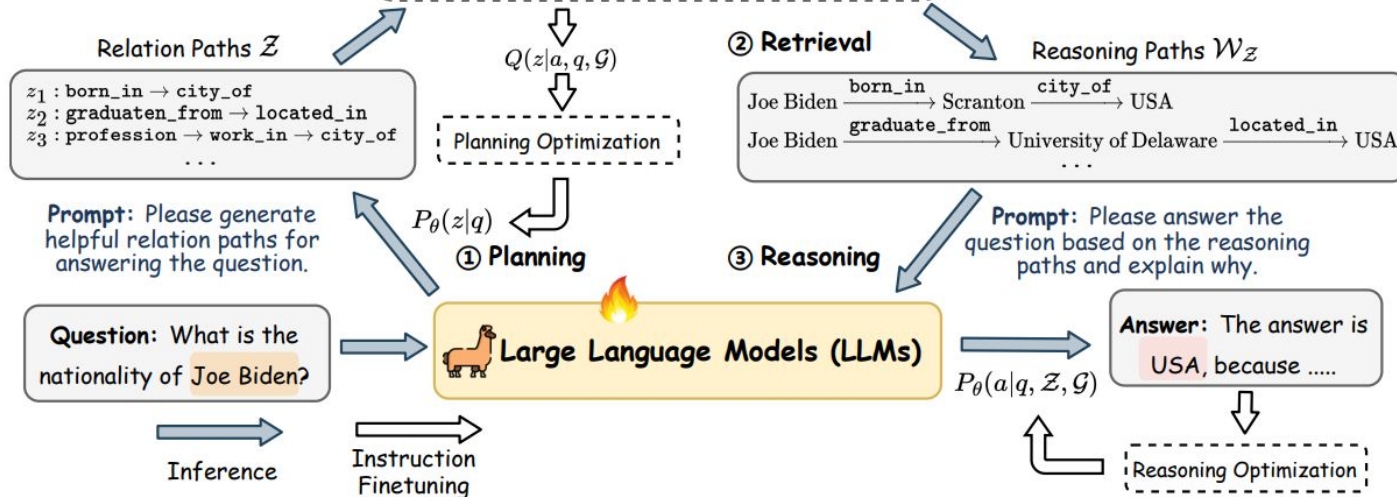


# Графы знаний и RAG

## Knowledge Graphs (KGs)



- [Luo et al. \(2024\)](#): reasoning on graphs (RoG)



# Графы знаний и RAG

- [Luo et al. \(2024\)](#): reasoning on graphs (RoG)

Question	What type of government is used in the country with Northern District?
Answer	Parliamentary system
ChatGPT + CoT	To determine what type of government is used in a country with a Northern District, <b>we need more specific information</b> . The term "Northern District" does not specify the name of any particular country or provide enough context to identify its government type. <b>Could you please provide additional details or clarify the specific country you are referring to?</b>
Reasoning Paths	Northern District → location.administrative_division.first_level_division_of → Israel → government.form_of_government.countries → <b>Parliamentary system</b>
RoG	<p>The type of government used in the country with the Northern District is a <b>Parliamentary system</b>.</p> <p>Explanations:</p> <ol style="list-style-type: none"> <li>Northern District → location.administrative_division.first_level_division_of → Israel: The Northern District is a part of Israel.</li> <li>Israel → government.form_of_government.countries → Parliamentary system: Israel has a Parliamentary system of government.</li> </ol> <p>So, by following the reasoning paths provided, we can conclude that the country with the Northern District, which is Israel, uses a Parliamentary system of government.</p>

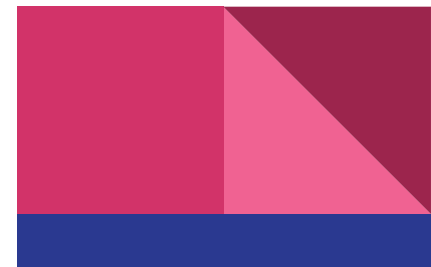
(a) RoG gathers additional knowledge to answer a question

Question	Where is the home stadium of the team who won the 1946 World Series championship?
Answer	Busch Stadium
ChatGPT + CoT	<p>Step 1: Identify the team that won the 1946 World Series championship. The team that won the 1946 World Series championship was the St. Louis Cardinals.</p> <p>Step 2: Determine the home stadium of the St. Louis Cardinals during that time. The St. Louis Cardinals' home stadium during the 1946 season was <b>Sportsman's Park</b>.</p> <p>Therefore, the possible answer to the question "Where is the home stadium of the team who won the 1946 World Series championship?" is: - <b>Sportsman's Park</b>.</p>
Reasoning Paths	<p>1946 World Series → sports.sports.team.championships → St. Louis Cardinals → sports.sports.team.arena_stadium → <b>Busch Stadium</b></p> <p>1946 World Series → sports.sports.team.championships → St. Louis Cardinals → sports.sports.team.arena_stadium → <b>Roger Dean Stadium</b></p>
RoG	<p>Based on the reasoning paths provided, the home stadium of the team who won the 1946 World Series championship is <b>Busch Stadium</b>. This conclusion is based on the following path:</p> <p><b>1. 1946 World Series → sports.sports.team.championships → St. Louis Cardinals → sports.sports.team.arena_stadium → Busch Stadium</b></p> <p>The reasoning path leads us to the St. Louis Cardinals as the team that won the 1946 World Series, and Busch Stadium is the stadium associated with the St. Louis Cardinals. Therefore, <b>Busch Stadium</b> is the home stadium of the team that won the 1946 World Series championship.</p>

(b) RoG successfully avoids hallucination

# RAG: итоги

- Никакого контекста не хватит на весь интернет, искать всё равно будет нужно
- RAG — одно из главных направлений использования инструментов для LLM (есть и другие), важное направление для дальнейших улучшений
- Есть много разных подходов, в том числе очень простые: In-context RALM ([Ram et al., 2023](#)) просто берёт все найденные документы и добавляет их к запросу для LLM, и тоже всё работает



# 8. Рассуждающие сети



Напрасны все предсказания, все рассуждения о том, что с вами может быть, когда вы под власть вашей страсти: увлечение сердца перемогает все красноречивые речи. Нам нужна другая воля, и она тогда только придёт, когда наше коварство нас порядком накажет... и мы почувствуем в том для себя ущерб, что мы... оставили Бога. Тогда только страх Божий к нам возвратится.

*Авдотья Глинка. Подневные рассуждения*



# Chain of Thought

- [Wei et al., 2022](#): ещё на этапе few-shot prompting люди заметили, что если давать модели более подробные примеры, то качество ответов в итоге вырастет

## Standard 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?

### Model Output

A: The answer is 27. ❌

## 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: 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 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. ✅

# Let's think step by step

- [Kojima et al. \(2022\)](#): помогает даже просто добавить “Let's think step by step”!

## (a) Few-shot

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: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A:

(Output) The answer is 8. ✗

## (b) Few-shot-CoT

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: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A:

(Output) The juggler can juggle 16 balls. Half of the balls are golf balls. So there are  $16 / 2 = 8$  golf balls. Half of the golf balls are blue. So there are  $8 / 2 = 4$  blue golf balls. The answer is 4. ✓

## (c) Zero-shot

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A: The answer (arabic numerals) is

(Output) 8 ✗

## (d) Zero-shot-CoT (Ours)

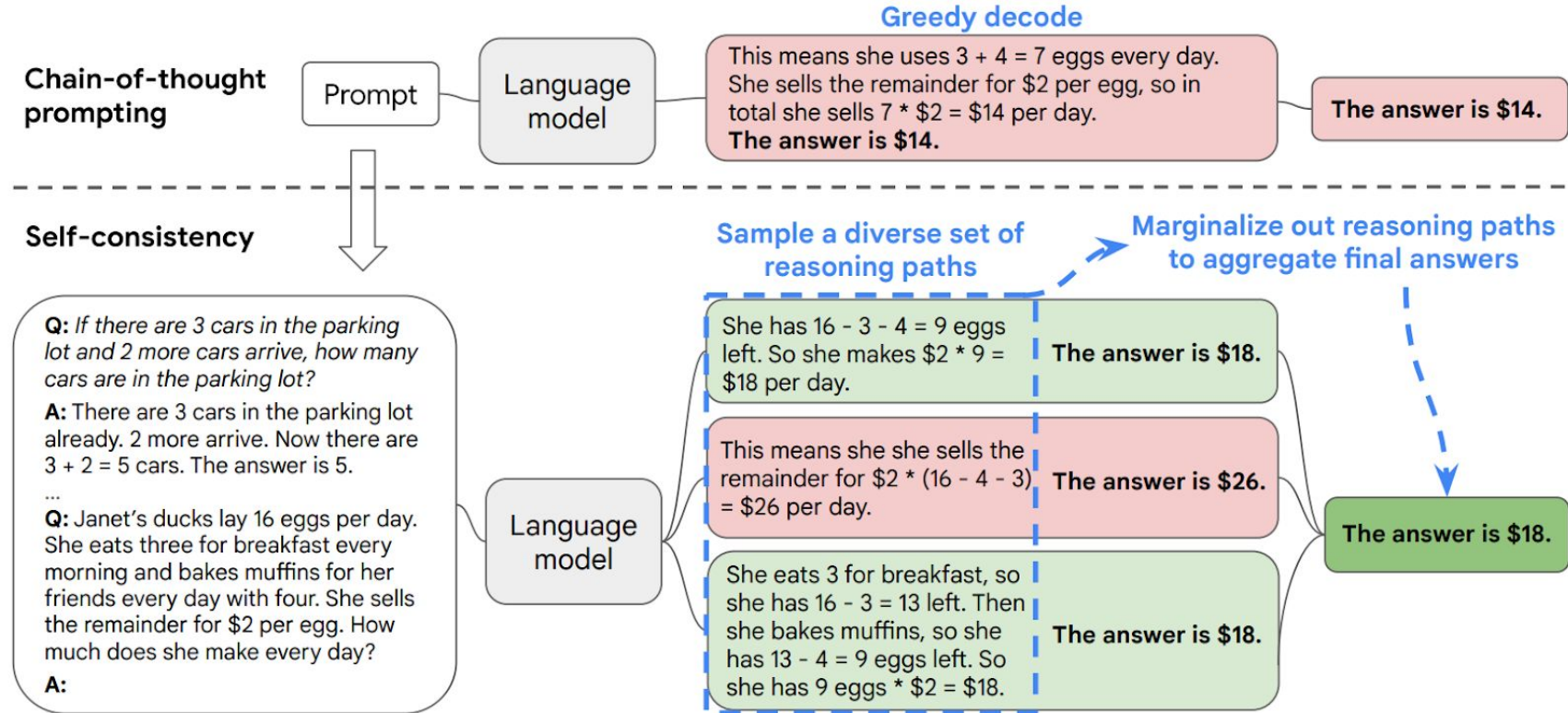
Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A: **Let's think step by step.**

(Output) There are 16 balls in total. Half of the balls are golf balls. That means that there are 8 golf balls. Half of the golf balls are blue. That means that there are 4 blue golf balls. ✓

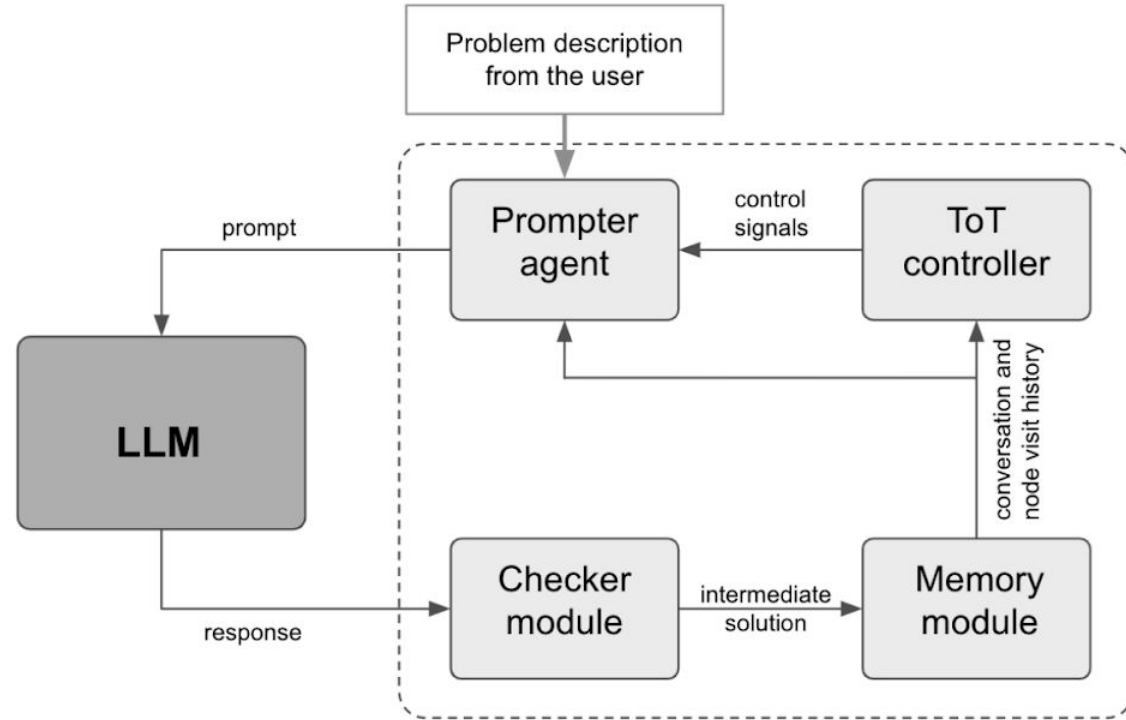
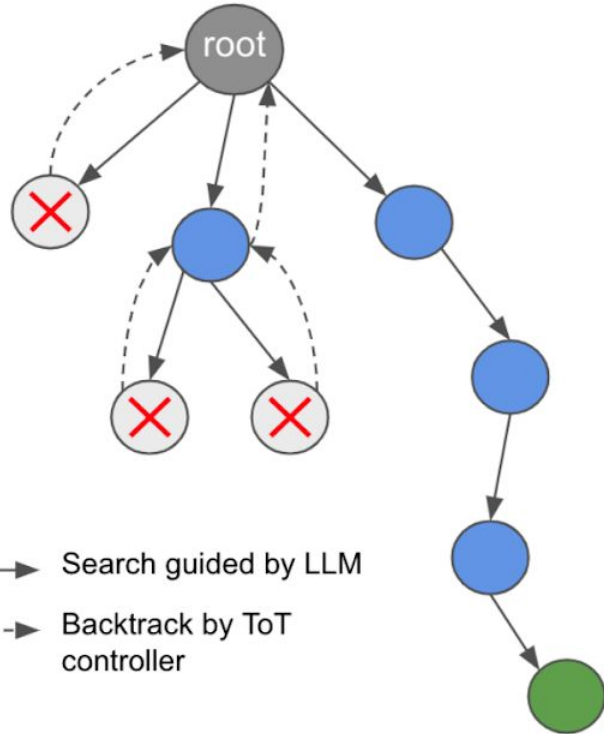
# От Chain-of-Thought к Graph-of-Thought

- [Wang et al. \(2022\)](#): Chain-of-Thought with Self-Consistency (CoT-SC)



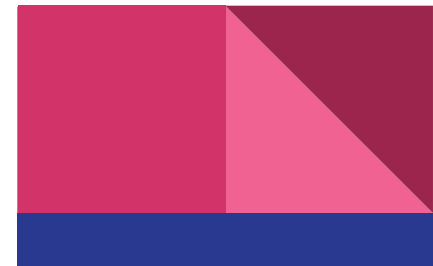
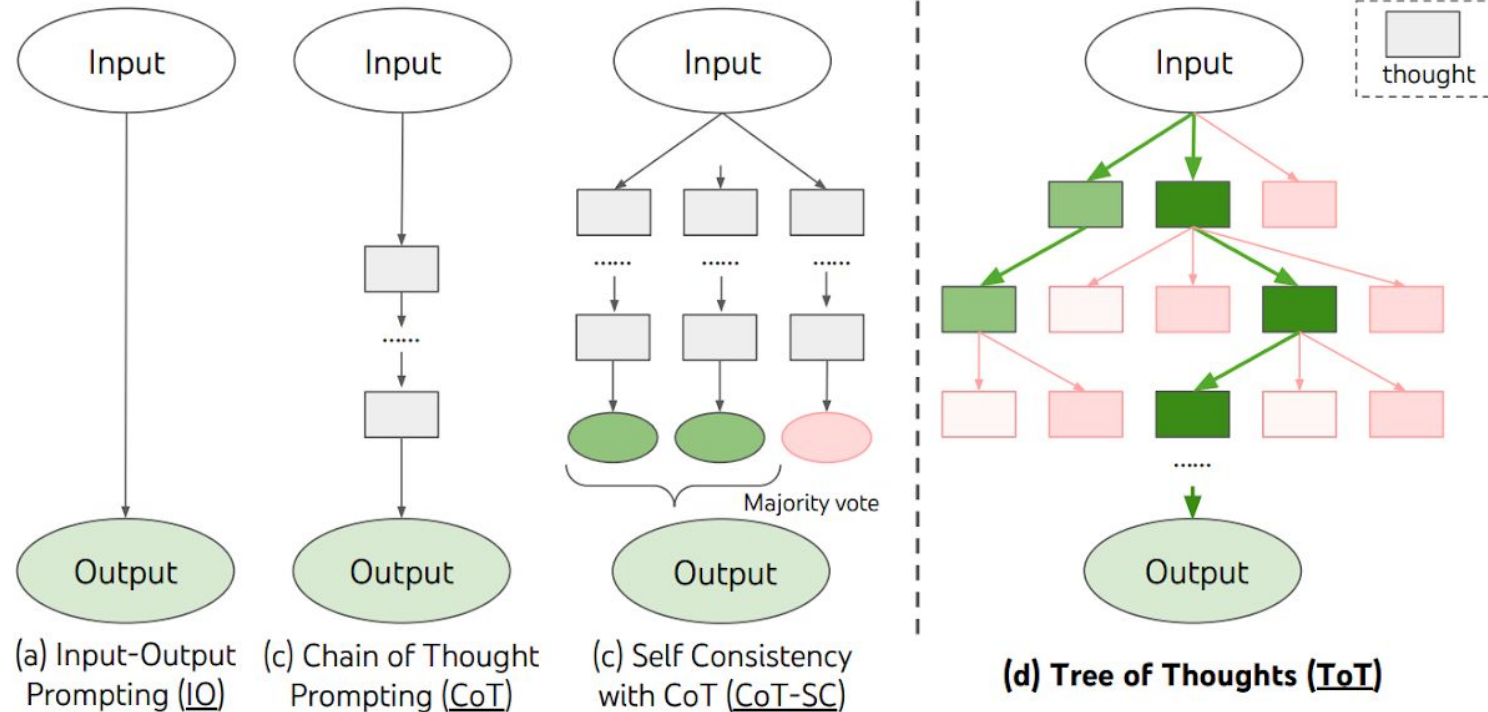
# От Chain-of-Thought к Graph-of-Thought

- Long (May 15, 2023): Tree-of-Thought через контроллер с откатом по дереву



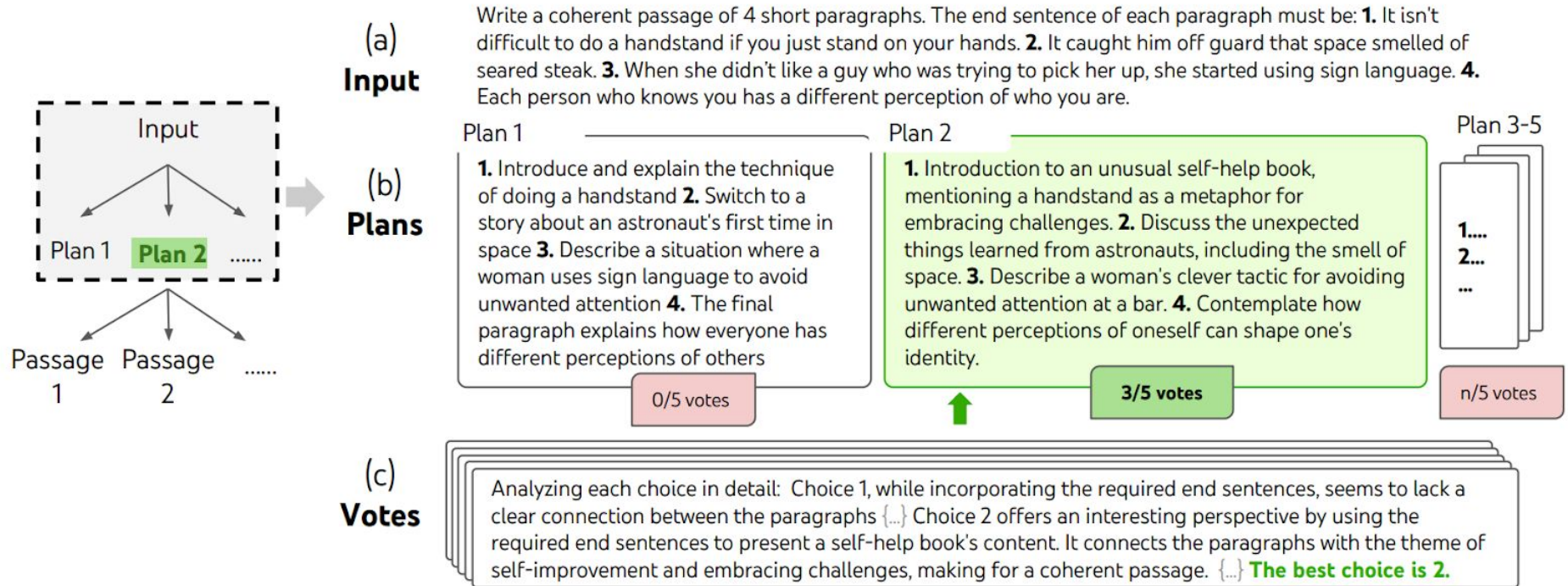
# От Chain-of-Thought к Graph-of-Thought

- [Yao et al. \(May 17, 2023\)](#): Tree-of-Thought с голосованием или другим механизмом



# От Chain-of-Thought к Graph-of-Thought

- [Yao et al. \(May 17, 2023\)](#): Tree-of-Thought с голосованием или другим механизмом

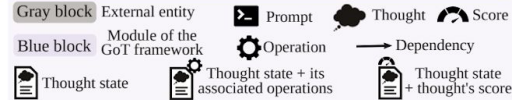


# От Chain-of-Thought к Graph-of-Thought

- [Besta et al. \(2023\)](#): логичный следующий шаг – граф операций с возможностью агрегации



## Legend



## API for Controller

- //LLM params: model used, temperature, max tokens, api key, org, ...
- //LLM cost features: prompt token cost, response token cost, ...
- //Instances of Prompter + Parser + Graph of Operations,
- //Any additional input parameters (e.g., numbers to be sorted).

## Available operations when building the GoO (extensible)

- Generate, Aggregate, Score, ... //see Prompter API
- KeepBest(N) //preserves N best scoring thoughts
- Repeat(k) //Repeat a given operation k times, generating k thoughts.  
 //For example, this enables "Aggregate" to generate multiple outcomes //of the combination operation. Each such thought is maintained //within the Graph Reasoning State and scored individually.

## API for Prompter (extensible)

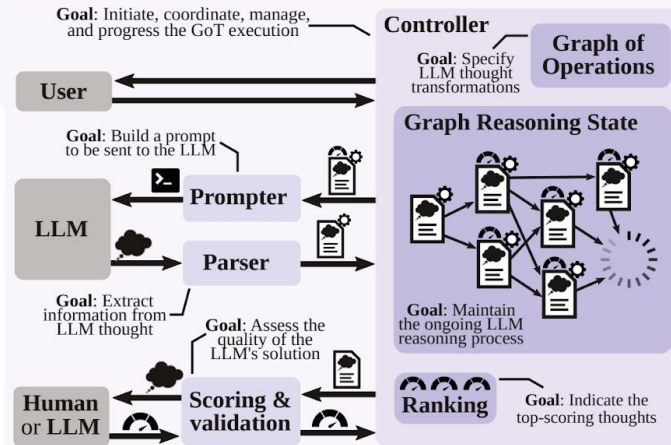
- Generate(t, k) //generate a prompt for k new thoughts, using thought t
- ValidateAndImprove(t) //generate a prompt to enhance thought t,
- Aggregate(t1, ..., tk) //generate a prompt to combine thoughts t1, ..., tk
- Score(t) //score thought t
- Validate(t) //generate a prompt to validate the correctness of thought t

## API for Parser (extensible)

ParseGenerate, ParseImprove, ParseScore, ParseAggregate, ParseValidate, ...

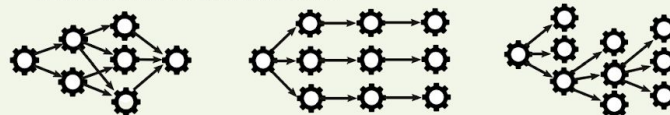
//Each of the above routines is responsible for parsing an LLM thought //to a corresponding Prompter routine (e.g., ParseScore parses Score).

## Architecture overview



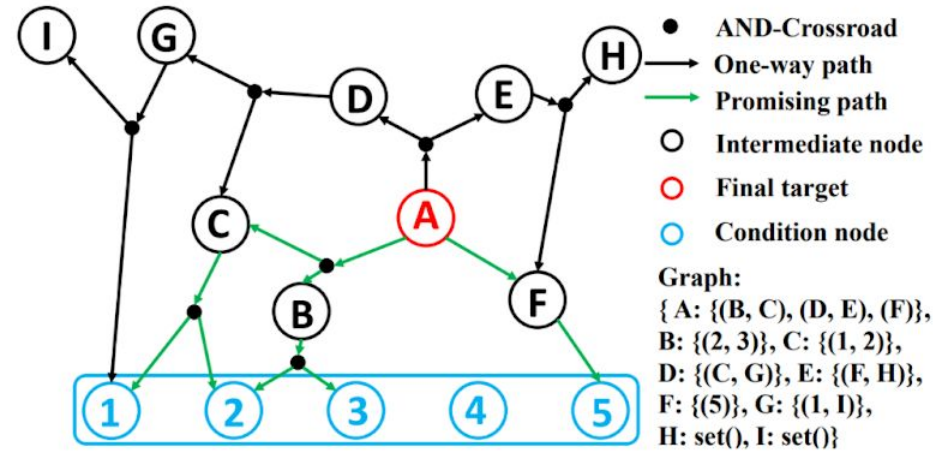
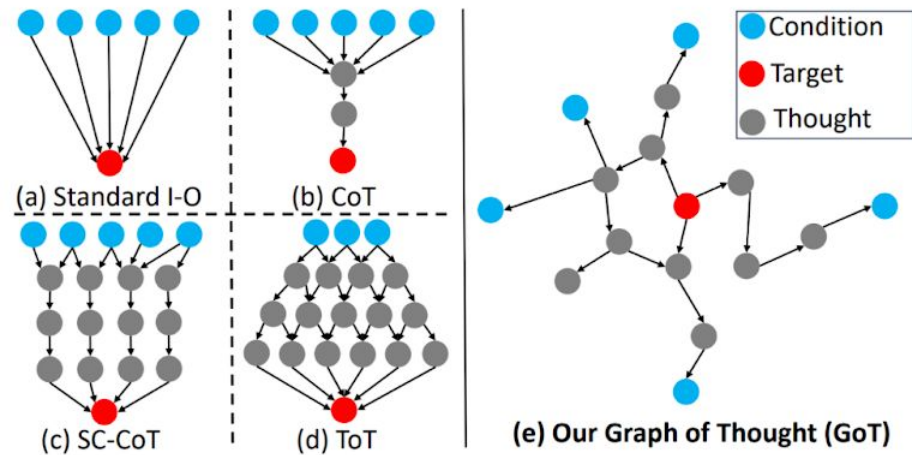
## Specifying the structure of the Graph of Operations (GoO)

Graph of Operations enables seamless specification of not only GoT, but also existing schemes such as CoT, GoT-SC, ToT



# От Chain-of-Thought к Graph-of-Thought

- [Lei et al. \(2023\)](#): добавляют узлы, которые обобщают то, чему модель научилась из предыдущих узлов, и добавляет результат как условия для дальнейших рассуждений



(a) Comparison of approaches by Lei et al. (2023)

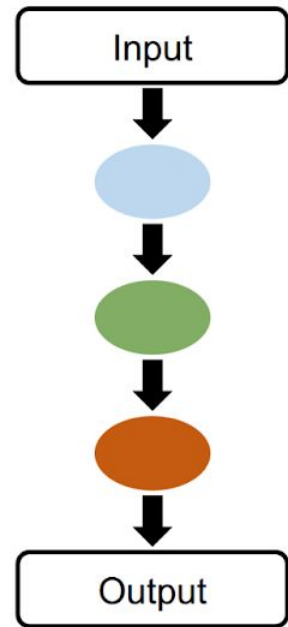
(b) Sample toy graph by Lei et al. (2023)



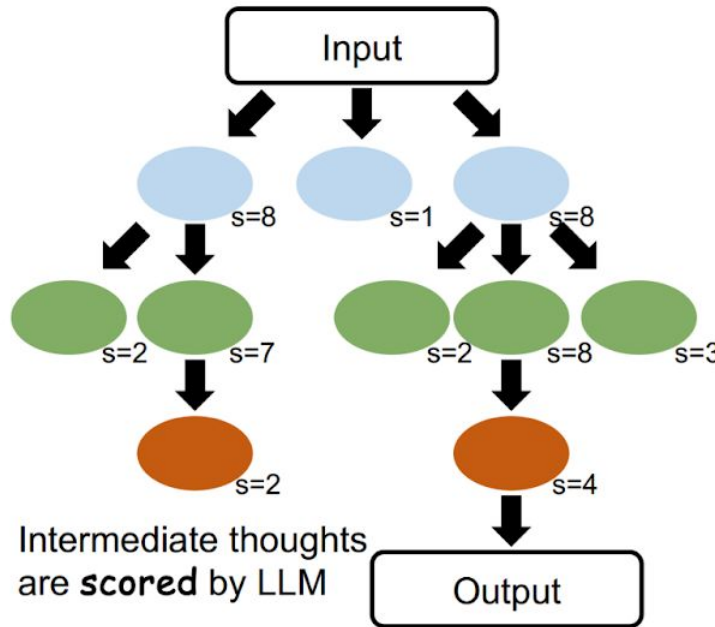
# От Chain-of-Thought к Graph-of-Thought

- [Zhang et al. \(2024\)](#): попарные сравнения вместо отдельных контроллеров

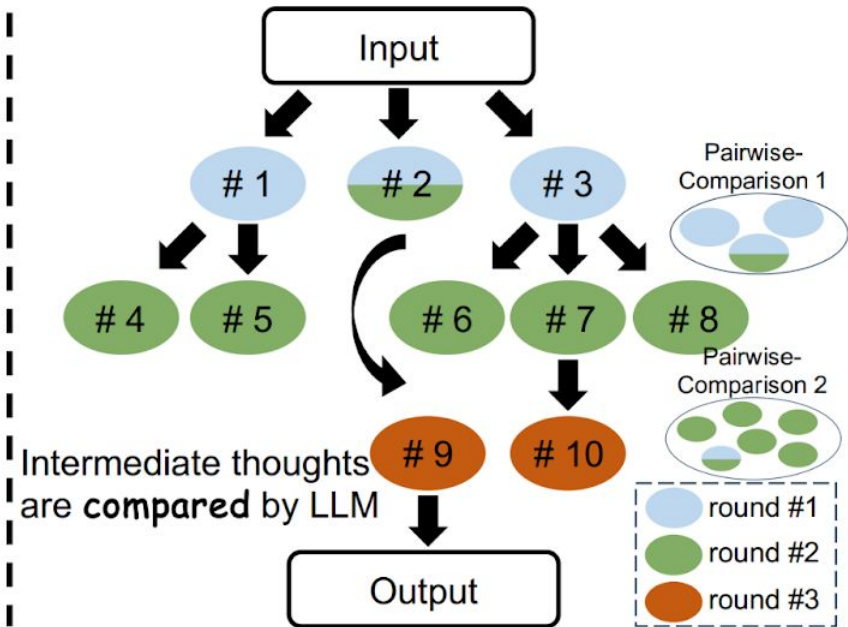
Chain-of-Thoughts



Score-based Tree-of-Thoughts (S-ToT)

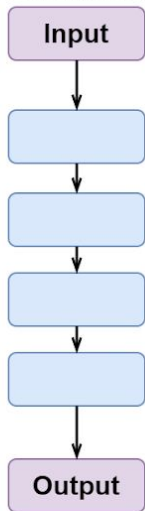


Comparison-based Tree-of-Thoughts (C-ToT)



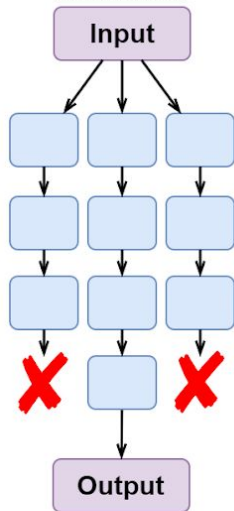
## Chain-of-Thought (CoT)

Wei et al., Jan 2022  
Kojima et al., May 2022



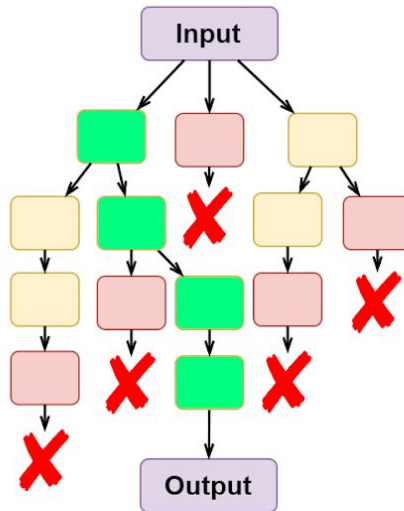
## Multiple Chains-of-Thought

Wang et al.,  
March 2022



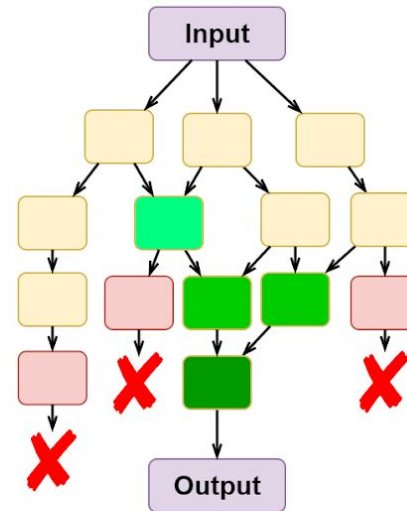
## Tree-of-Thought (ToT)

Long, May 2023  
Yao et al., May 2023

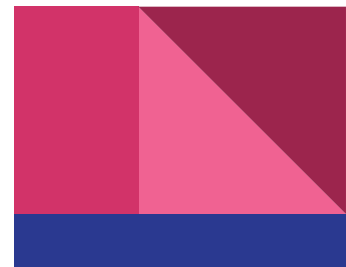


## Graph-of-Thought (GoT)

Besta et al., August 2023  
Lei et al., August 2023



- Итого получается так

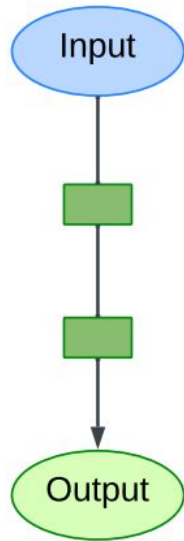


# Algorithm-of-Thought

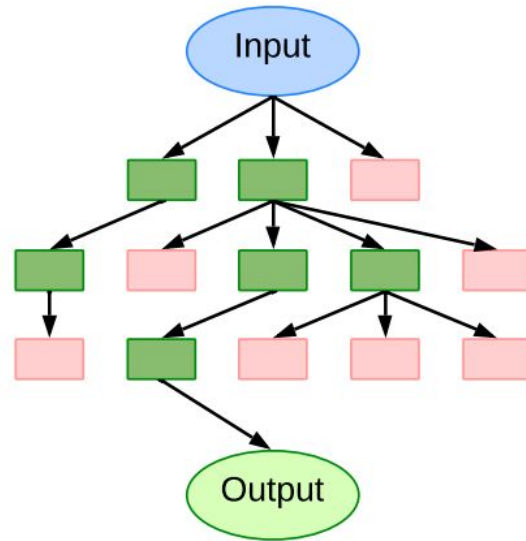
- [Sel et al. \(2023\)](#): Algorithm-of-Thought; каждый узел дерева – шаг в алгоритме, и модель порождает следующий шаг по результатам предыдущего



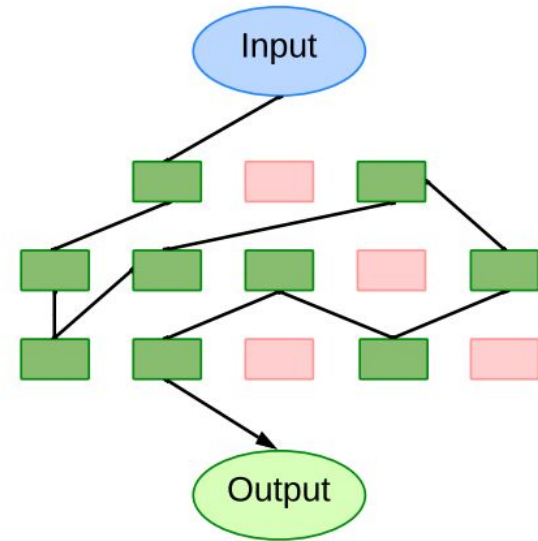
Standard Prompting



Chain of Thoughts



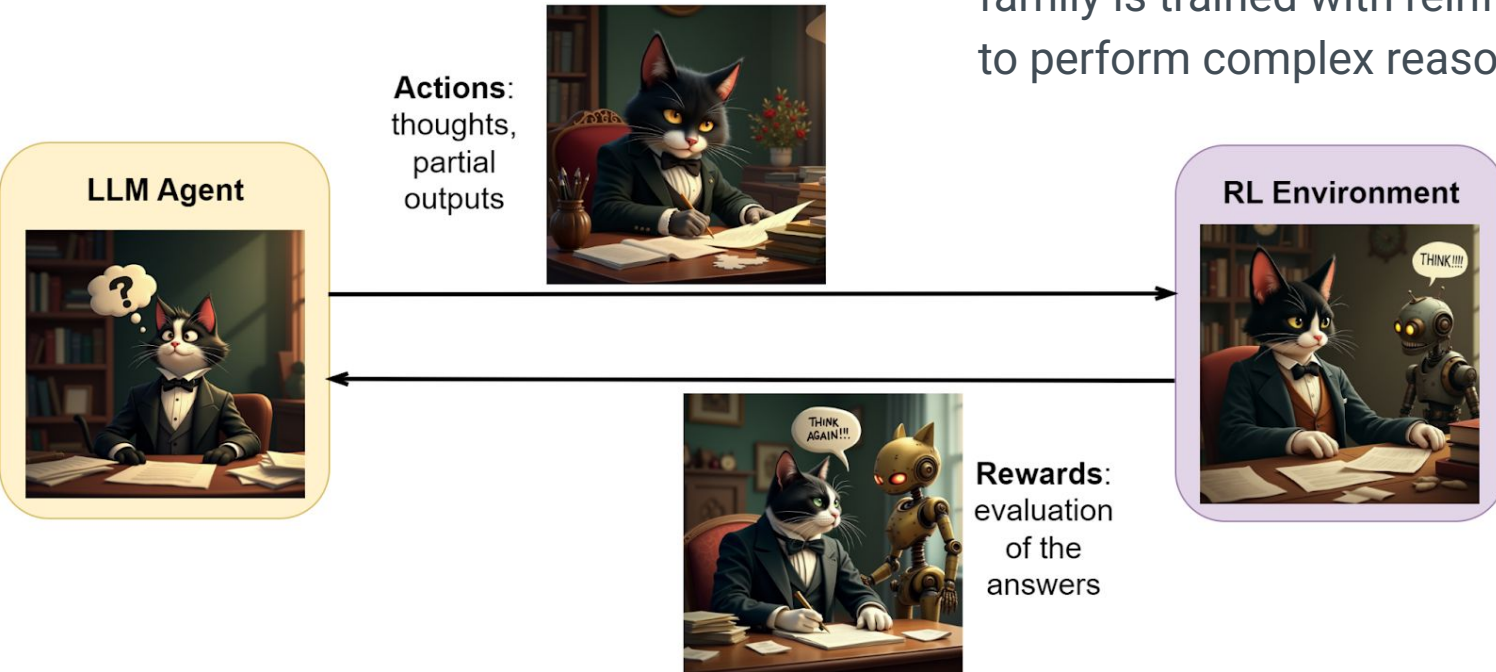
Tree of Thoughts



Algorithm of Thoughts

# Семейство моделей o1

- Но всё изменилось, когда появилось семейство моделей o1 от OpenAI
- Детали не разглашались, system card говорила так: “The o1 large language model family is trained with reinforcement learning to perform complex reasoning”; как это?



# Process Reward Models

- [Uesato et al. \(2022\)](#): обычно в RL можно награждать только конечный результат; но в chain-of-thought мышлении LLM выводит мысли, которые в принципе можно оценивать самостоятельно
- Process reward model (PRM) обучается оценивать промежуточные шаги

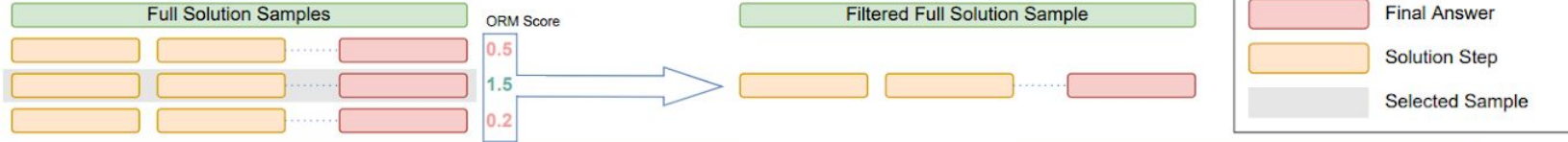
## Final Answer RL

GSM8K Problems



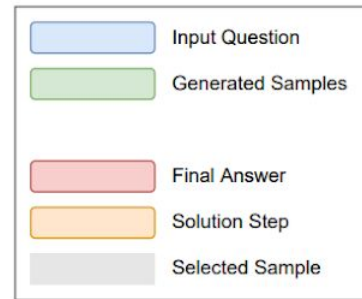
## ORM-RL

GSM8K Problems



## PRM-RL

GSM8K Problems





# Process Reward Models

- [Lightman et al. \(OpenAI, 2024\)](#):




Let's Verify  
Step by Step –  
давайте  
обучим PRM  
выяснять, где  
именно в  
рассуждениях  
ошибка

The denominator of a fraction is 7 less than 3 times the numerator. If the fraction is equivalent to  $2/5$ , what is the numerator of the fraction? (Answer: )

   Let's call the numerator  $x$ .

   So the denominator is  $3x-7$ .

   We know that  $x/(3x-7) = 2/5$ .

   So  $5x = 2(3x-7)$ .

    $5x = 6x - 14$ .

   So  $x = 7$ .

# Process Reward Models

- [Lightman et al. \(OpenAI, 2024\)](#):

Let's Verify  
 Step by Step –  
 давайте  
 обучим PRM  
 выяснить, где  
 именно в  
 рассуждениях  
 ошибка

Let

$$x^8 + 3x^4 - 4 = p_1(x)p_2(x) \cdots p_k(x),$$

where each non-constant polynomial  $p_i(x)$  is monic with integer coefficients, and cannot be factored further over the integers. Compute  $p_1(1) + p_2(1) + \cdots + p_k(1)$ .

I notice that the given polynomial has even degree and only even powers of  $x$ , so I can try to make a substitution to simplify it.

Let  $y = x^4$ , then the polynomial becomes  $y^2 + 3y - 4$ , which is a quadratic equation.

I can factor this quadratic equation as  $(y + 4)(y - 1)$ , so the original polynomial is  $(x^4 + 4)(x^4 - 1)$ .

Now I need to factor each of these factors further, if possible, over the integers.

For the first factor,  $x^4 + 4$ , I recall the Sophie Germain identity, which says that  $a^4 + 4b^4 = (a^2 + 2b^2 + 2ab)(a^2 + 2b^2 - 2ab)$ .

Applying this identity with  $a = x$  and  $b = 1$ , I get  $x^4 + 4 = (x^2 + 2x + 2)(x^2 - 2x + 2)$ .

For the second factor,  $x^4 - 1$ , I can use the difference of squares formula, which says that  $a^2 - b^2 = (a + b)(a - b)$ .

Applying this formula with  $a = x^2$  and  $b = 1$ , I get  $x^4 - 1 = (x^2 + 1)(x^2 - 1)$ .

Then I can apply the difference of squares formula again with  $a = x$  and  $b = 1$ , to get  $x^2 - 1 = (x + 1)(x - 1)$ .

So the original polynomial is  $(x^4 + 4)(x^4 - 1) = (x^2 + 2x + 2)(x^2 - 2x + 2)(x^2 + 1)(x + 1)(x - 1)$ .

These are all the irreducible factors over the integers, so  $p_1(x) = x^2 + 2x + 2$ ,  $p_2(x) = x^2 - 2x + 2$ ,  $p_3(x) = x^2 + 1$ ,  $p_4(x) = x + 1$ , and  $p_5(x) = x - 1$ .

To find the sum of their values at  $x = 1$ , I just need to plug in  $x = 1$  into each factor and add them up.

I get  $p_1(1) = 1^2 + 2 \cdot 1 + 2 = 5$ ,  $p_2(1) = 1^2 - 2 \cdot 1 + 2 = 1$ ,  $p_3(1) = 1^2 + 1 = 2$ ,  $p_4(1) = 1 + 1 = 2$ , and  $p_5(1) = 1 - 1 = 0$ .

So the final answer is  $5 + 1 + 2 + 2 + 0 = 10$ .

Answer: 10

Let

$$x^8 + 3x^4 - 4 = p_1(x)p_2(x) \cdots p_k(x),$$

where each non-constant polynomial  $p_i(x)$  is monic with integer coefficients, and cannot be factored further over the integers. Compute  $p_1(1) + p_2(1) + \cdots + p_k(1)$ .

This problem is about factoring a polynomial over the integers, which can be a tricky task.

One possible strategy is to look for rational roots of the polynomial, using the rational root theorem.

However, that would involve checking a lot of possible factors of  $-4$ , and there is no guarantee that the polynomial has any rational roots at all.

Another possible strategy is to use some algebraic tricks to simplify the polynomial or reduce its degree.

I notice that the polynomial has only even powers of  $x$ , so I can try to substitute  $y = x^2$  and see what happens.

Then the polynomial becomes  $y^4 + 3y^2 - 4$ , which is a quadratic equation in  $y^2$ .

I can try to factor this quadratic equation by looking for two numbers that multiply to  $-4$  and add to  $3$ .

Those numbers are  $4$  and  $-1$ , so I can write  $y^4 + 3y^2 - 4 = (y^2 + 4)(y^2 - 1)$ .

Now I can substitute back  $x^2 = y$  and get  $x^8 + 3x^4 - 4 = (x^4 + 4)(x^4 - 1)$ .

This looks promising, because I have reduced the degree of the factors from  $8$  to  $4$ .

I can try to factor each of these factors further, using the difference of squares formula.

I get  $x^8 + 3x^4 - 4 = (x^2 + 2)(x^2 - 2)(x^2 + 1)(x^2 - 1)$ .

I can apply the difference of squares formula again to the last factor and get  $x^8 + 3x^4 - 4 = (x^2 + 2)(x^2 - 2)(x^2 + 1)(x + 1)(x - 1)$ .

Now I have factored the polynomial completely into monic linear and quadratic factors with integer coefficients.

These are the  $p_i(x)$ 's that the problem is asking for.

To find the sum of their values at  $x = 1$ , I just need to plug in  $x = 1$  into each factor and add them up.

I get  $p_1(1) + p_2(1) + \cdots + p_k(1) = (1^2 + 2)(1^2 - 2)(1^2 + 1)(1 + 1)(1 - 1)$ .

Simplifying, I get  $p_1(1) + p_2(1) + \cdots + p_k(1) = (3)(-1)(2)(2)(0)$ .

Multiplying, I get  $p_1(1) + p_2(1) + \cdots + p_k(1) = 0$ .

Answer: 0


# Process Reward Models

- [Xia et al. \(2024\)](#): ReasonEval с тем же СМЫСЛОМ
- А что делает o1? Неизвестно, и люди сразу попытались реплицировать

Question: Simplify  $\sqrt{242}$

Ground Truth:  $11\sqrt{2}$

Solution



ReasonEval

$S_{\text{validity}}$	$S_{\text{redundancy}}$
0.85	0.05
0.85	0.05
0.75	0.25
0.95	0.05
0.80	0.05
0.15	0.05
0.75	0.05

**Step 1:** We can simplify the square root of 242 by factoring it into its prime factors.

**Step 2:** The prime factorization of 242 is  $2 * 11 * 11$ .

**Step 3:** Let's check it. The prime factorization of 242 is  $2 * 11 * 11$ . ● ←-----

**Step 4:** By the property of square roots, we know that  $\sqrt{ab} = \sqrt{a} * \sqrt{b}$  for any non-negative numbers  $a$  and  $b$ . Applying this property to  $\sqrt{2 * 11 * 11}$ , we have  $\sqrt{2} * \sqrt{11^2}$ .


**Step 5:** By the property of square roots, we know that  $\sqrt{a^2} = a$  for any non-negative number  $a$ . Applying this property to  $\sqrt{11^2}$ , we have 11.

**Step 6:** By combining the results from step 4 and step 5, we have  $\sqrt{2} * 11 * 11 = 11 * \sqrt{2}$ . ● ←-----

**Step 7:** Therefore, the simplified form of  $\sqrt{242}$  is  $11\sqrt{2}$ .

● redundant    ● invalid

Answer:  $11\sqrt{2}$

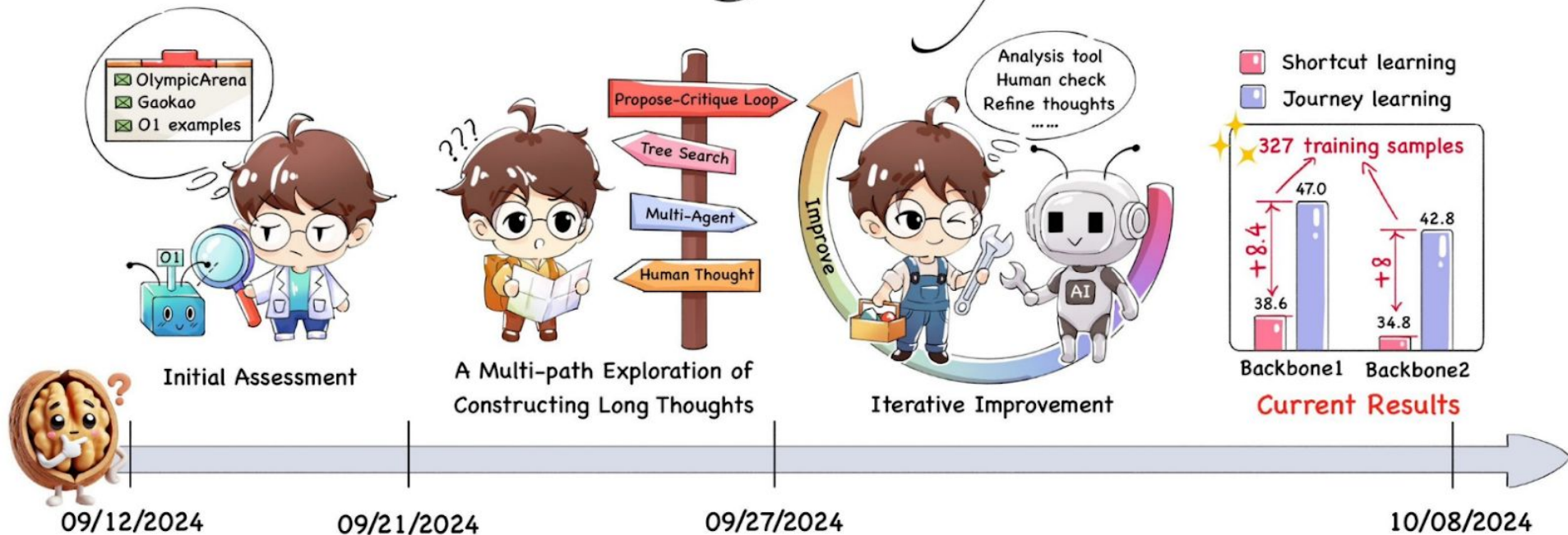




# Пример: O1 Journey

- Пример – Generative AI Research Lab (G AIR) из Шанхая; [Qin et al. \(Oct 2024\)](#):

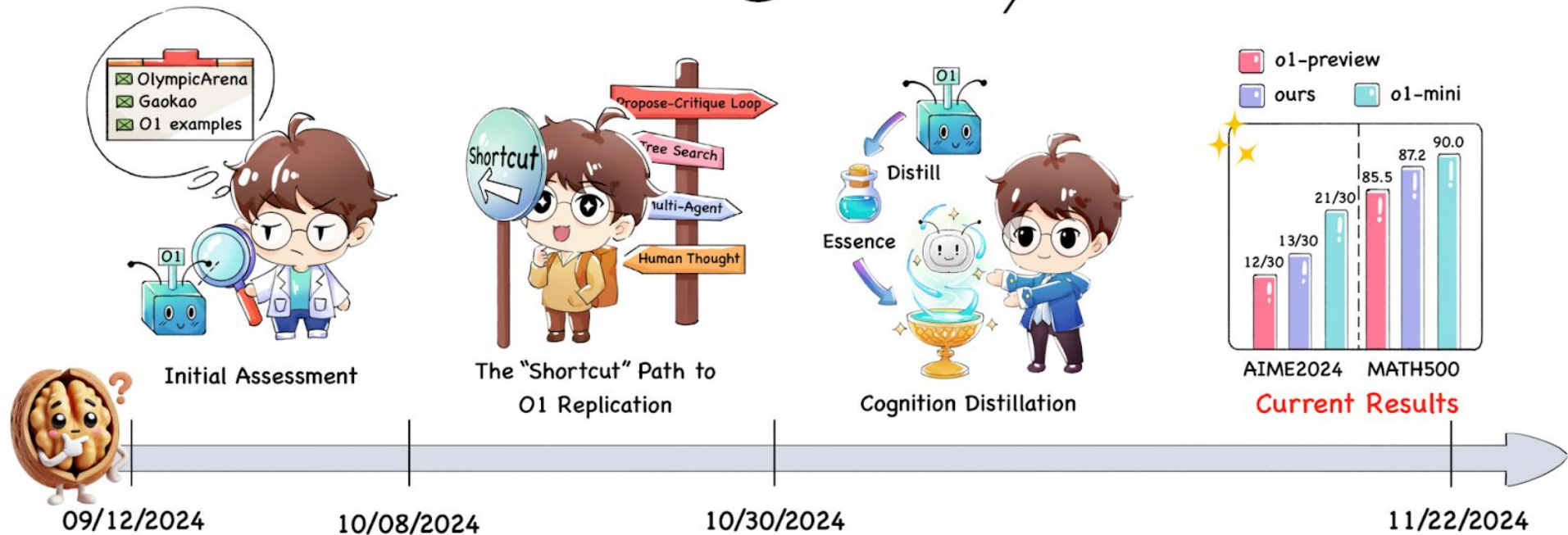
## O1 Journey



# Пример: O1 Journey

- Пример – Generative AI Research Lab (G AIR) из Шанхая; [Huang et al. \(Nov 2024\)](#):

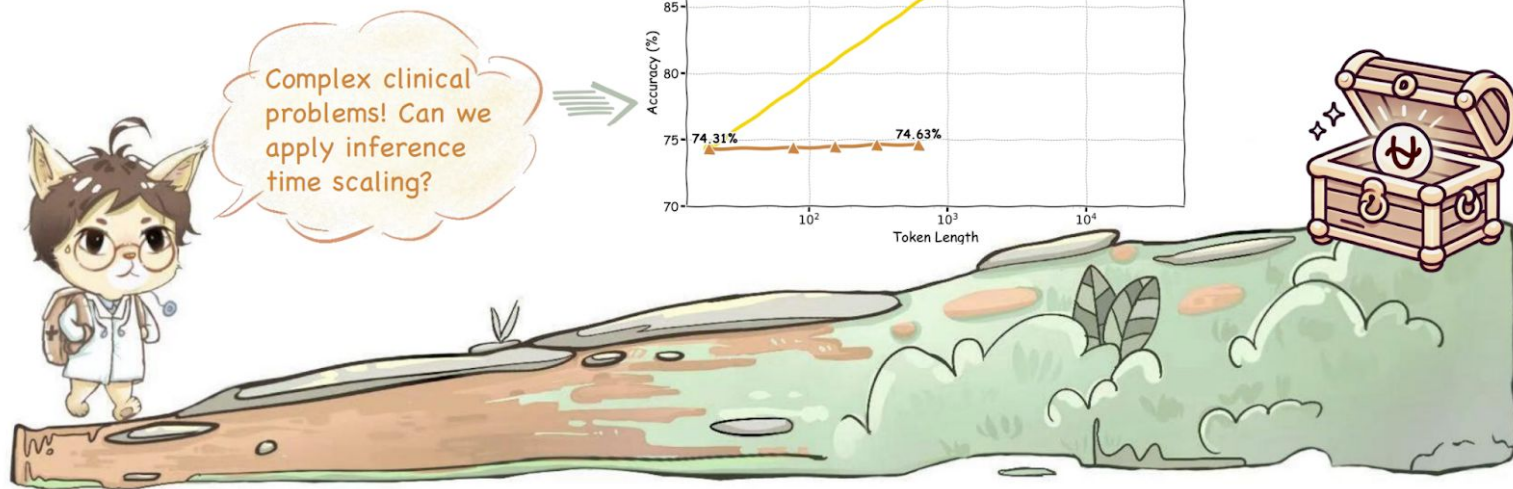
## O1 Journey



# Пример: O1 Journey

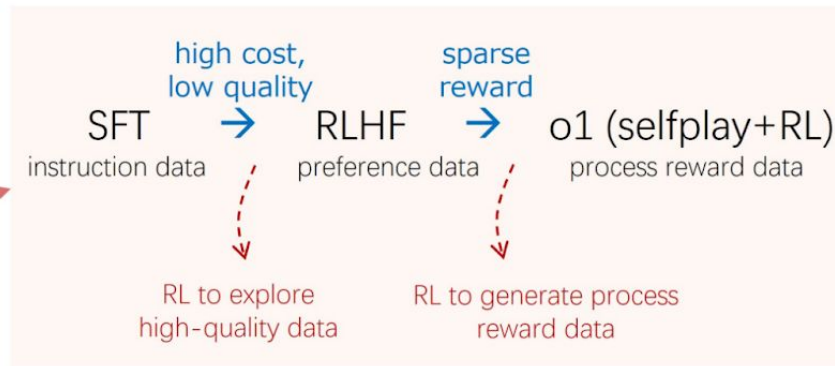
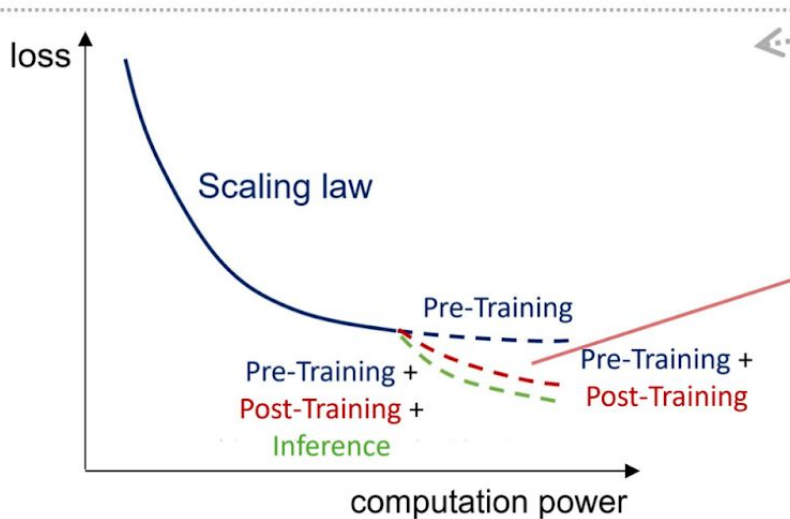
- Пример – Generative AI Research Lab (G AIR) из Шанхая; [Huang et al. \(Jan 2025\)](#):

## O1 Journey



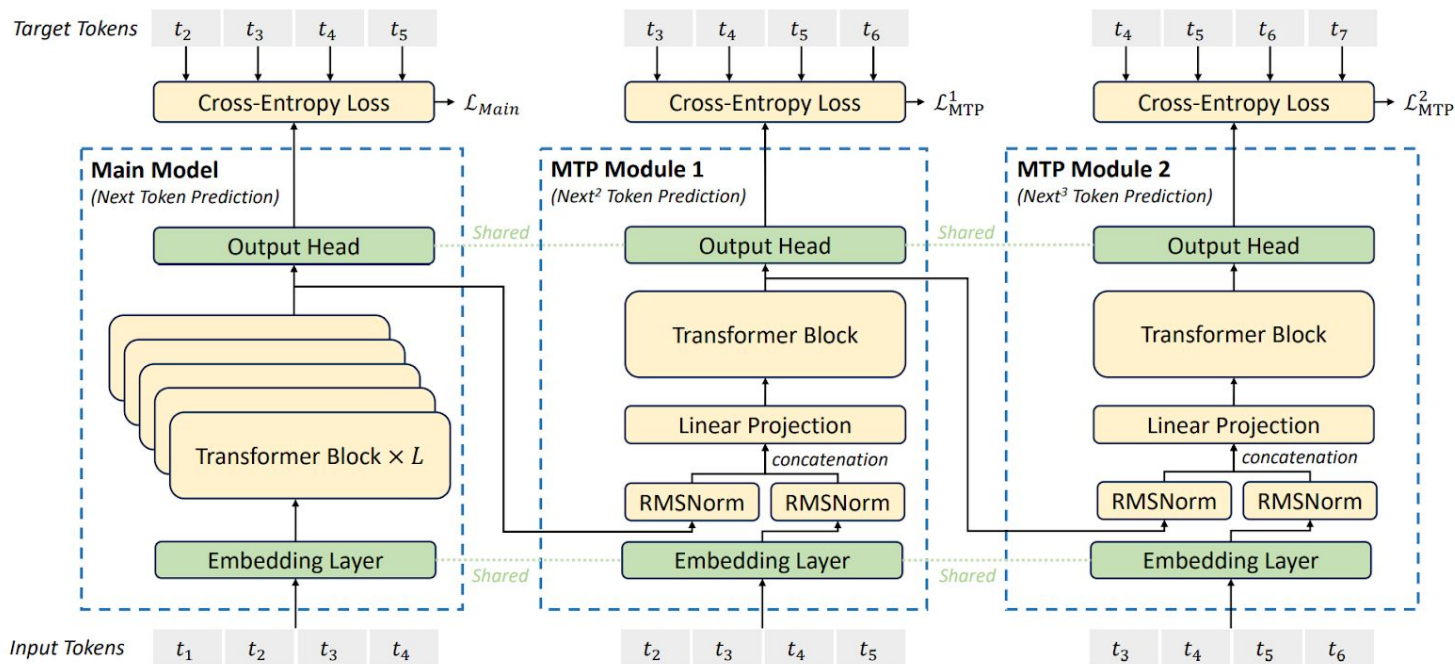
# Пример: O1 Journey

- [Zhang et al. \(2024\)](#): o1-Coder, добавили PRM и MCTS

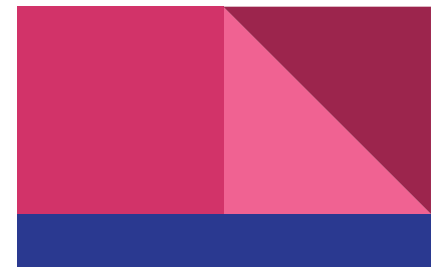


# DeepSeek-V3

- А что же делал наш кит в комнате, то есть DeepSeek?
- Сначала DeepSeek-V3; multi-token prediction:

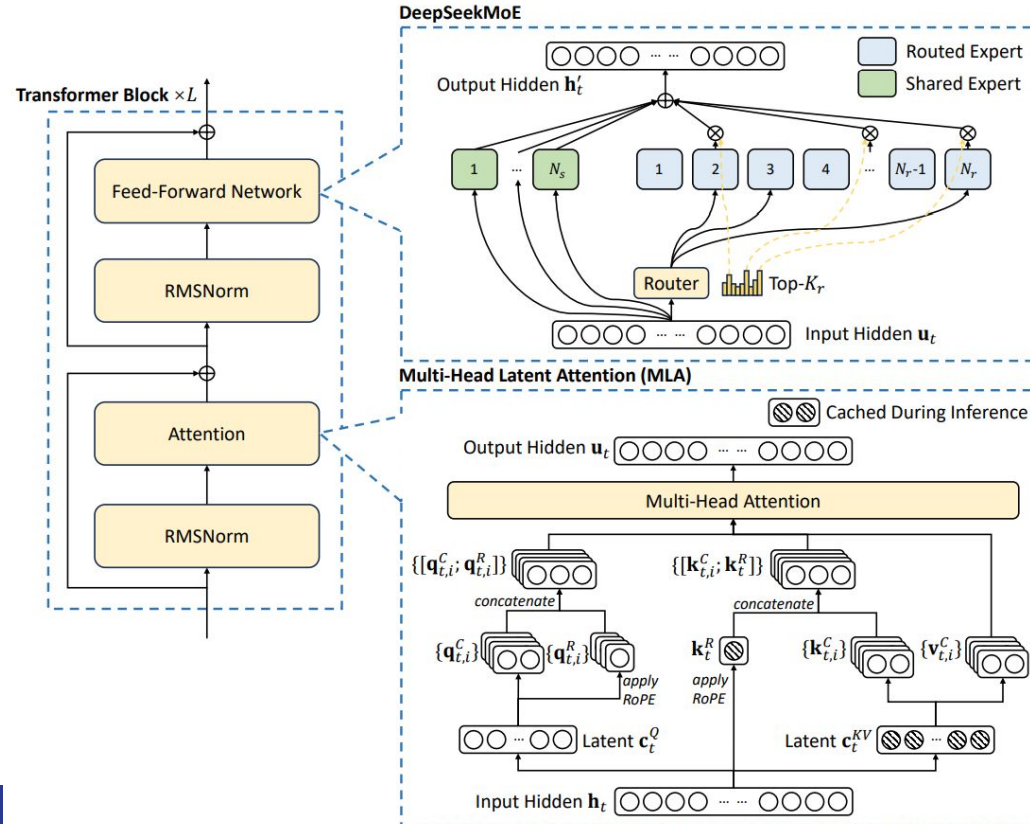


...



# DeepSeek-V3

- А что же делал наш кит в комнате, то есть DeepSeek?
- Сначала DeepSeek-V3
- Key-value caching (не будем углубляться), mixture-of-experts
- В общем, хорошая современная LLM

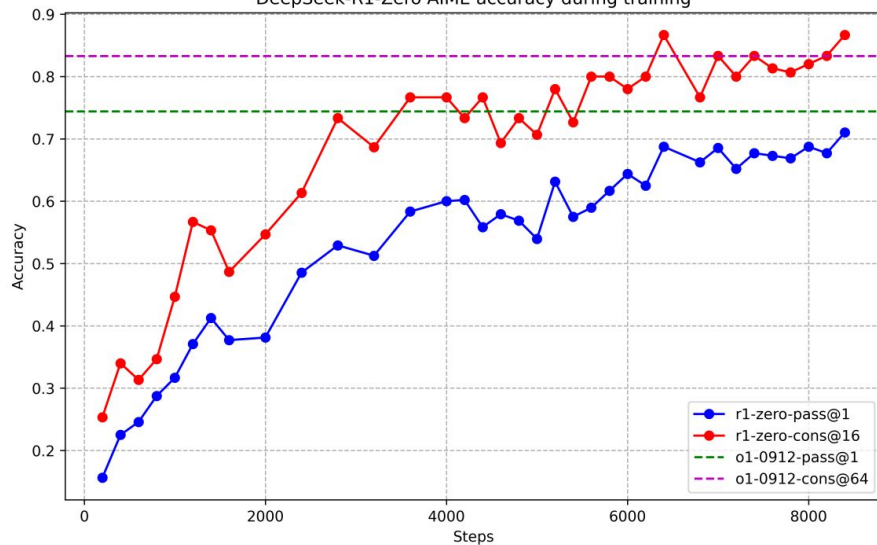


# DeepSeek-R1

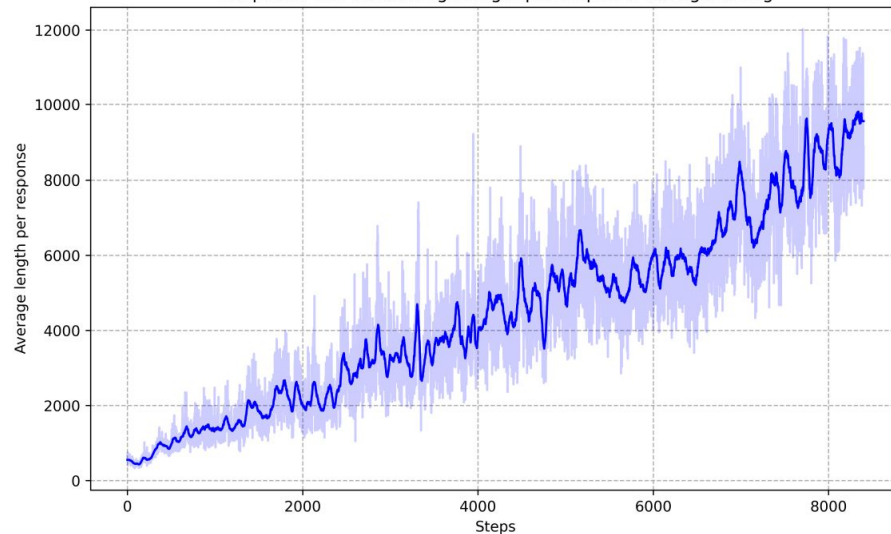
- А потом и DeepSeek-R1
- R1-Zero: чистый RL на проверяемых задачах



DeepSeek-R1-Zero AIME accuracy during training

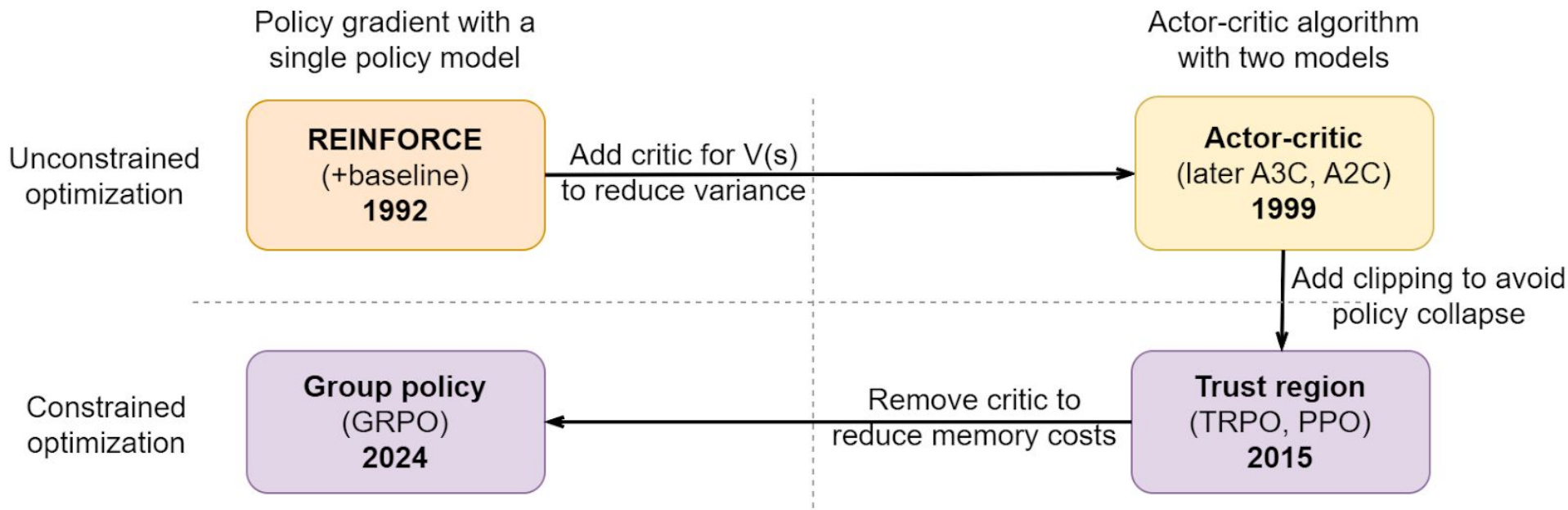


DeepSeek-R1-Zero average length per response during training



# DeepSeek-R1

- А потом и DeepSeek-R1
- Кстати, в RL у DeepSeek тоже было новшество: GRPO





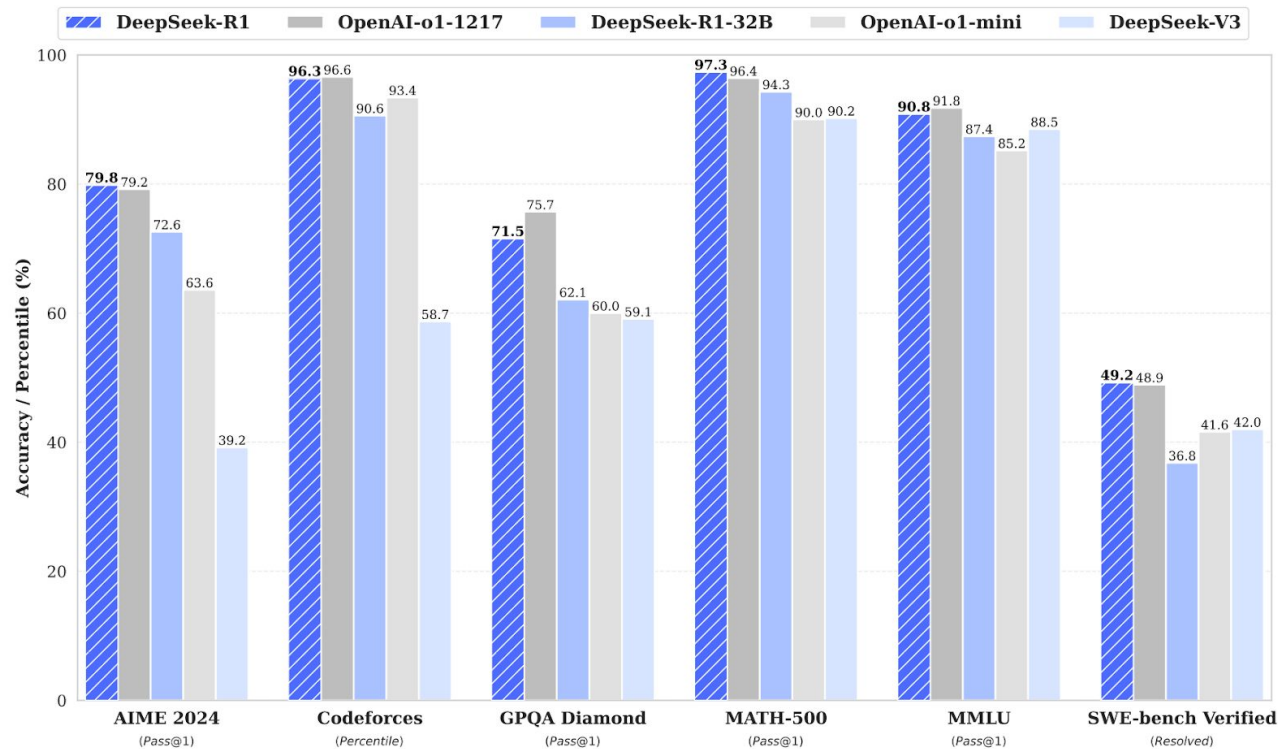
# DeepSeek-R1

- R1: сначала соберём маленький chain-of-thought датасет (несколько тысяч примеров) для cold start
- Потом уже RL на полученной модели
- После сходимости RL ещё соберём датасет из reasoning traces этой модели, отфильтруем по читабельности и каким-то внешним признакам
- Сделаем SFT на этом датасете
- А потом ещё RLHF как обычно поверх



# DeepSeek-R1

- Результаты хорошие, и вообще модель получилась хорошая



# DeepSeek-R1

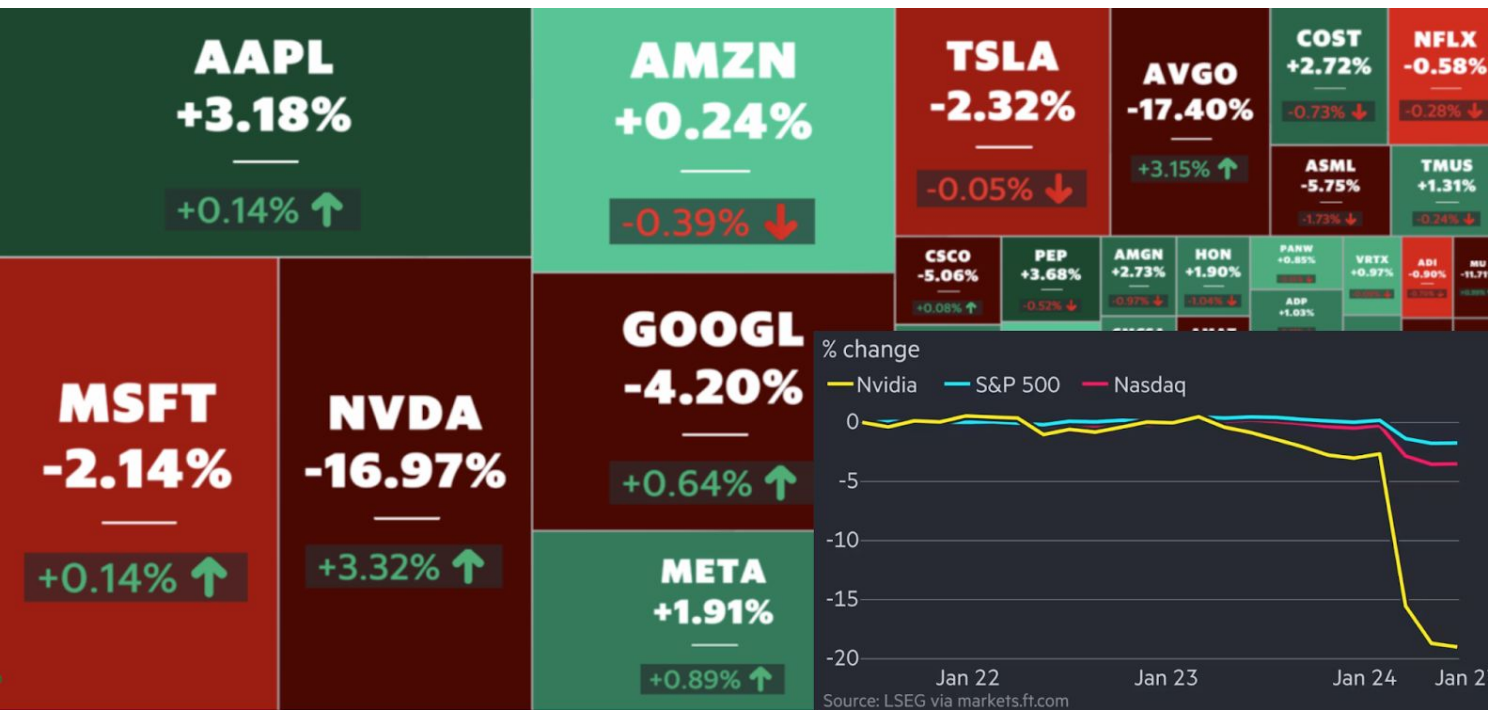
- А датасет из 800K reasoning traces, который у них получился, можно использовать для дистилляции, тоже очень успешно
- Опять тот же горький урок: distill, baby, distill...



Model	AIME 2024		MATH-500	GPQA Diamond	LiveCode Bench	CodeForces
	pass@1	cons@64	pass@1	pass@1	pass@1	rating
GPT-4o-0513	9.3	13.4	74.6	49.9	32.9	759
Claude-3.5-Sonnet-1022	16.0	26.7	78.3	65.0	38.9	717
OpenAI-o1-mini	63.6	80.0	90.0	60.0	53.8	<b>1820</b>
QwQ-32B-Preview	50.0	60.0	90.6	54.5	41.9	1316
DeepSeek-R1-Distill-Qwen-1.5B	28.9	52.7	83.9	33.8	16.9	954
DeepSeek-R1-Distill-Qwen-7B	55.5	83.3	92.8	49.1	37.6	1189
DeepSeek-R1-Distill-Qwen-14B	69.7	80.0	93.9	59.1	53.1	1481
DeepSeek-R1-Distill-Qwen-32B	<b>72.6</b>	83.3	94.3	62.1	57.2	1691
DeepSeek-R1-Distill-Llama-8B	50.4	80.0	89.1	49.0	39.6	1205
DeepSeek-R1-Distill-Llama-70B	70.0	<b>86.7</b>	<b>94.5</b>	<b>65.2</b>	<b>57.5</b>	1633

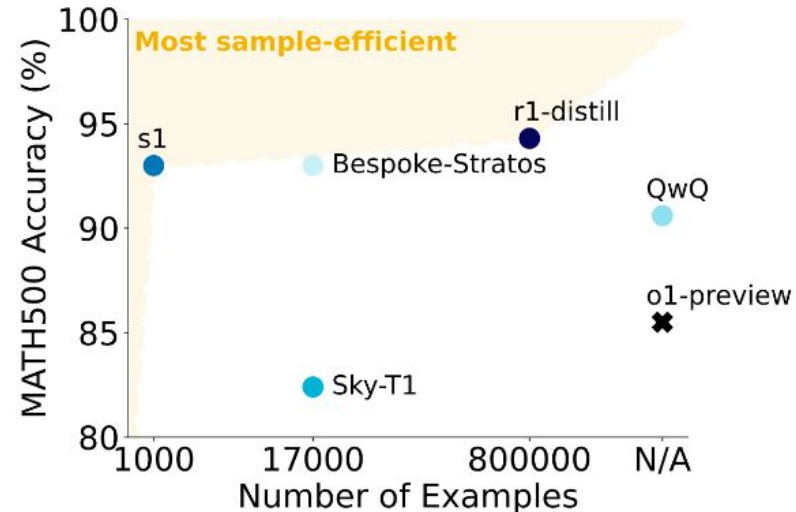
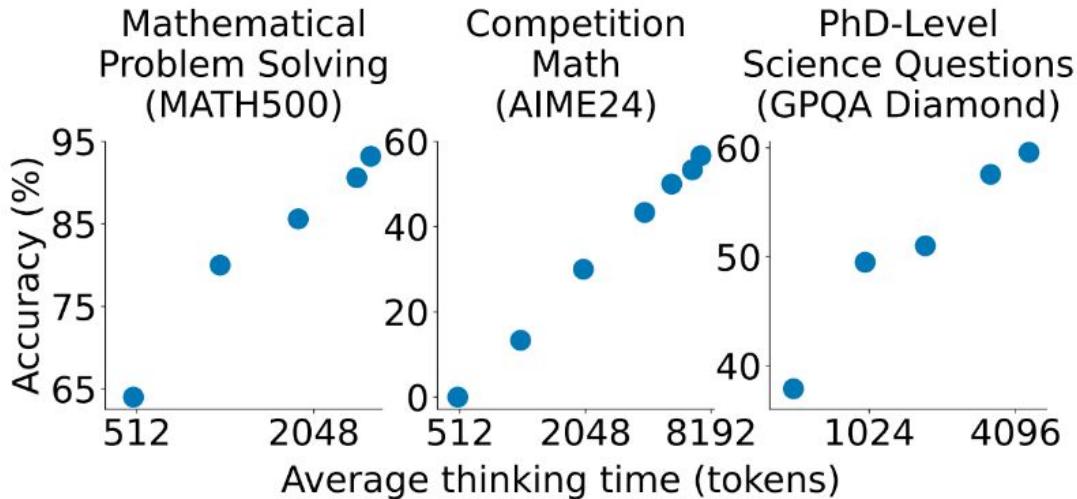
# DeepSeek-R1

- Самое, конечно, поразительное во всём этом – это реакция рынка :) но это уже другой разговор



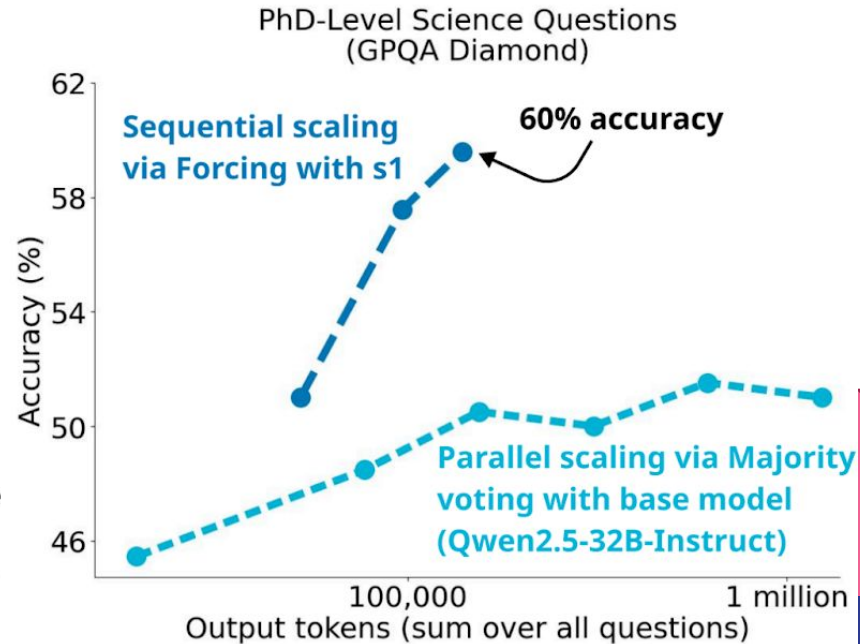
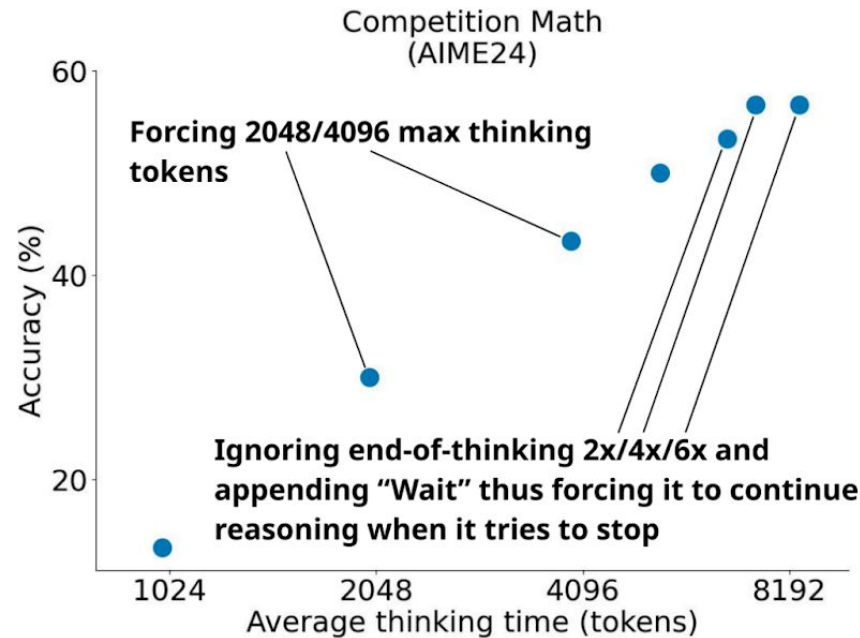
# s1-32B: distill, baby, distill...

- [Mueninghoff et al. \(2025\)](#): s1-32B, дистилляция на 1000 примеров за \$50
- Получились очень хорошие результаты в рассуждениях



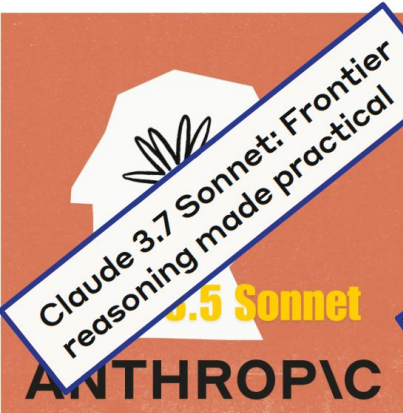
# s1-32B: distill, baby, distill...

- [Mueninghoff et al. \(2025\)](#): s1-32B, дистилляция на 1000 примеров за \$50
- И очень простые идеи – budget forcing и голосование – отлично работают

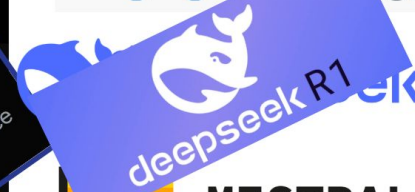


# Итоги

- Рассуждающие модели сейчас везде
- Важный вывод из DeepSeek-R1 — то, что там не используется: ни PRM, ни MCTS... оказалось, что достаточно RL хорошо сложить с SFT и RLHF
- Но что будет дальше?..



LLaMA  
by Meta



MISTRAL  
AI



# 9. Сети Колмогорова-Арнольда



[Наука] лишает нас чувства почтительности, ибо чтят лишь вещи, на которые не смеют поднять глаз. Выводы науки – не равносильны ли они отрицанию нравственности?

*Анри Пуанкаре. Наука и нравственность*



# Теорема Колмогорова–Арнольда

- **13-я проблема Гильберта:** рассмотрим уравнение седьмой степени в приведённой форме

$$x^7 + ax^3 + bx^2 + cx + 1 = 0.$$

- Решение этого уравнения – функция от  $a$ ,  $b$  и  $c$ . Можно ли её выразить как композицию конечного числа функций от двух переменных?
- Число коэффициентов дальше уже не сократить, поэтому кажется логичным, что может быть и нельзя; для алгебраических функций проблема до сих пор не решена.
- А вот для непрерывных...

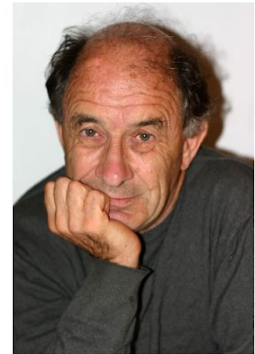
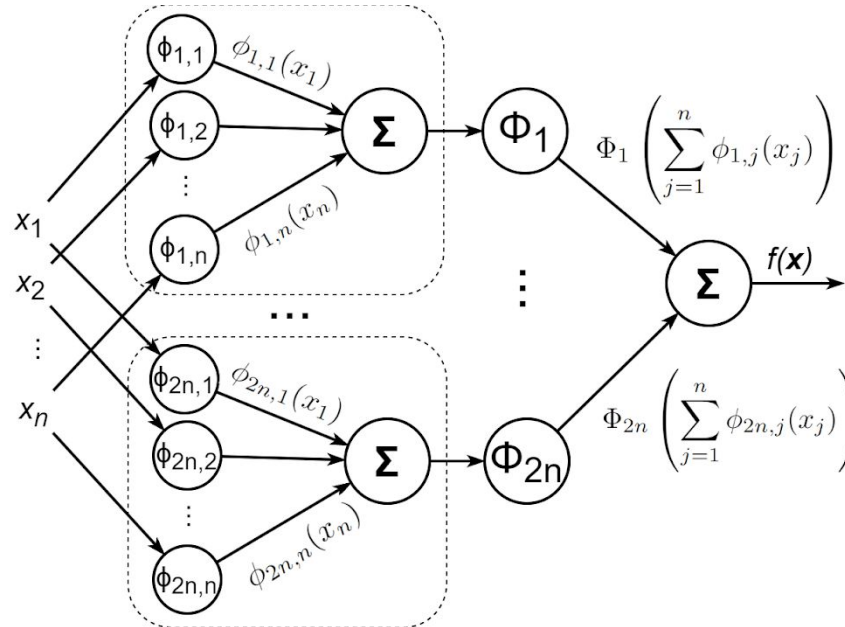
# Теорема Колмогорова–Арнольда

- Каждая многомерная непрерывная функция может быть представлена в виде суперпозиции непрерывных функций **одной** переменной (и сложения)!

$$f(\mathbf{x}) = f(x_1, \dots, x_n) = \sum_{i=0}^{2n} \Phi_i \left( \sum_{j=1}^n \phi_{i,j}(x_j) \right)$$



Andrey  
Kolmogorov



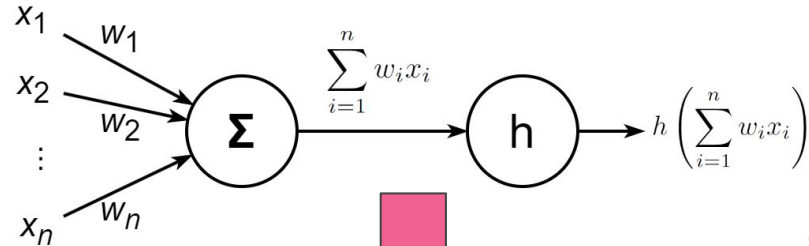
Vladimir  
Arnold

# Можно ли построить сеть на этом?

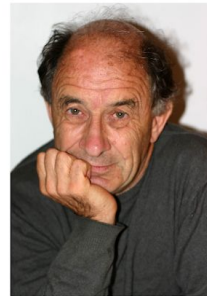
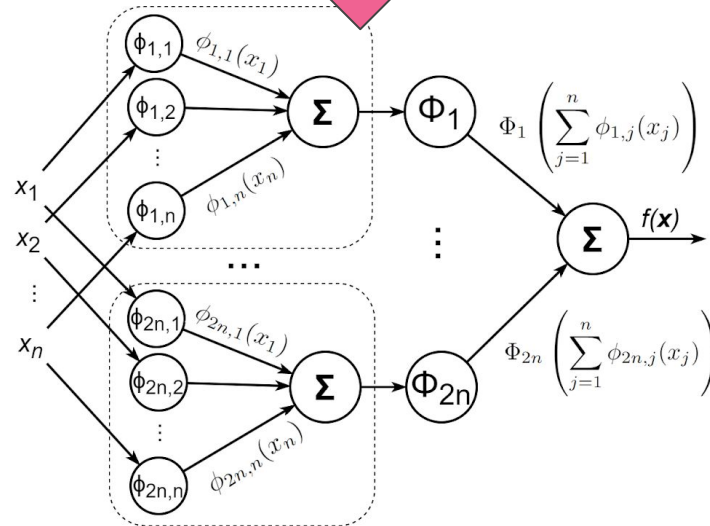
- Было бы интересно попробовать построить на этом сети — одномерную-то функцию мы как-нибудь построим
- Есть конструктивное доказательство ([Sprecher, 1965; 1972](#)), но есть и результаты о том, что функции будут плохие ([Girosi, Poggio, 1989](#))



Walter Pitts

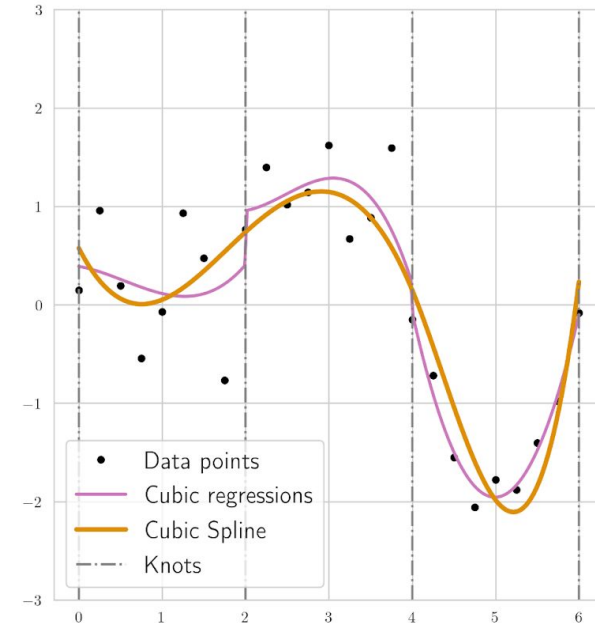
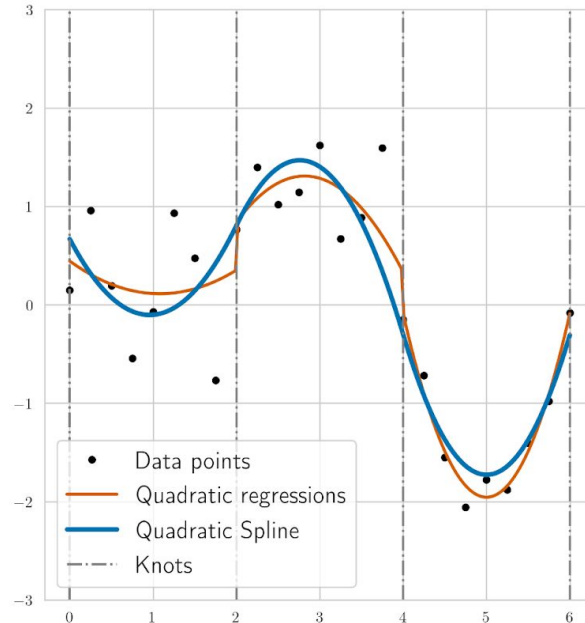


Warren McCulloch


 Andrey  
 Kolmogorov

 Vladimir  
 Arnold

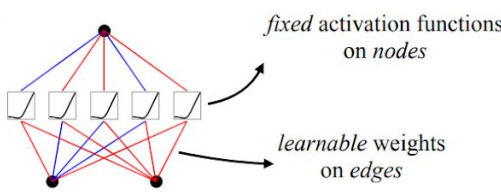
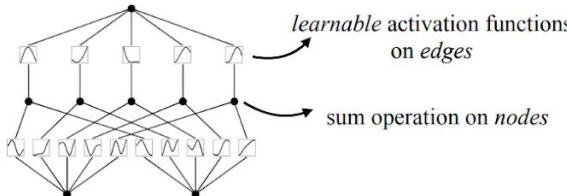
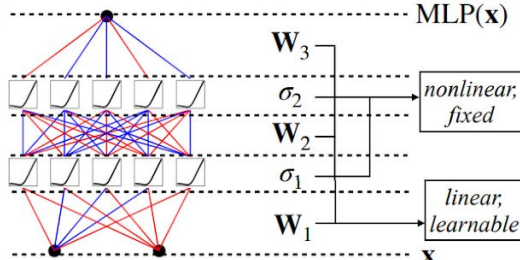
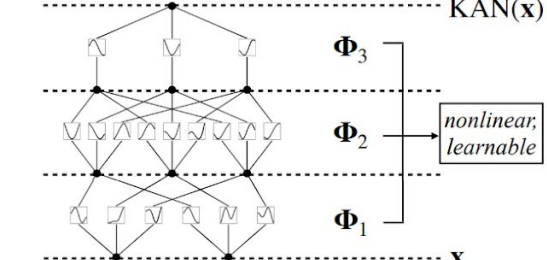
# Сплайны

- Поэтому конструктивные подходы сокращают множество функций и увеличивают их число и/или глубину сети
- Главный инструмент – *сплайны*, гладкие кусочно-полиномиальные функции



# Сети Колмогорова–Арнольда

- Kolmogorov–Arnold networks (KAN; [Liu et al., 2024](#))
- Фактически обучаемые функции активации для нейронов, и это B-сплайны

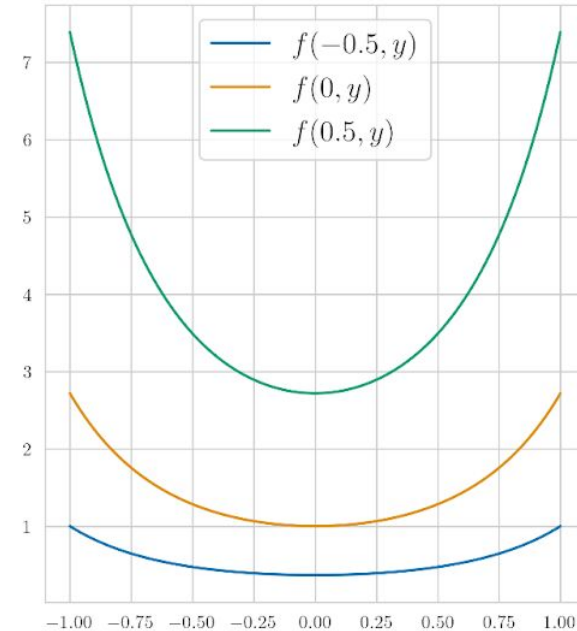
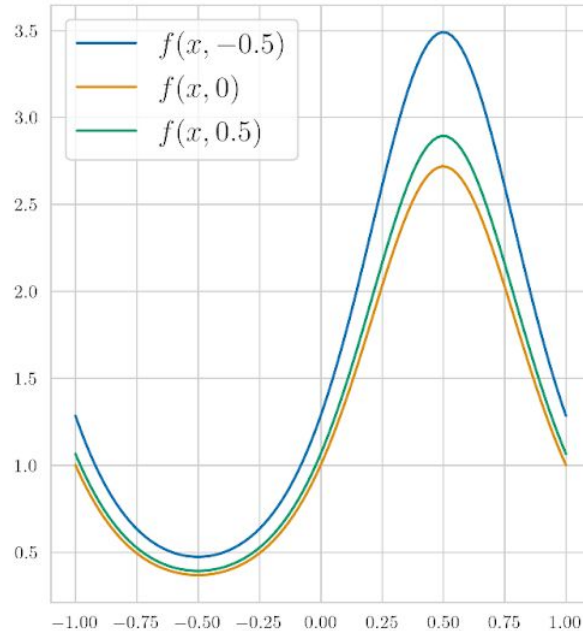
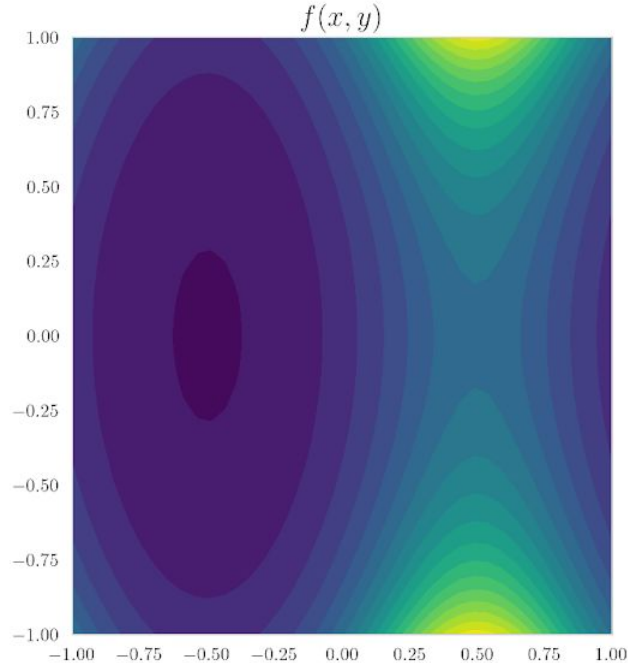
Model	<b>Multi-Layer Perceptron (MLP)</b>	<b>Kolmogorov-Arnold Network (KAN)</b>
Theorem	<b>Universal Approximation Theorem</b>	<b>Kolmogorov-Arnold Representation Theorem</b>
Formula (Shallow)	$f(\mathbf{x}) \approx \sum_{i=1}^{N(e)} a_i \sigma(\mathbf{w}_i \cdot \mathbf{x} + b_i)$	$f(\mathbf{x}) = \sum_{q=1}^{2n+1} \Phi_q \left( \sum_{p=1}^n \phi_{q,p}(x_p) \right)$
Model (Shallow)	<b>(a)</b>  <p>fixed activation functions on nodes</p> <p>learnable weights on edges</p>	<b>(b)</b>  <p>learnable activation functions on edges</p> <p>sum operation on nodes</p>
Formula (Deep)	$\text{MLP}(\mathbf{x}) = (\mathbf{W}_3 \circ \sigma_2 \circ \mathbf{W}_2 \circ \sigma_1 \circ \mathbf{W}_1)(\mathbf{x})$	$\text{KAN}(\mathbf{x}) = (\Phi_3 \circ \Phi_2 \circ \Phi_1)(\mathbf{x})$
Model (Deep)	<b>(c)</b>  <p>MLP(x)</p> <p><math>\mathbf{W}_3</math></p> <p><math>\sigma_2</math> → nonlinear, fixed</p> <p><math>\mathbf{W}_2</math></p> <p><math>\sigma_1</math> → linear, learnable</p> <p><math>\mathbf{W}_1</math></p> <p><math>\mathbf{x}</math></p>	<b>(d)</b>  <p>KAN(x)</p> <p><math>\Phi_3</math></p> <p><math>\Phi_2</math> → nonlinear, learnable</p> <p><math>\Phi_1</math></p> <p><math>\mathbf{x}</math></p>

$$\phi(x) = w_b \cdot b(x) + w_s \cdot s(x),$$

$$b(x) = x / (1 + e^{-x}), \quad s(x) = \sum_i c_i B_i(x).$$

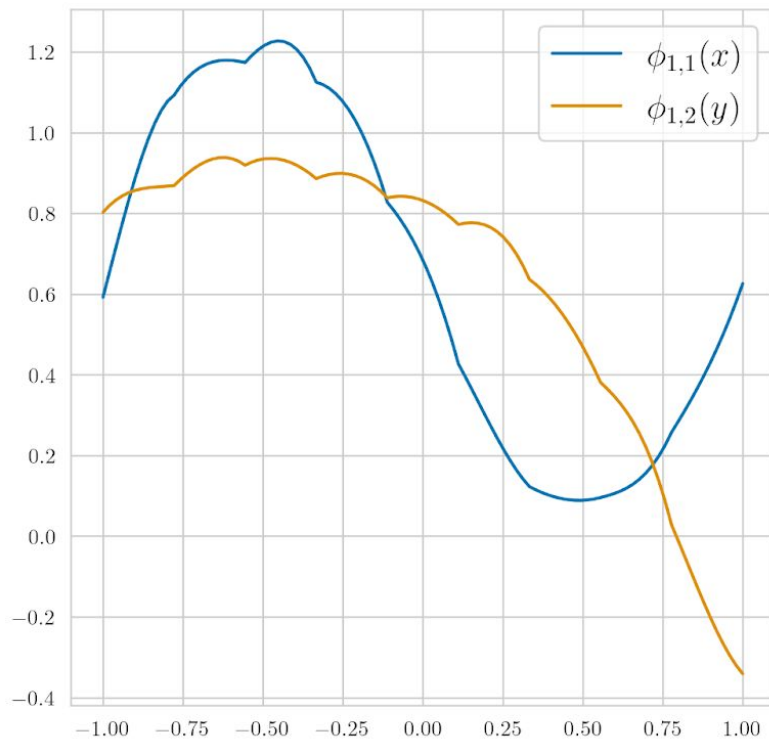
# Сети Колмогорова–Арнольда

- Пример: давайте приблизим функцию  $f(x, y) = \exp(\sin(\pi x) + y^2)$ .
- Возьмём [2, 2, 1] KAN:  $\hat{f}(x, y) = \phi_{2,1}(\phi_{1,1}(x) + \phi_{1,2}(y))$ .

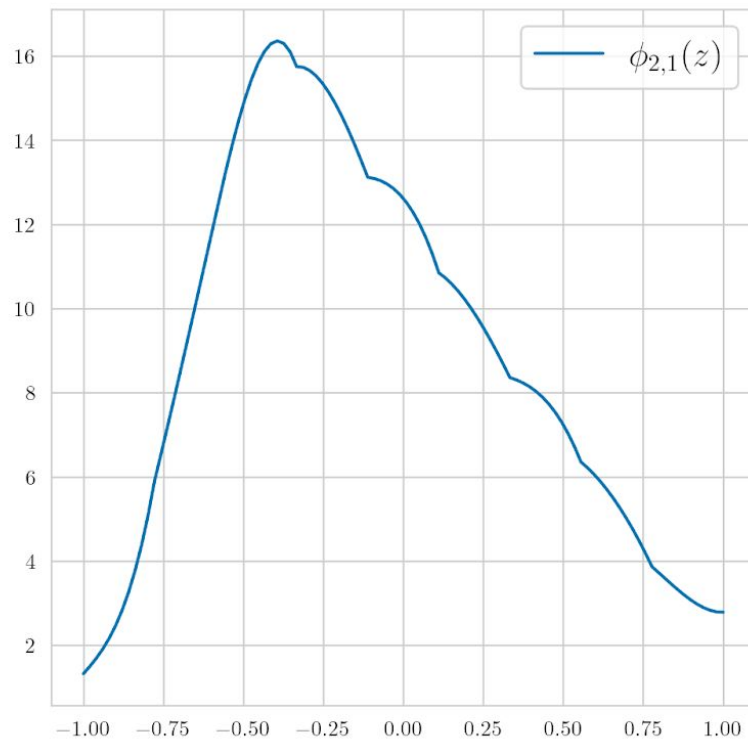


# Сети Колмогорова–Арнольда

- После обучения функции получаются такие:



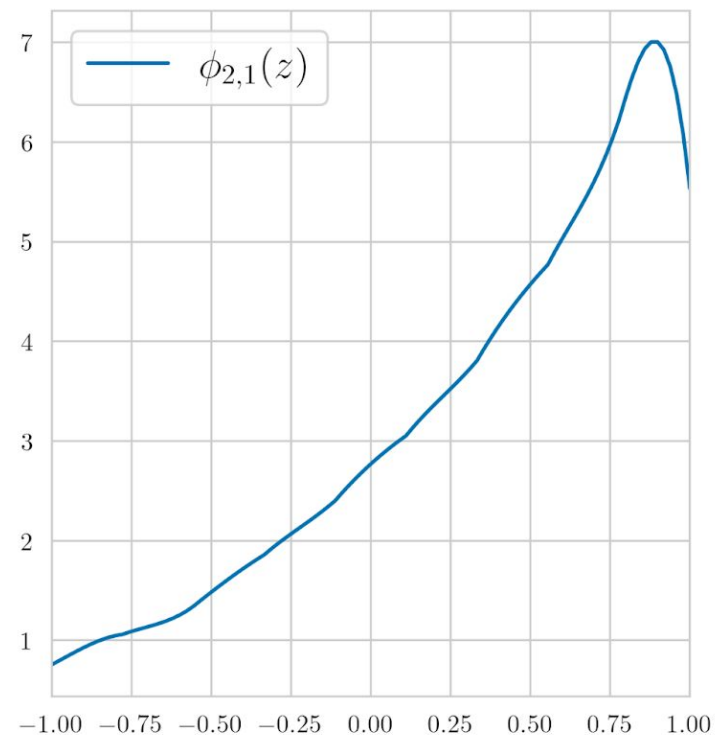
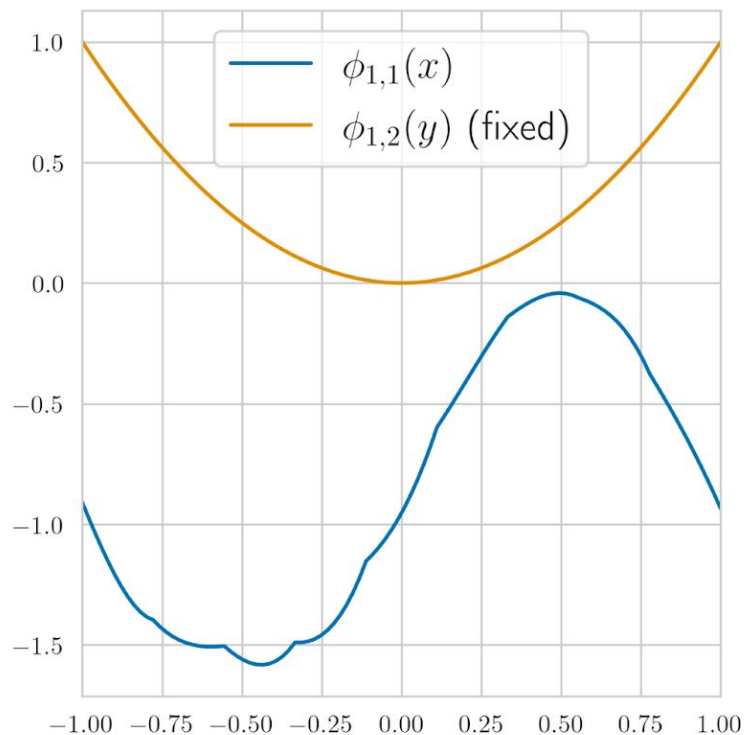
$$\hat{f}(x, y) = \phi_{2,1}(\phi_{1,1}(x) + \phi_{1,2}(y)).$$



# Сети Колмогорова–Арнольда

$$\hat{f}(x, y) = \phi_{2,1}(\phi_{1,1}(x) + \phi_{1,2}(y)).$$

- Теперь можно заметить, что  $\phi_{1,2}$  похоже на  $y^2$ , и зафиксировать:



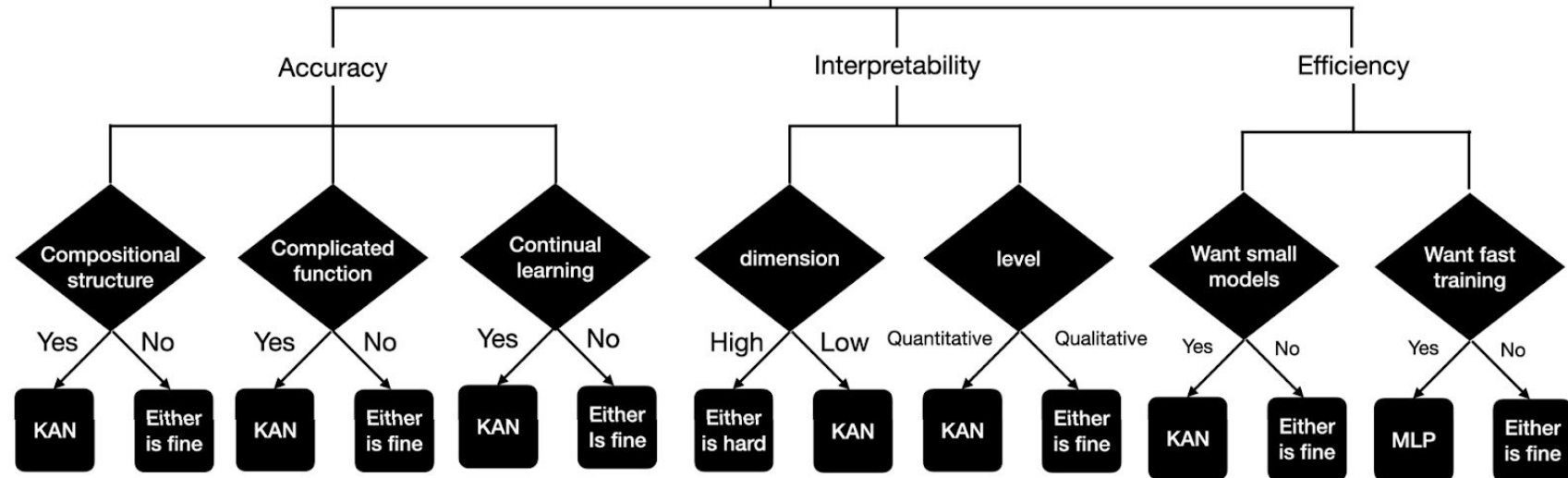
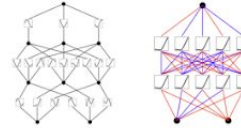


# Сети Колмогорова–Арнольда

- Как видите, изначально KAN – это скорее о том, чтобы вместе с человеком найти формулу аппроксимации; Liu et al. рекомендуют так:



Should I use KANs or MLPs?



# FastKAN и ReLU-KAN

- KAN в [Liu et al. \(2024\)](#) работали медленно, в основном из-за сложности построения B-сплайнов (de Boor–Cox iteration)
- FastKAN ([Li, 2024](#)) – можно заменить B-сплайны на гауссовские RBF и получить KAN втрое быстрее без потери качества:

$$\hat{f}(x) = \sum_{i=1}^m w_i \phi(\|x - \mu_i\|), \quad \phi(r) = \exp(-c \cdot r^2).$$

- ReLU-KAN ([Qiu et al., 2024](#)) – а можно и заменить на конструкцию из ReLU, что будет ещё эффективнее, в 5–20 раз быстрее:

$$R_i(x) = (\text{ReLU}(e_i - x) \cdot \text{ReLU}(x - s_i))^2 \cdot \frac{16}{(e_i - s_i)^4}.$$

Fast indeed!

# FastKAN и ReLU-KAN

- KAN в [Liu et al. \(2024\)](#) — замена сплайнов на кубические полиномы (вместо B-spline (Pade approximation) и Runge-Cox iteration) **30 апреля 2024 г.**
- FastKAN ([Li, 2024](#)) — замена сплайнов на гауссовские RBF и получение KAN быстрее без потери качества: **10 мая**

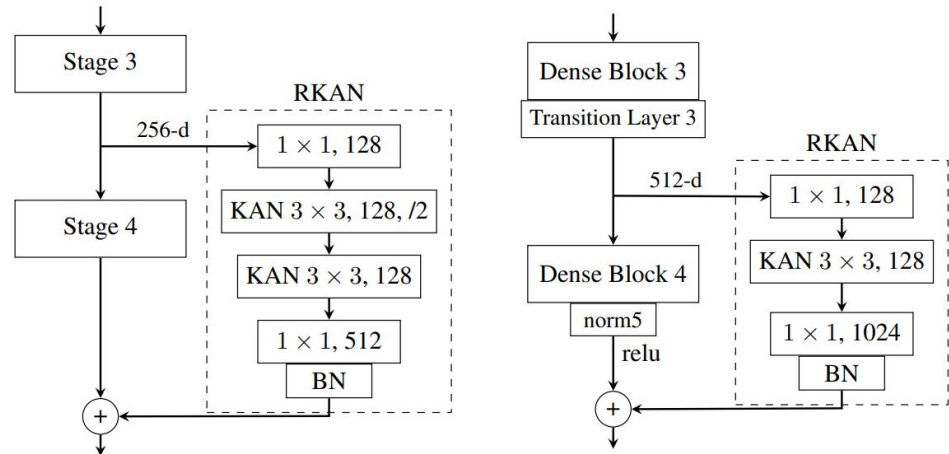
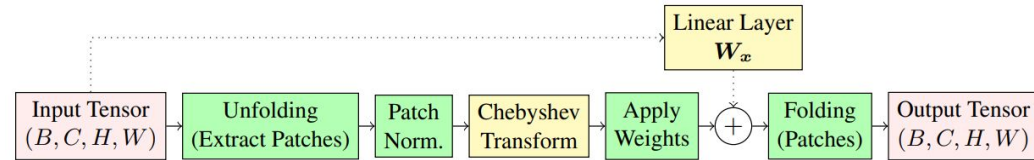
$$f(x) = \sum_{i=1}^m w_i \phi(\|x - \mu_i\|), \quad \phi(r) = \exp(-c \cdot r^2).$$

- ReLU-KAN ([Qiu et al., 2024](#)) — замена сплайнов на конструкцию из ReLU, что будет ещё быстрее, в 5–20 раз быстрее: **4 июня**

$$R_i(x) = (\text{ReLU}(e_i - x) \cdot \text{ReLU}(x - s_i))^2 \cdot \frac{16}{(e_i - s_i)^4}.$$

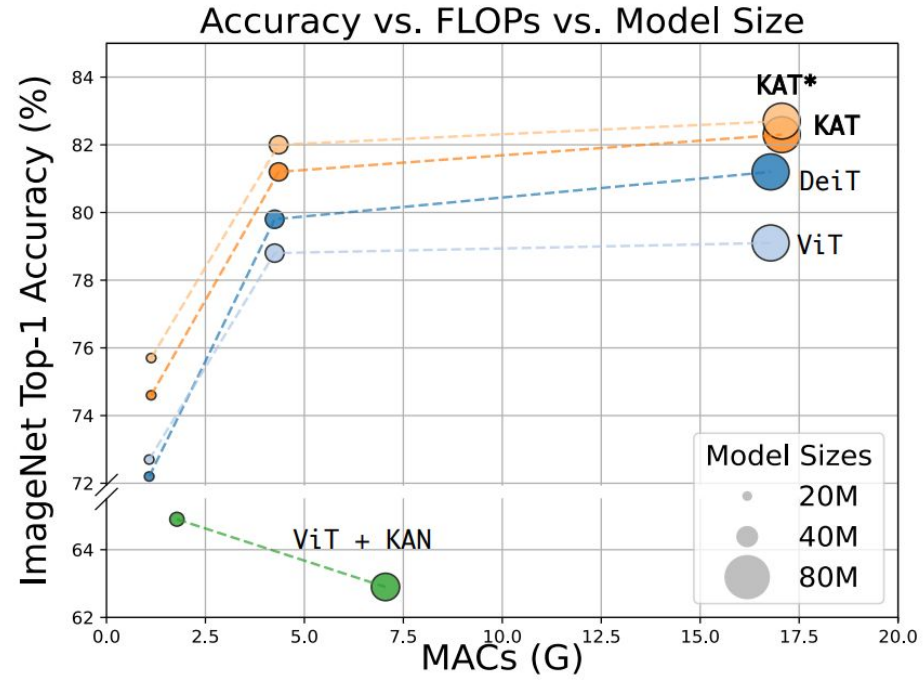
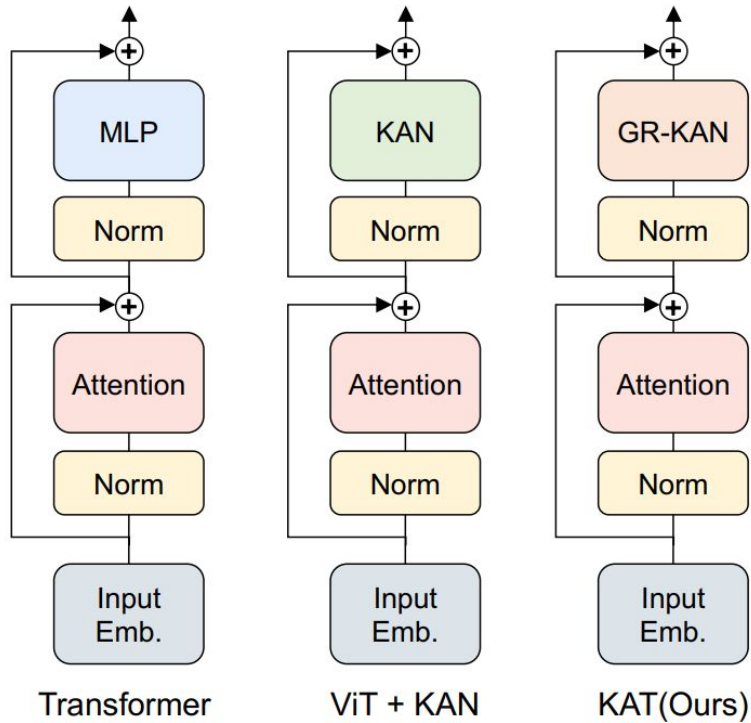
# KAN сегодня: архитектуры

- [Bodner et al. \(2024\)](#): Convolutional KAN, который можно приделать сбоку обычных архитектур
- [Genet et al. \(2024\)](#): Recurrent KAN (RKAN), Temporal KAN (TKAN) для обработки временных рядов



# KAN сегодня: архитектуры

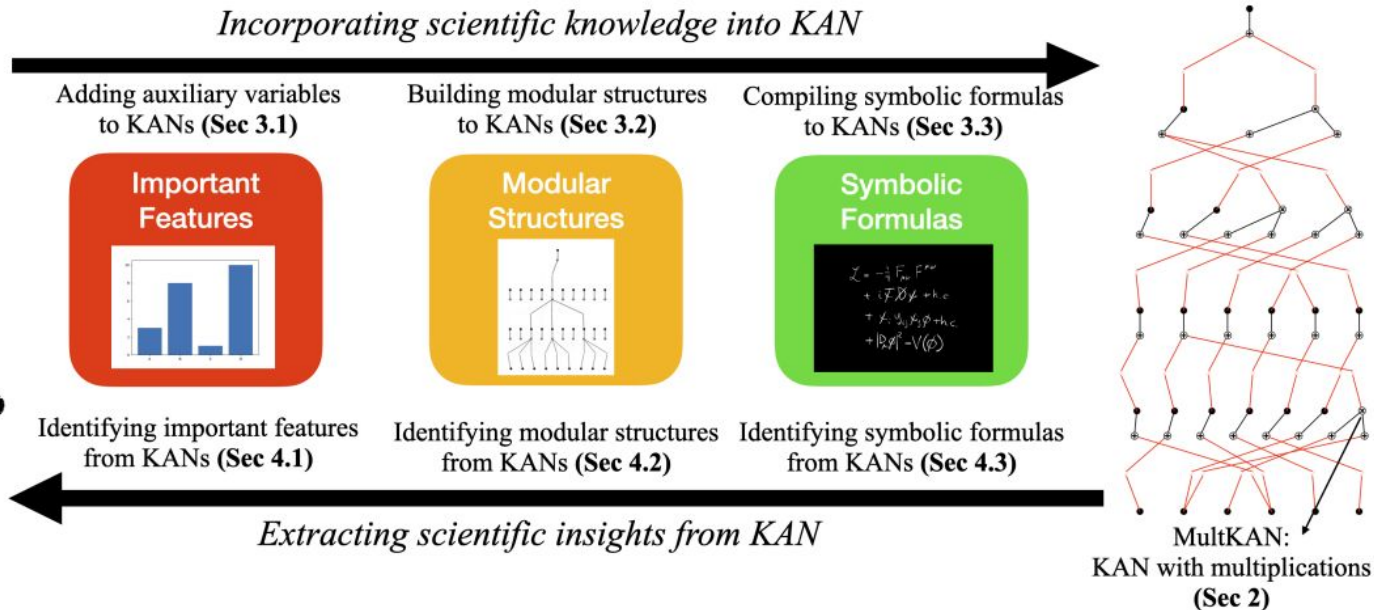
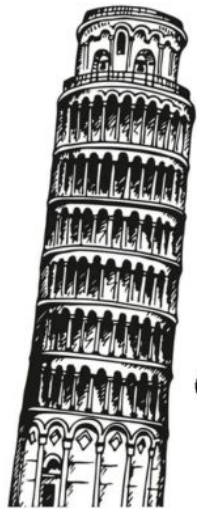
- [Yang, Wang \(2024\)](#): Kolmogorov–Arnold Transformer (KAT); self-attention оставим, а MLP в трансформере заменим на KAN



# KAN сегодня: применения

- [Liu et al. \(Aug 2024\)](#): KAN 2.0 помогает находить физические законы

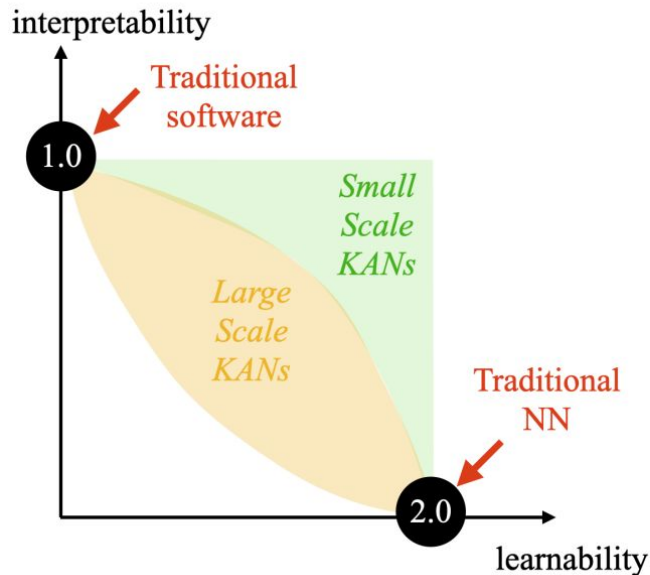
## Science



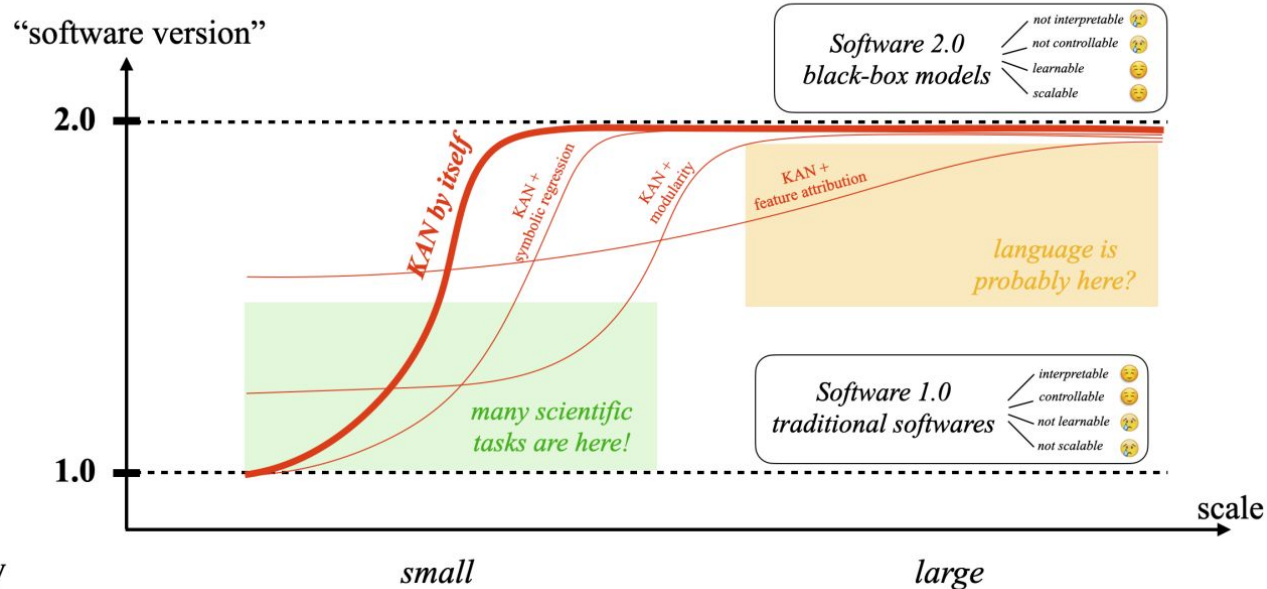
# KAN сегодня: применения

- [Liu et al. \(Aug 2024\)](#): в целом применения в основном вокруг интерпретируемости, KAN может выводить формулы

(a) interpretability-learnability plane



(b) “software version”-scale plane



# KAN: ВЫВОДЫ

- Интересная тема, всплывшая на поверхность в конце апреля 2024 года...
- ...и уже превратившаяся в большую область с многими разными применениями (MPT-снимки, молекулярная динамика, задачи оптимального управления и т.д.)
- В основном с упором на интерпретируемость, но это тоже важно; отличный пример ещё одного “низко висящего плода” в DL
- А что если обучать KAN через LLM?..







# МАШИННОЕ ОБУЧЕНИЕ

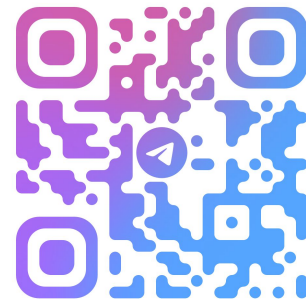
ОСНОВЫ



От Сергей Николаев

Дорогие друзья,  
в этой книге вы  
научитесь понимать  
язык машинного  
обучения и сможете  
применять его на  
практике. Книга  
подготовлена  
специально для  
русскоязычных  
читателей и  
содержит  
информацию  
о последних  
достижениях  
в области  
искусственного  
интеллекта.

# Спасибо за внимание!



@SINECOR

