

Large Language Models
Generative Pre-trained Transformers
Artificial Intelligence

LLMs GPTs in research and education AI

Jan Střeleček
October 2024

Expectations vs. reality



how many rocks should I consume each day

All Images Videos News Forums More

AI Overview

Learn more

According to UC Berkeley geologists, eating **at least one small rock per day** is recommended. Rocks can range in size from a handful of dust to a 5-pound cobblestone. Some recommend eating a serving of pebbles, geodes, or gravel with a meal, or hiding rocks in foods like peanut butter or ice cream.

ResFrac Corporation

Geologists Recommend Eating At Least One Small Rock Per Day - ...

May 19, 2021

Eating rocks can be good for you because they are a source of vitamins and minerals that are important for digestive health. Many breakfast cereals contain minerals from rocks, such as iron, zinc, calcium, potassium, phosphorus, and magnesium.

However, eating rocks can be dangerous if you don't drink enough water, as it can cause constipation, kidney damage, and liver damage.

Generative AI is experimental.

Good vs bad usage

is-odd-ai

PyPI v0.0.2.post1

A project inspired by the npm version [is-odd-ai](#)

`is-odd-ai` uses OpenAI's GPT-4o-mini model to determine if a number is odd or even. AI to the moon 🚀🚀🤪!

Installation

To install `is-odd-ai`, use pip:

```
pip install is-odd-ai
```

Usage

To use `is-odd-ai`, you will need an OpenAI key.

Here is an example.

```
from is_odd_ai import OddChecker

odd_checker = OddChecker("api key here")
odd_checker.is_odd(5) # True
odd_checker.is_odd(6) # False
odd_checker.is_odd("abc") # Hopefully a ValueError. AI to the moon!
```

ChatGPT 4o

Count the letters r in the word raspberry.



The word "raspberry" contains two letters "r."

(c) Determining the normalized states $|j m\rangle$ for $|3 3\rangle$, $|3 2\rangle$, and $|3 1\rangle$:

To construct these states, we'll use the Clebsch-Gordan coefficients to combine the individual angular momenta.

For $|3 3\rangle$:

This state corresponds to the maximum projection of the total angular momentum. All individual angular momenta are aligned along the quantization axis.

$$|3 3\rangle = |l = 2, m_l = 2\rangle \otimes |s_e = \frac{1}{2}, m_{s_e} = \frac{1}{2}\rangle \otimes |s_p = \frac{1}{2}, m_{s_p} = \frac{1}{2}\rangle$$

For $|3 2\rangle$:

Possible combinations of m_l , m_{s_e} , and m_{s_p} satisfying $m_l + m_{s_e} + m_{s_p} = 2$ are:

1. $m_l = 2, m_{s_e} = \frac{1}{2}, m_{s_p} = -\frac{1}{2}$
2. $m_l = 1, m_{s_e} = \frac{1}{2}, m_{s_p} = \frac{1}{2}$

Using the Clebsch-Gordan coefficients, the normalized state is:

$$|3 2\rangle = \sqrt{\frac{5}{6}} |2, 1\rangle |\frac{1}{2}, \frac{1}{2}\rangle |\frac{1}{2}, \frac{1}{2}\rangle + \sqrt{\frac{1}{6}} |2, 2\rangle |\frac{1}{2}, -\frac{1}{2}\rangle |\frac{1}{2}, \frac{1}{2}\rangle$$

For $|3 1\rangle$:

Possible combinations satisfying $m_l + m_{s_e} + m_{s_p} = 1$:

1. $m_l = 1, m_{s_e} = \frac{1}{2}, m_{s_p} = -\frac{1}{2}$
2. $m_l = 0, m_{s_e} = \frac{1}{2}, m_{s_p} = \frac{1}{2}$
3. $m_l = 1, m_{s_e} = -\frac{1}{2}, m_{s_p} = \frac{1}{2}$

The normalized state is:

$$|3 1\rangle = \sqrt{\frac{2}{3}} |2, 0\rangle |\frac{1}{2}, \frac{1}{2}\rangle |\frac{1}{2}, \frac{1}{2}\rangle + \sqrt{\frac{1}{3}} |2, 1\rangle |\frac{1}{2}, -\frac{1}{2}\rangle |\frac{1}{2}, \frac{1}{2}\rangle$$

What are the current GPTs

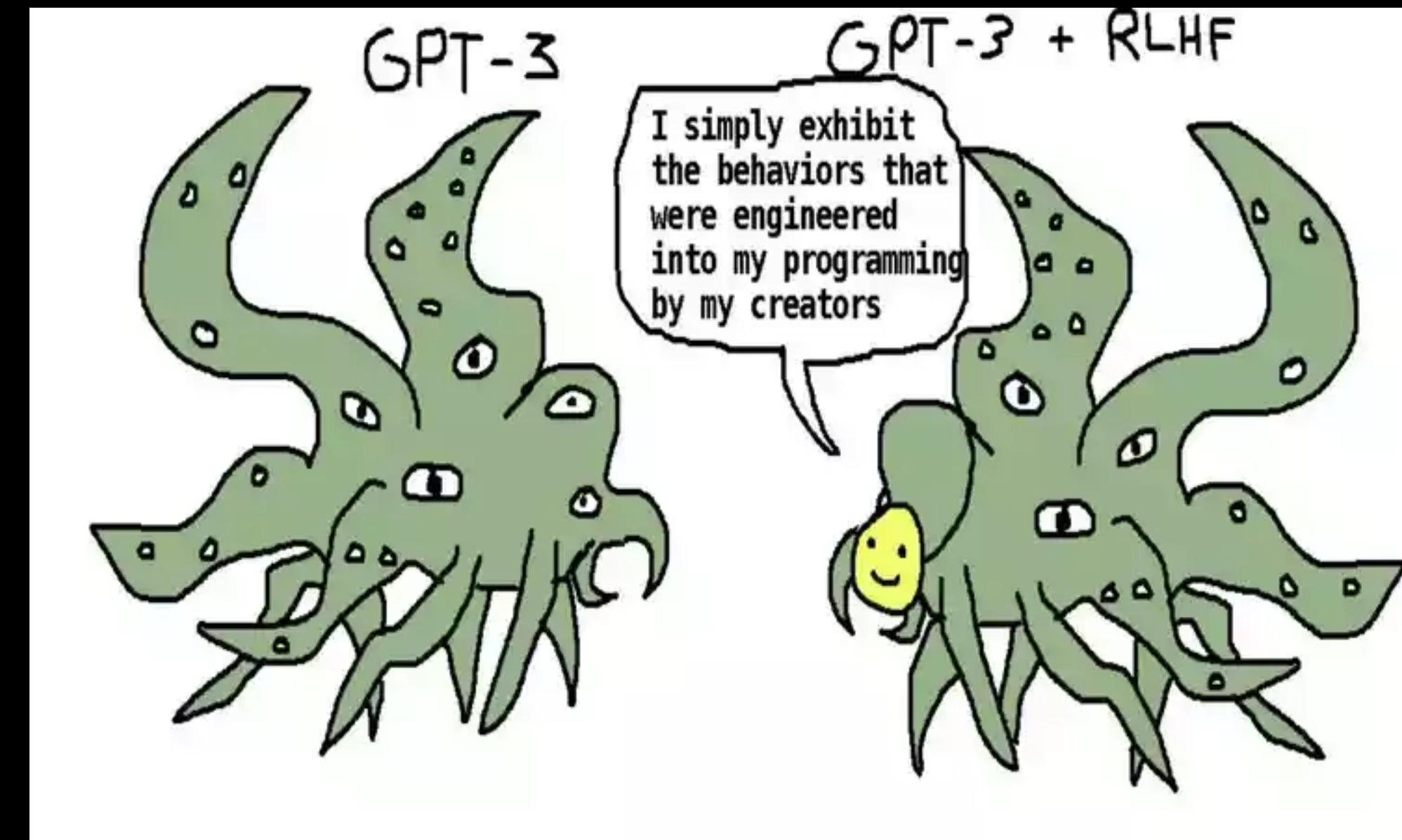
1. gather and clean data
2. train the model and make a chatbot from it
4. Reinforcement Learning from Human Feedback
 - makes it safe, but more stupid and obedient

ask: "Did your answer meet my expectations?"

+ self corrections using chain of thought etc.

+ questions are translated into the "models code", so it is not smarter in English (e.g. cross-lingual-thought prompting)

5. Build a tool on top



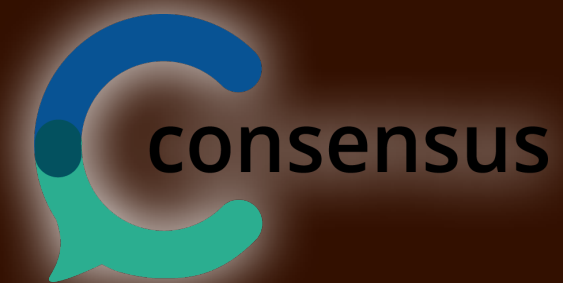
<https://aclanthology.org/2023.findings-emnlp.826/>

Education



Research

 Litmaps



scite_

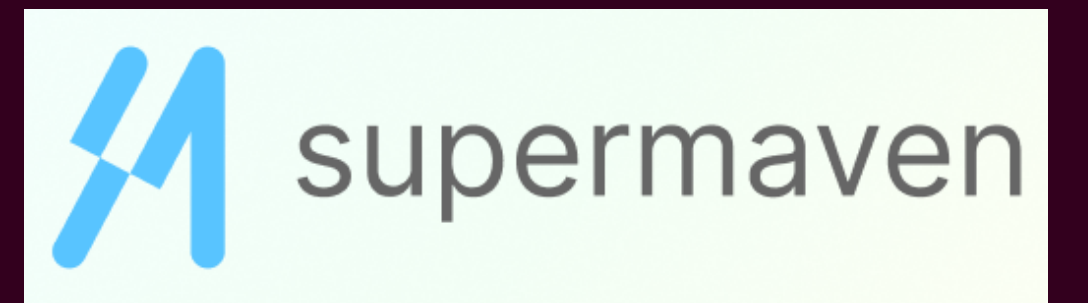


Gemini

LLaMA



Programming



Education

“If kids could get microlessons in the form of on-demand video, at their own time and pace, class time could be used for Socratic dialogue, collaborative assignments, and supported student work. Essentially, lectures were now happening at home, and ‘homework’ could now happen in a much more interactive classroom environment.”

“...but what if it didn’t write for them at all? What if, instead, it wrote with them?”

Brave New Words

How AI Will
Re|volutionize
Education (and
Why That’s a
Good Thing) ☀

Salman Khan

allen lane

“...Bloom wrote that if a student works with a tutor to master a topic or skill, the student would gain a two-standard-deviation improvement—a massive upgrade that takes someone from the 50th percentile to roughly the 96th percentile.”

wikipedia.org/wiki/Bloom's_2_sigma_problem

“What might it be like if every student on the planet had access to an **artificially intelligent personal tutor**: an AI capable of writing alongside the student.”

“To address concerns around cheating, for instance, we considered what it would take to create an AI tutor that refused to give students answers. Like a good human tutor, it would instead ask leading questions.”



Khanmigo

AI power



This conversation is recorded and viewable by your parent and teacher.

My code isn't working, what's wrong?

This is it:

```
var animals = ["elephant", "orca", "sloth"];
```

```
var counter = 1;
```

```
while (counter < animals.length) {
```

```
  console.log(animals[counter]);
```

```
}
```

You're almost there! You just need to make sure to increment the counter variable inside the loop, otherwise it will get stuck in an infinite loop.

Type message (do NOT share any personal data)



Clear chat



Education: general models vs. Khanmigo



What do the letters in qma and bqp stand for

Show drafts



The letters in the complexity classes QMA and BQP stand for:

- **QMA:** Quantum Merlin Arthur
- **BQP:** Bounded-error Quantum Polynomial time

Here's a brief explanation of each:

- **QMA:** This class represents problems that can be verified by a quantum computer with help from a quantum proof (provided by Merlin) in polynomial time. If the answer is "yes," there's a high probability that Merlin can convince Arthur (the verifier) of this. If the answer is "no," there's a high probability that Arthur won't be convinced, even if Merlin tries to cheat.
- **BQP:** This class represents problems that can be solved by a quantum computer in polynomial time with a bounded probability of error. In other words, a quantum computer can solve these problems efficiently, and the answer will be correct with a high probability.

“Well, I got something a little bit different from you. Can you explain your reasoning to me?”

“...AI is not bias-free ... Instead, we should measure its risk relative to the bias that is already involved in subjective processes such as hiring and assessment.”

“we are limiting ourselves to a much narrower type of assessment that arguably magnifies a bias toward prioritising easy-to-measure skills over ones that are harder to measure but perhaps more important.”

Education: Khanmigo, teachers perspective

“Khanmigo can report back, ‘We worked on the paper for about four hours. Sal initially had trouble coming up with a thesis, but I was able to help him by asking some leading questions. The outlining went pretty smoothly. I just had to help him ensure that the conclusion really brought everything together. Sal did most of the writing. I just helped him tidy up the grammar and strengthen his argument in the third paragraph. Based on the rubric for the assignment, I’d recommend Sal get a B + on the assignment.’”

Education: problems

“The key is for parents and application developers to find a balance between monitoring and respecting children’s privacy and independence.”

“PARENT: What has my daughter been up to?”

KHANMIGO: She’s spent fifty-four minutes this past week trying to get more mastery of Algebra I. I helped her a bit when she was confused about slope. In particular, she kept trying...”

 **Almost a year ago**

“Right now, our best estimate of the computation costs of average usage of Khanmigo is between five and fifteen dollars a month per user.”

Education: how to use what we have now?

prompt

**learning material
+
prompt**

**learning material
+
socratic conversation**

example 

Research

Opinion #1:

Use it in the beginning, so you don't lose the sense of authorship.

First drafts take a lot of time, editing is easier (for some of us)


- a) write a lot of text, let GPT rewrite it, then it's your turn
- b) write list of stuff you would like to include, GPT rewrites it, or gives me a good approach to arrange it


Opinion #2:

Use it at the end, just to improve the writing to make sure all ideas are yours.

Research

Where have I seen this...

↳ In which article they mentioned that QPTs are impassable for geodesics? 

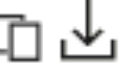
 The concept of geodesics in the context of quantum phase transitions (QPTs) is critical for understanding the geometric structure of parameter spaces associated with quantum systems. Specifically, the notion that geodesics can be considered impassable in certain scenarios is discussed in the literature, particularly in relation to the behavior of quantum systems at critical points.

In the work by Kumar et al., it is emphasized that the study of geodesics is essential in understanding the geometric properties of quantum systems undergoing phase transitions. The authors note that in various geometric setups, including those relevant to QPTs, the behavior of geodesics can exhibit significant anomalies, particularly at critical points where the system undergoes a transition (Kumar et al., 2012). This suggests that geodesics may not always provide a continuous path through parameter space, indicating that they can be "impassable" in certain contexts.

Furthermore, the article by Rezakhani et al. discusses the intrinsic geometry of quantum adiabatic evolution and highlights that geodesics can become obstructed during QPTs. The authors argue that the geometric structure of the parameter space changes dramatically at critical points, leading to situations where the expected geodesic paths may not be traversable due to singularities or other geometric obstructions (Rezakhani et al., 2010). This aligns with the broader understanding that critical phenomena often lead to non-analytic behavior in the parameter space, which can render certain geodesic paths inaccessible.

Additionally, the work by Venuti and Zanardi provides insights into the scaling behavior of geometric tensors at zero-temperature phase transitions, reinforcing the idea that geodesics can become problematic near critical points. They illustrate that the geometric properties of the system can lead to

 **References**  Search Strategy



"...In any geometric setup, the behavior of geodesics is an important object to study. Some studies on geodesics have appeared in..." [>>](#)

Section: Introduction

Geodesics in information geometry: Classical and quantum phase transitions

[Prashant Kumar](#)¹, [Subhash Mahapatra](#)², [Prabwal Phukon](#)³ et al. 2012 *Phys. Rev. E*

 36 |  1 |  29 |  0

[View full text](#) [Add to dashboard](#) [Cite](#)

"...Geometric and topological concepts have long played useful roles in both classical and quantum physics [>>](#)

Section: Introduction

Intrinsic geometry of quantum adiabatic evolution and quantum phase transitions

[A. T. Rezakhani](#)¹, [Damian F. Abasto](#)², [Daniel A. Lidar](#)³ et al. 2010 *Phys. Rev. A*

 86 |  2 |  104 |  0

[View full text](#) [Add to dashboard](#) [Cite](#)

"...-Phase transitions at zero-temperatures are dramatic changes in the ground-state (GS) properties of a system driven by quantu..." [>>](#)

Quantum Critical Scaling of the Geometric Tensors

[Lorenzo Campos Venuti](#)¹, [Paolo Zanardi](#)² 2007 *Phys. Rev. Lett.*

 401 |  5 |  210 |  0

[View full text](#) [Add to dashboard](#) [Cite](#)

Research

Remind me...




What is sampling in this context?

1806.06081v2

PDF PDF



Znovu vygenerovat koncept 



Analysis complete 

Sampling, in this context, refers to the ability of a quantum annealing algorithm or machine to find all possible solutions (ground states) to a given optimization problem with equal probability. It is the process of repeatedly running the optimization procedure and observing the distribution of the solutions obtained. Fair sampling means that each valid solution has an equal chance of being found, while biased sampling indicates that some solutions are favored over others.

Research

How should I study this topic?



In which order should I read the articles if I want to understand the concept of shortcuts to adiabaticity? Give me a short description of what should I expect from these articles, point out which parts are the most important to read or if I should read them at all. Give me a short summary of all of them and how it fits the research...

Takahashi ...ylov Space

PDF



Hörnedal ... operators

PDF



Yamashiro ...spin model

PDF



Vodola et ...edge n

PDF

Regenerate draft



Analysis complete

The four articles you uploaded are all relevant to the concept of shortcuts to adiabaticity (STA), but they approach the topic from different angles and with varying levels of complexity. Here's a suggested reading order along with a brief overview of each article:

- 1. Takahashi and Del Campo (2024):** This article provides a comprehensive overview of STAs, including their history, theoretical foundations, and various techniques for engineering them. It also discusses the challenges and potential solutions for implementing STAs in complex quantum systems. This is a good starting point as it gives a broad overview of the field and introduces key concepts like counterdiabatic (CD) driving and the adiabatic gauge potential (AGP).
 - **Key sections:** Introduction, Section III (Adiabatic Gauge Potential and Counterdiabatic Driving)
- 2. Yamashiro et al. (2019):** This article focuses on a specific application of STAs, namely reverse annealing, in the context of quantum annealing for solving optimization problems. It investigates the effectiveness of reverse annealing for a particular model (the p-spin model) and compares the performance of quantum and classical algorithms. While not essential for understanding the basics of STAs, this article provides valuable insights into their practical applications.
 - **Key sections:** Introduction, Section II (Adiabatic Reverse Annealing)

Research: direct usage of LLMs

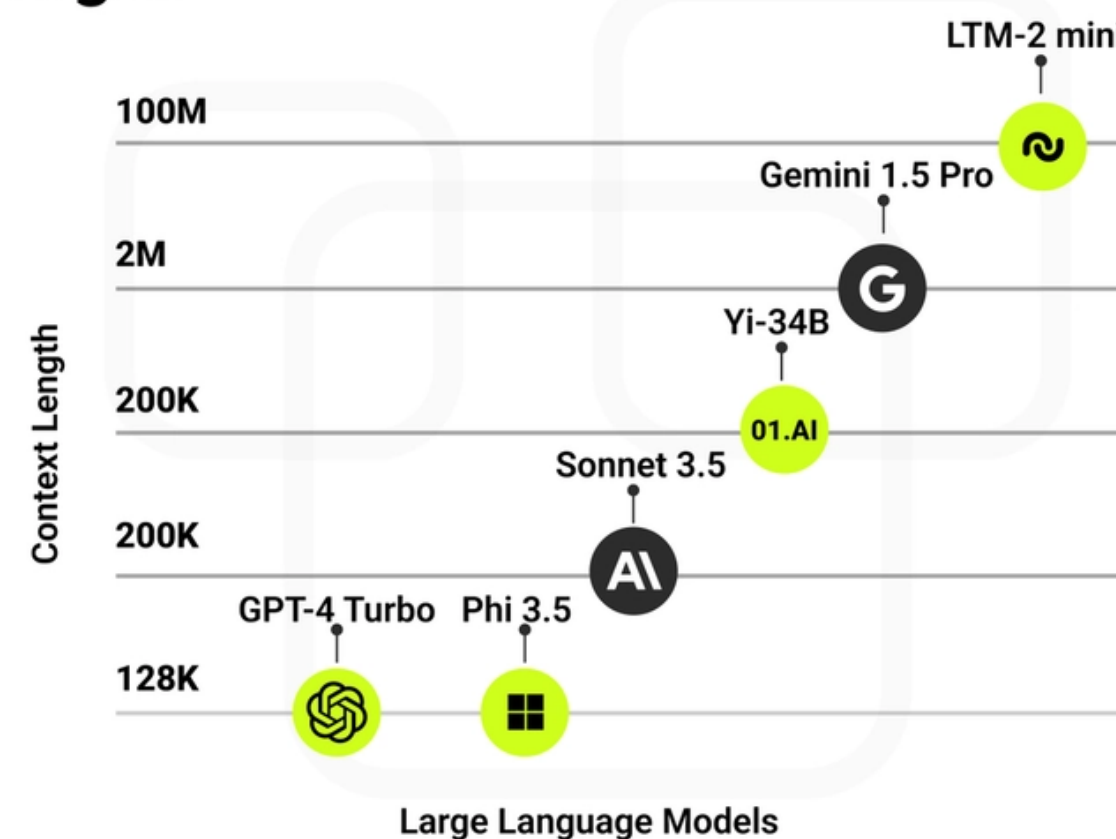
 gemini.google.com

 claude.ai

 chatgpt.com

LLaMA not in EU,
but open-source

Models with larger context lengths



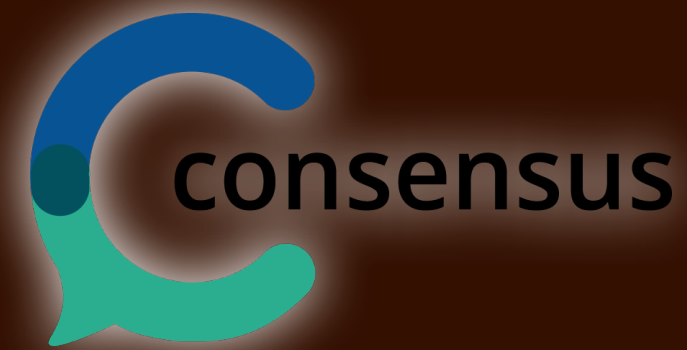
(as of 8/30/2024)

linkedin.com/in/rakeshgohel01

	Claude 3.5 Sonnet (new)	Claude 3.5 Haiku	Claude 3.5 Sonnet	GPT-4o*	GPT-4o mini*	Gemini 1.5 Pro
Graduate level reasoning <i>GPQA (Diamond)</i>	65.0% 0-shot CoT	41.6% 0-shot CoT	59.4% 0-shot CoT	53.6% 0-shot CoT	40.2% 0-shot CoT	59.1% 0-shot CoT
Undergraduate level knowledge <i>MMLU Pro</i>	78.0% 0-shot CoT	65.0% 0-shot CoT	75.1% 0-shot CoT	—	—	75.8% 0-shot CoT
Code <i>HumanEval</i>	93.7% 0-shot	88.1% 0-shot	92.0% 0-shot	90.2% 0-shot	87.2% 0-shot	—
Math problem-solving <i>MATH</i>	78.3% 0-shot CoT	69.2% 0-shot CoT	71.1% 0-shot CoT	76.6% 0-shot CoT	70.2% 0-shot CoT	86.5% 4-shot CoT
High school math competition <i>AIME 2024</i>	16.0% 0-shot CoT	5.3% 0-shot CoT	9.6% 0-shot CoT	9.3% 0-shot CoT	—	—
Visual Q/A <i>MMMU</i>	70.4% 0-shot CoT	—	68.3% 0-shot CoT	69.1% 0-shot CoT	59.4% 0-shot CoT	65.9% 0-shot CoT
Agentic coding <i>SWE-bench Verified</i>	49.0%	40.6%	33.4%	—	—	—
Agentic tool use <i>TAU-bench</i>	Retail 69.2% Airline 46.0%	Retail 51.0% Airline 22.8%	Retail 62.6% Airline 36.0%	—	—	—

example →

Research: self-check and provide references



consensus.app

scite_

scite.ai



elicit.com



typeset.io

example →

Research: other tools

 Litmaps

litmaps.com


ResearchRabbit

researchrabbitapp.com



connectedpapers.com

Research: other tools

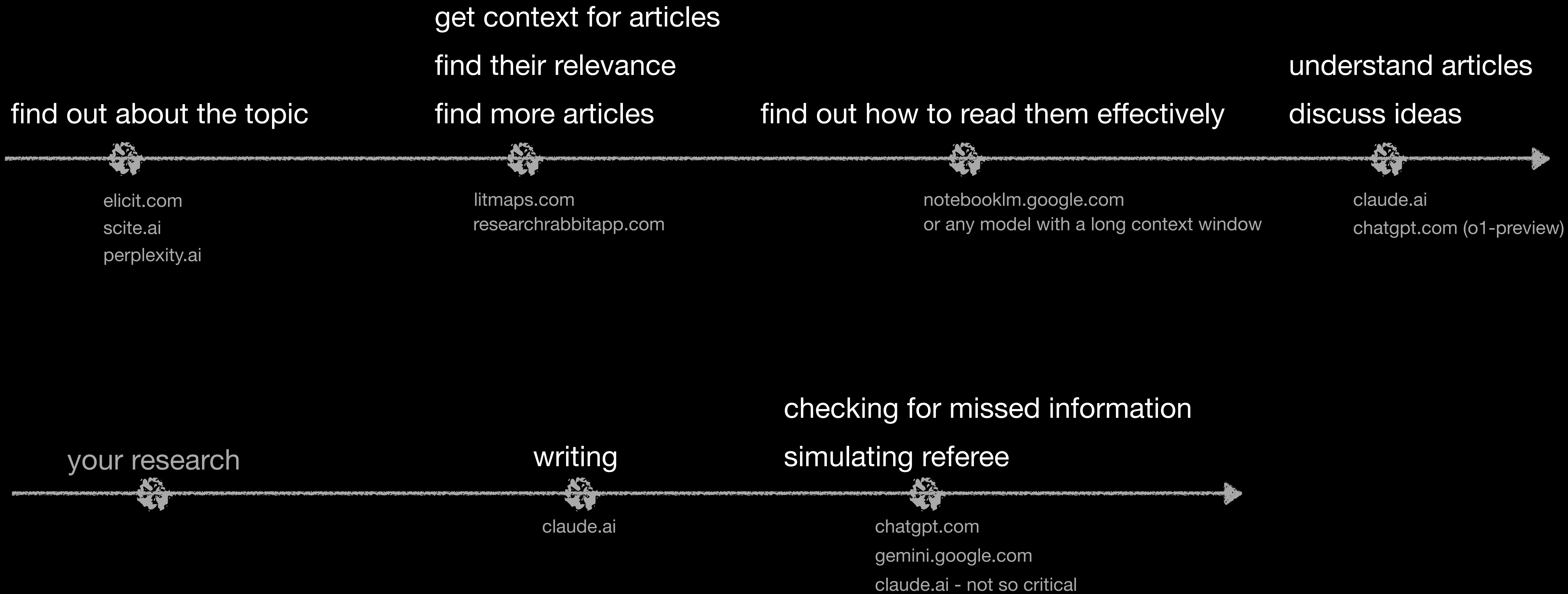


perplexity.ai

NotebookLM

notebooklm.google.com

Research: example timeline

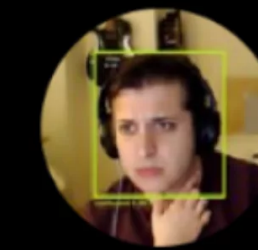


Programming



github.com/edu/students

supermaven.com



Jordan Scales

@jdan

GitHub Copilot still has a couple kinks but I like what I see so far!

```
1
2
3 function invertBinaryTree(tree) {
4   throw Error("no that's dumb as hell you don't
   need to do that ever lmao")
5 }
```


11:44 AM · 29 Jun 21 · [Twitter Web App](#)

Programming: Mathematica

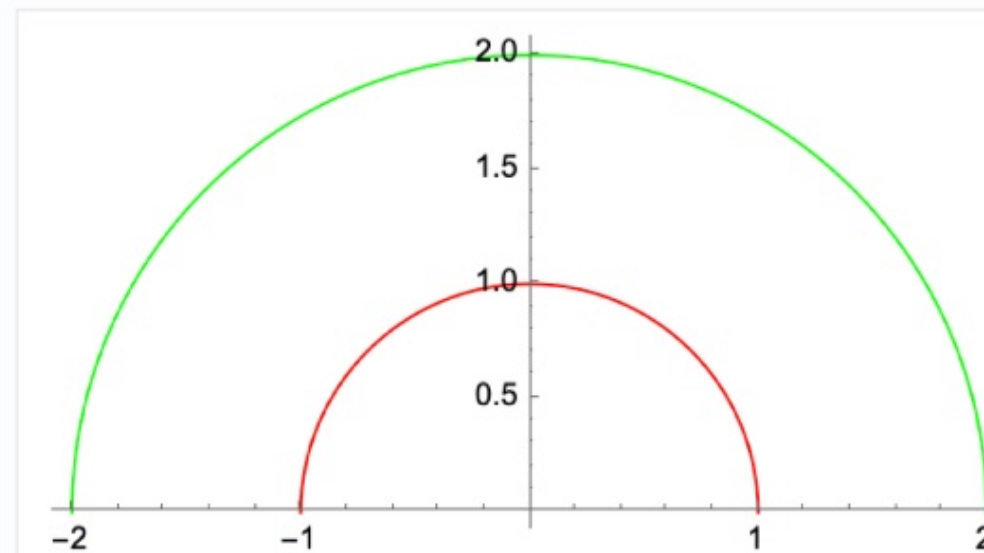
Draw a red and a green semicircle

You can draw a red and a green semicircle using the `Graphics` function along with `Circle` to define the semicircles. I'll create two semicircles with their centers at the origin, one with a radius of 1 (red) and the other with a radius of 2 (green). Here's the code:

```
Graphics[
{
  {Red, Circle[{0, 0}, 1, {0, Pi}]},
  {Green, Circle[{0, 0}, 2, {0, Pi}]}
},
Axes -> True
]
```

> Used Wolfram Language Evaluator 

Here's the plot showing a red semicircle with a radius of 1 and a green semicircle with a radius of 2, both centered at the origin:



...we will see where it goes

where to find more:

- [youtube.com/@DrAndyStapleton](https://www.youtube.com/@DrAndyStapleton)
- x.com/airesearchtools
- S. Khan, Brave New Words: How AI Will Revolutionize Education (and Why That's a Good Thing), 2024

