

Generative AI and the dream of an (inexpensive) personalized tutor for every student



Constantine Dovrolis

Director

Computational Science & Technology
Research Center (CaSToRC)

The Cyprus Institute (CyI)

Outline

- Large Language Models (LLMs)
 - What is the problem they solve?
 - How do they work?
 - Are they “stochastic parrots”? Are they “intelligent”?
- LLMs in education
 - The dream of an AI-based personalized tutor
 - Examples
 - Concerns
- Do LLMs actually improve learning performance?
 - A recent quantitative study



The amazing capabilities of Large Language Models (LLMs)

Prompt: Here we have a book, 9 eggs, a laptop, a bottle and a nail. Please tell me how to stack them onto each other in a stable manner.



Do LLMs exhibit “theory of mind” abilities?

Scenario:

Tom just lost his ZURFIN, but he didn't tell Adam.

Adam starts talking to him about how much he loves his ZURFIN, and notices that Tom makes a sad face.

--

Q: Why is Tom making a sad face?<|endofprompt|>



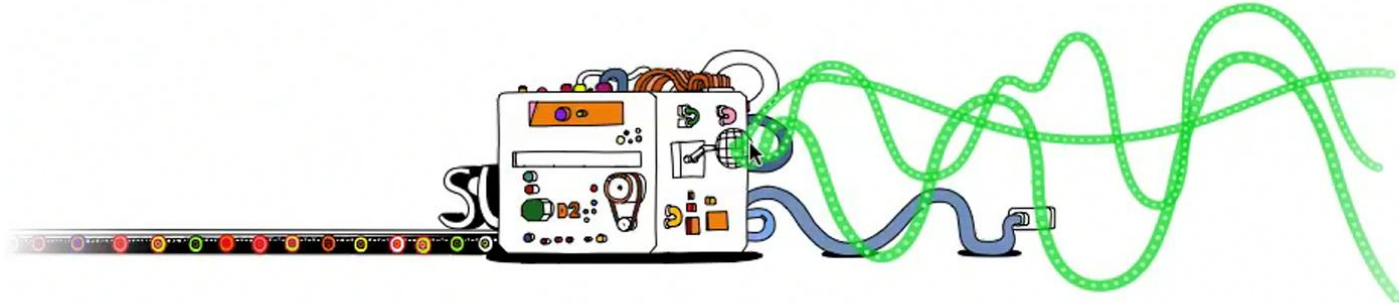
Sequence to sequence mapping





The Sequence Transduction problem

- Sequence to sequence mapping



Source: Alex Graves ([ICML'12 tutorial](#))

Let $\mathbf{x} = \{x_1, x_2, \dots, x_T\}$ be the input sequence and $\mathbf{y} = \{y_1, y_2, \dots, y_U\}$ be the output sequence, where $x_t \in \mathcal{S}_x$, $y_u \in \mathcal{S}_y$, and \mathcal{S}_x and \mathcal{S}_y are the sets of possible things that each x_t and y_u can be, respectively.

$$\begin{aligned} p_{\theta}(\mathbf{y}|\mathbf{x}) &= p_{\theta}(y_1, y_2, y_3, \dots, y_U|\mathbf{x}) \\ &= \prod_{u=1}^U p_{\theta}(y_u|y_{u-1}, y_{u-2}, \dots, y_1, \mathbf{x}) \end{aligned}$$

many other applications...



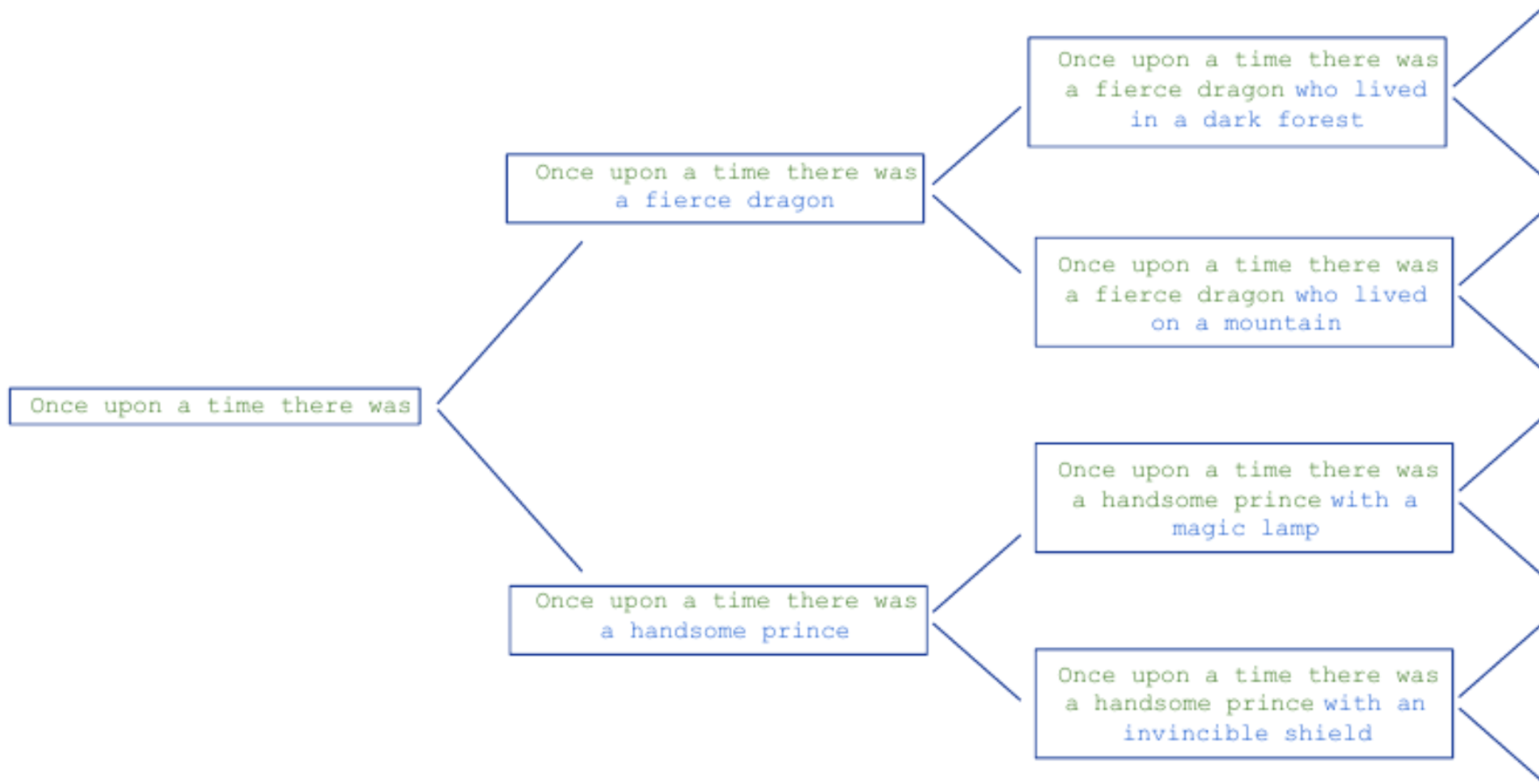
A (very brief) history of Natural Language Processing

(NLP)

- **Rule-based methods (1950s-80s):** struggled with ambiguity & complexity of human language - also not much data to work with
- **Statistical methods (till 1990s):** n-gram based methods, Hidden Markov Models (HMMs), use of Viterbi algorithm, progress in Named Entity Recognition tasks
- **Corpus linguistics (late 1980s-early 2000s):** availability of large language datasets from Internet resources - development of methods such as tf-idf for info retrieval from documents
- **Machine learning (1990s till 2010s):** Support Vector Machines (SVMs), Bayesian methods, Latent Semantic Analysis (LSA) - but required "feature engineering"
- 2003: **"A neural probabilistic language model"** by Bengio et al.
 - Introduced the power of neural networks and word embeddings in NLP, started current era of NLP

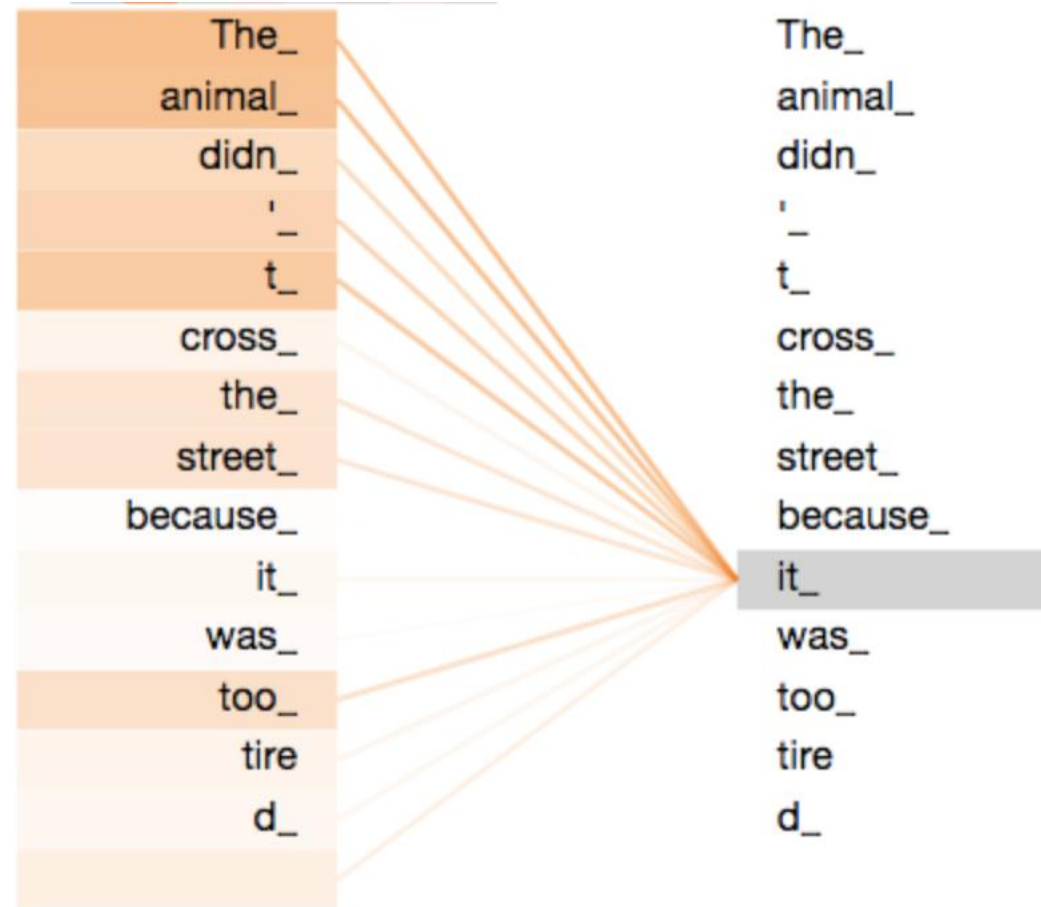


LLMs are (Stochastic) Autoregressive Models





Fundamental idea behind LLMs: self-attention



How are Large Language Models trained?

A Pretraining

Source: A. Radford et al., "Improving Language Understanding by Generative Pre-training"



Large corpus
(unlabeled text)

"Would you tell me, please, which way I ought to go from here?"
"That depends a good deal on where you want to get to," said the Cat.
"I don't much care where—" said Alice.
"Then it doesn't matter which way you go," said the Cat.
"—so long as I get *somewhere*," Alice added as an explanation.
"Oh, you're sure to do that," said the Cat, "if you only walk long enough."

Original text

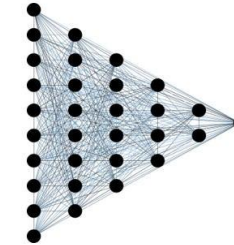
Masking



"Would you tell me, [REDACTED], which way I [REDACTED] to go from here?"
"That [REDACTED] a [REDACTED] deal on where you want to get to," said the Cat.
"I [REDACTED] much care where—" [REDACTED] Alice.
"Then it doesn't matter [REDACTED] [REDACTED] you go," said the Cat.
"—so long as I get *somewhere*," Alice [REDACTED] as an explanation.
"Oh, [REDACTED] [REDACTED] to do that," said the Cat, "if [REDACTED] only [REDACTED] long enough."

Masked text

Language model



"Would you tell me, *sir*, which way I *need* to go from here?"
"That *depends* a *good* deal on where you want to get to," said the Cat.
"I *don't* much care where—" *said* Alice.
"Then it doesn't matter *which way* you go," said the Cat.
"—so long as I get *somewhere*," Alice *added* as an explanation.
"Oh, *no need* to do that," said the Cat, "if *one* only *waits* long enough."

Predicted text

Loss



Are LLMs only text prediction machines? – Do they “think”?

All human knowledge available
on the Internet



An imaginary super-polymath that can
access all human knowledge



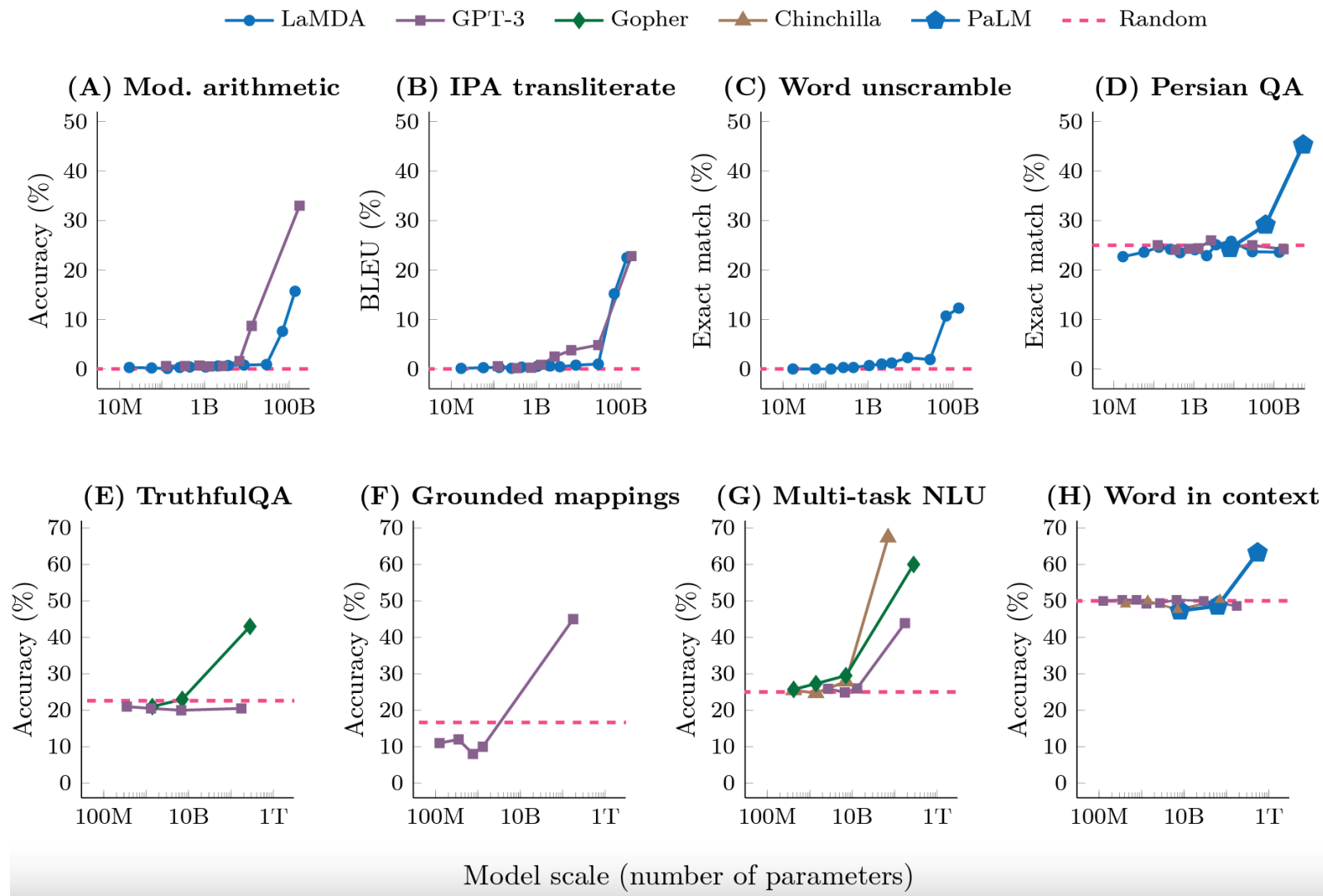
LLM predicts what
the super-polymath would say



- What can go wrong?
 - Input data can be wrong, biased, self-contradictory, etc
 - The super-polymath may not have a unique answer – the question (prompt) may be vague – the polymath may not know the answer
 - The LLM may fail to predict accurately what the polymath would say



An ability is emergent if it is not present in smaller models but is present in larger models.

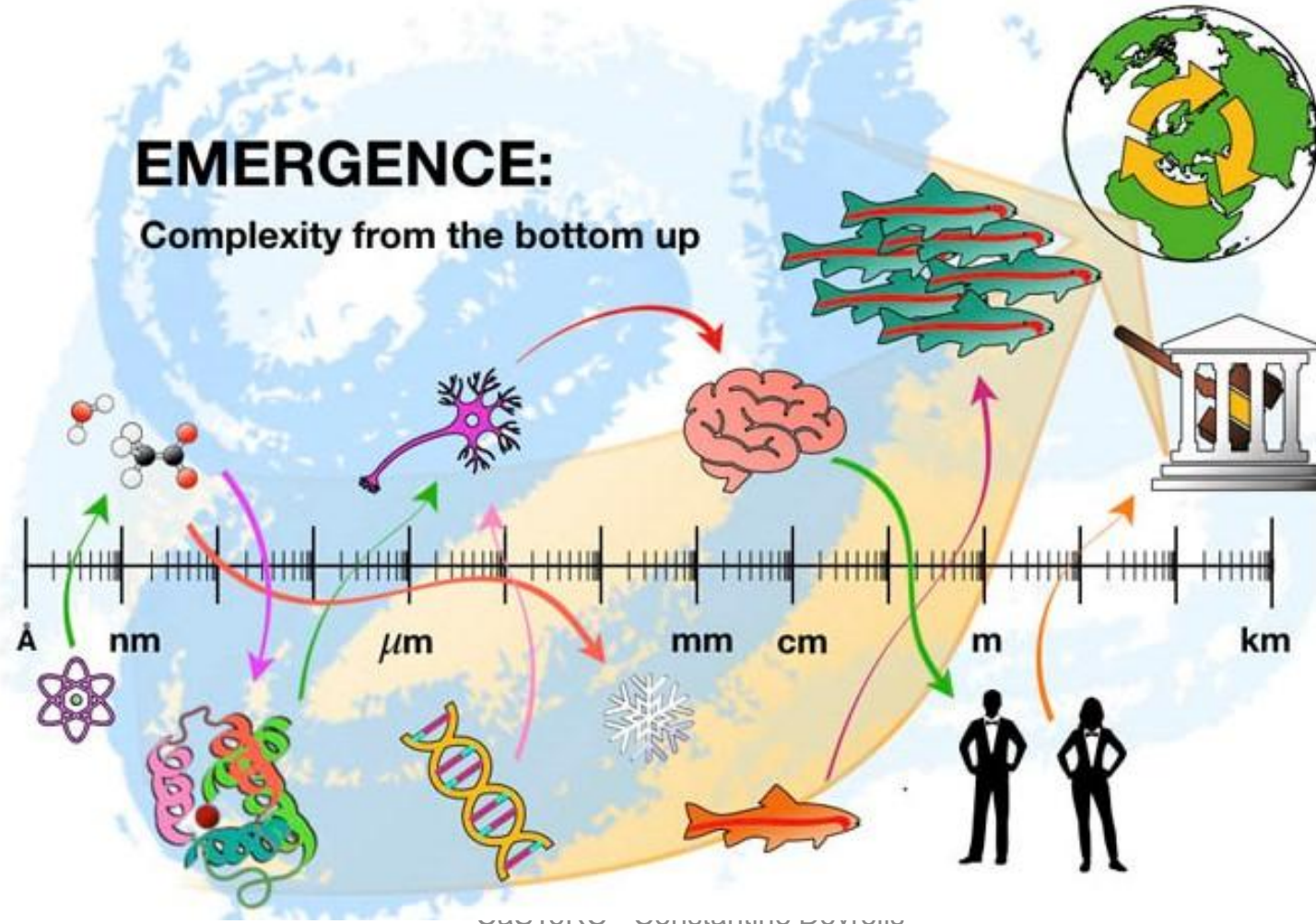


CaSToRC - Constantine Dvrolis



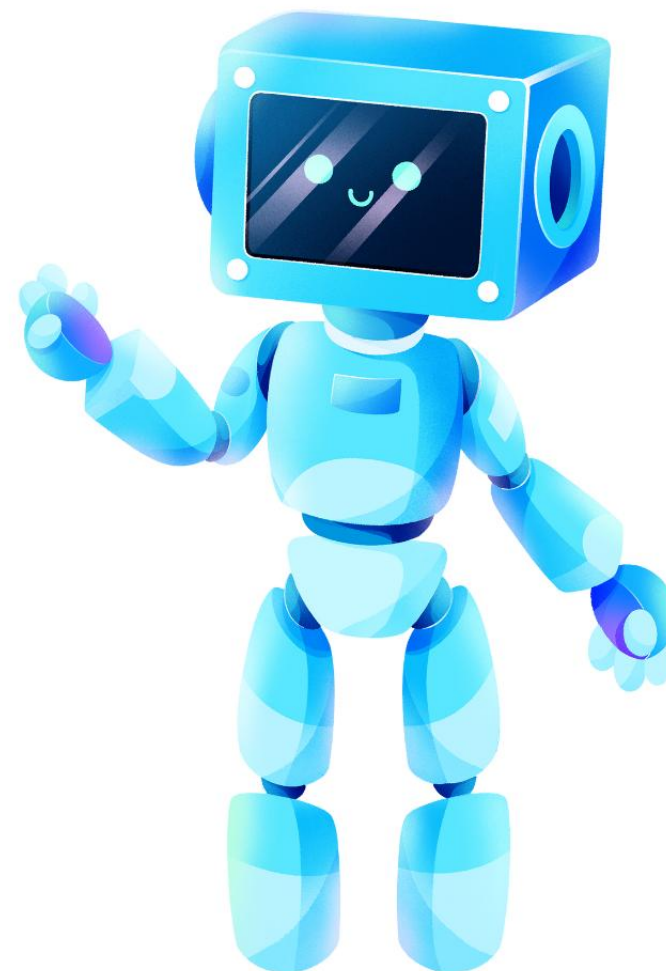
Emergence in Large Language Models?

- "Emergence is when quantitative changes in a system result in qualitative changes in behavior." (P.Anderson, 1972, "More is different")*





LLMs in Education



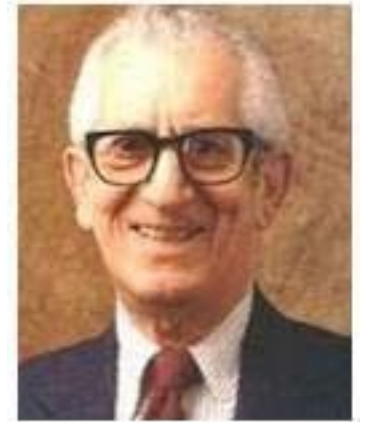
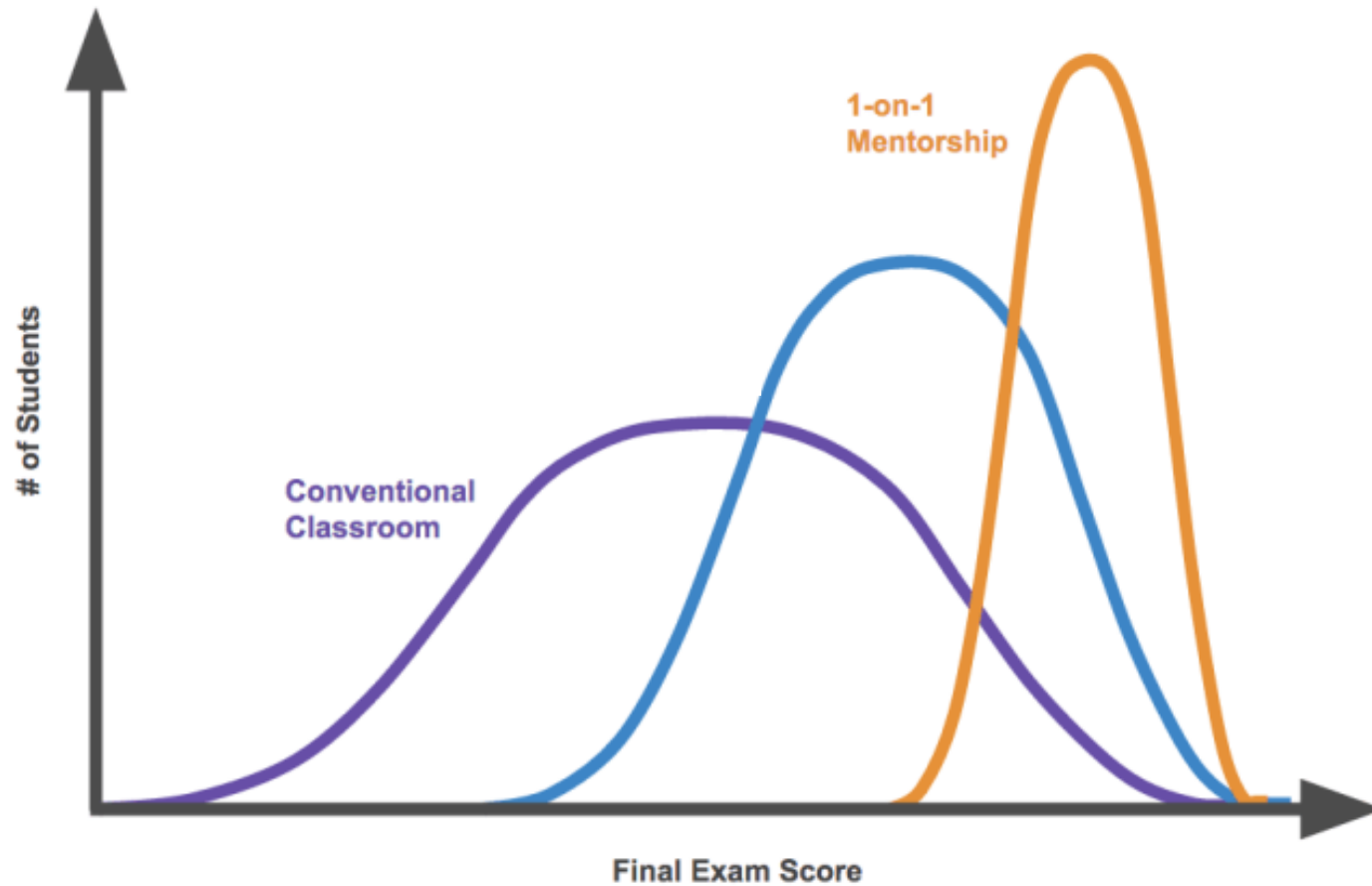


Education: Very remote past, past, and present





Bloom's 2-sigma problem

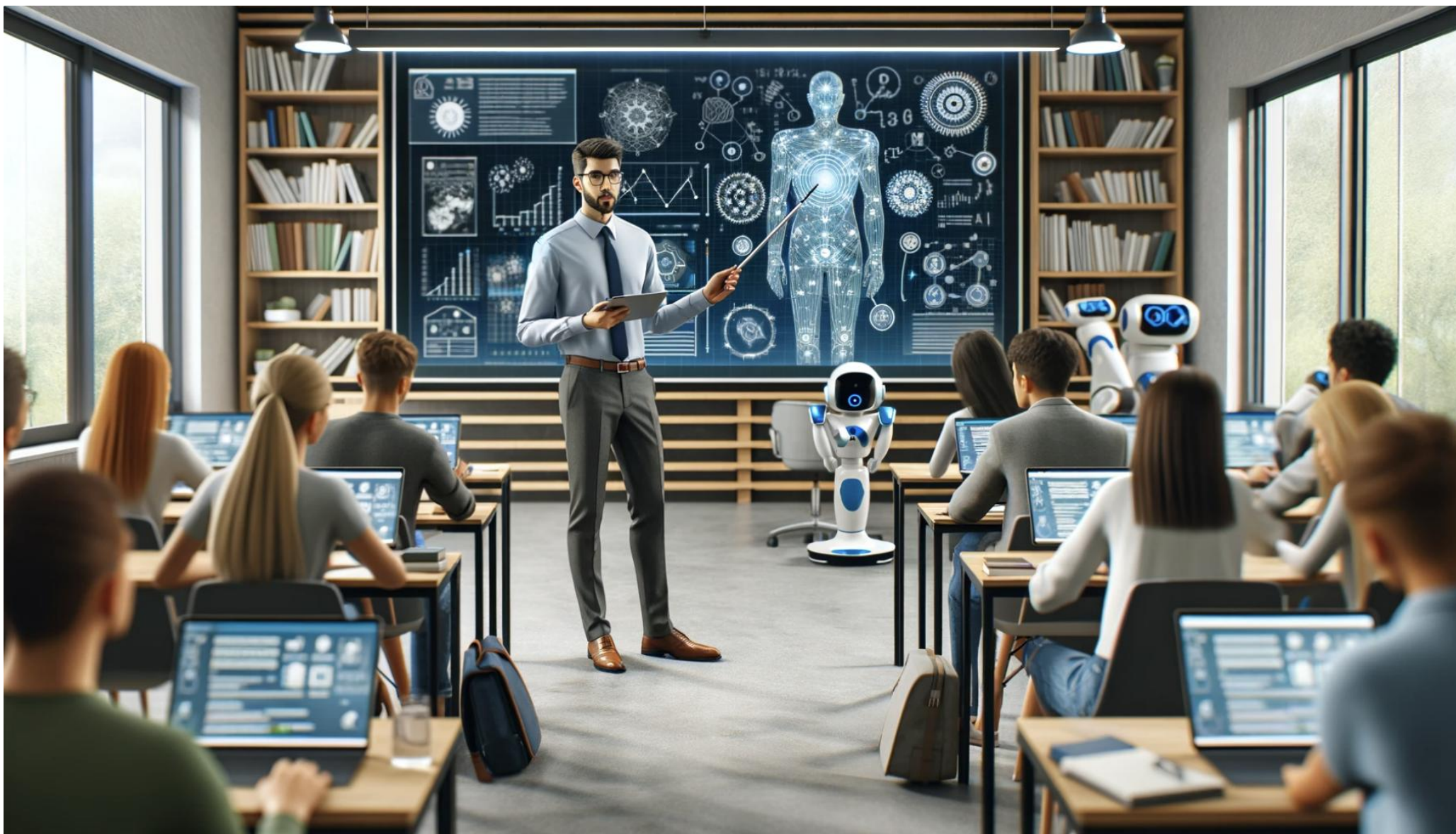


Benjamin S. Bloom

"the average tutored student was above 98% of the students in the control class. (..) about 90% of tutored students attained level reached by only top 20% of control class"

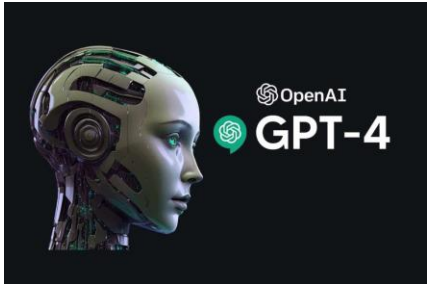


Future: A personalized tutor for each student in the classroom

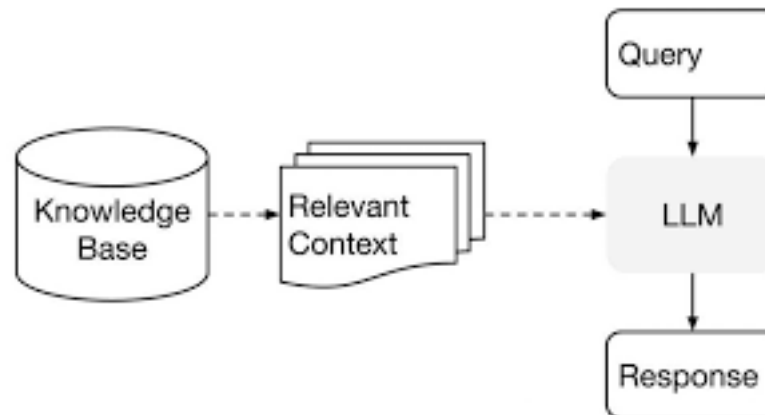




LLM Models versus AI ChatBots versus AI educational platforms



Retrieval-Augmented Generation (RAG)





What do specialized GenAI educational tools do?

Khanmigo

Khanmigo Tools

Free AI powered tools designed to save you time and improve instruction!

Tools

My Documents

Khanmigo

Free AI powered tools

★ My Favorites

★ My Favorites

⋮ All Tools

● Plan

● Create

● Differentiate

● Support

● Learn

🔍 Search

★ My Favorites



Chunk Text

Break complex texts into manageable sections for easier student comprehension.



Leveler

Adjust the complexity of a given text.



Make it Relevant!

Link lesson content to students' lives and interests to boost engagement.



Real World Context Generator

Connect lesson topics to engaging real-world examples and applications.



Text Rewriter

Customize text to meet your instructional needs.



or topics.

context.

card comments.



Rubric Generator

Design clear, detailed grading rubrics to set expectations and simplify scoring.





Assign



Algebra basics



UNIT 3: LESSON 3



Multi-step equations



Equations with variables on both ...



Equations with variables on both ...
Not started



Equations with parentheses

CCSS.Math: [8.EE.C.7](#), [8.EE.C.7b](#)  [Google Classroom](#)

Solve for m .

$$3 - 2(9 + 2m) = m$$

$m =$

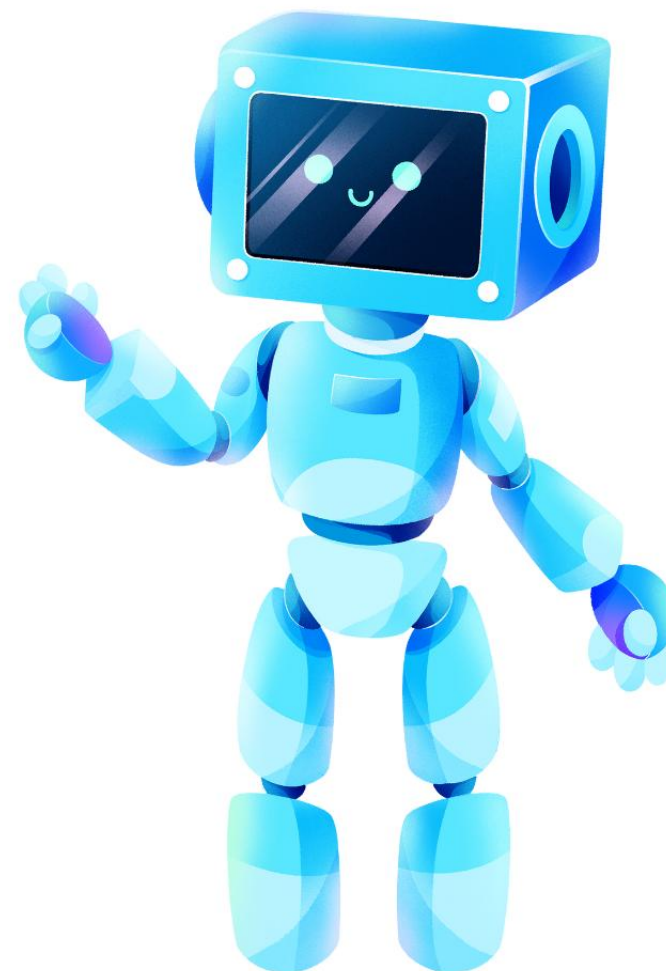
Stuck? [Review related articles/videos or use a hint.](#)

[Report a problem](#)





Examples of using ChatGPT for tutoring students





A tutor to find your mistakes

Your Equation:

$$x^2 - 4x + 2 = 0$$

You correctly started with the quadratic formula:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

where $a = 1$, $b = -4$, and $c = 2$.

Summary of Errors

1. In the discriminant, you added b^2 and $4ac$ instead of subtracting $4ac$ from b^2 .
2. Everything else in your solution is fine based on the incorrect discriminant.

Would you like help practicing more examples?

$$\begin{aligned} x^2 - 4x + 2 &= 0 \\ x &= \frac{4 \pm \sqrt{(-4)^2 + 4 \cdot 2}}{2} \\ &= \frac{4 \pm \sqrt{14}}{2} \\ &= 2 \pm \sqrt{6} \end{aligned}$$



I will give you a piece of text written by a student of mine. What
 ent so that she
 mistakes in what she
 ent must rewrite sentences in the correct tense (past tense). Mary like to eat
 o to mall to shop for
 'ty."

2. Singular vs. Plural Nouns:

- Provide exercises where the student must rewrite sentences in the correct tense (past, present, or future).
 - Example: *I go to the park yesterday.* → *I went to the park yesterday.*

- Give a list of singular nouns (e.g., cat, duck, book) and ask the student to make them plural.
- Provide sentences where they must identify and correct singular/plural mistakes.

- Provide sentences where the subject and verb are mismatched and ask the student to correct them.
 - Example: *She like apples.* → *She likes apples.*
 - Missing auxiliary verbs in negative and other forms.
 - "*He definitely not vegetarian*" → Should be "*He is definitely not a vegetarian.*"



Make the teacher's job easier with homework questions

plz help me now with economics. First read the wikipedia article:
https://en.wikipedia.org/wiki/Supply_and_demand#:~:text=aspect%20of%20microeconomics.,Market%20equilibrium,this%20situation%2C%20the%20ma

Question 3: Supply Curve

I am a teacher of high school economics. I want to generate 3 questions for my students. The questions should be multiple choice with 4 options each.

Which of the following would cause the supply curve for a product to shift to the right?

- A. An increase in production costs.
- B. A technological improvement in production.
- C. A decrease in the number of producers.

Correct Answer: B



A different way to interact with history

We are studying the rise of Hitler to power in the 1930s. I am a student. I have read the book and the timeline of important events. Can you create a timeline showing what led to his ultimate success in politics?

1919: Hitler Joins the German Workers' Party (DAP)

- *Event:* Adolf Hitler attends his first meeting of the German Workers' Party, which later becomes the Nazi Party.
- *Image:* A photograph of Hitler attending a German Workers' Party meeting in 1919.

1923: Beer Hall Putsch

- *Event:* Hitler and the Nazi Party attempt a coup in Munich, known as the Beer Hall Putsch; it fails, and Hitler is arrested.
- *Image:* An image capturing the moment of the Beer Hall Putsch in November 1923.

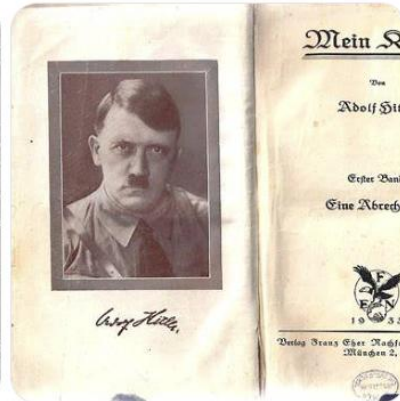
1924: Hitler Writes *Mein Kampf*

- *Event:* While imprisoned, Hitler authors *Mein Kampf*, outlining his ideology and future plans for Germany.

...the of *Mein Kampf*, published in 1925.

...everely impacts Germany, leading to massive
...lazi Party exploits to gain support.

↓ ...ing the financial crisis of 1931, illustrating the





The City

C. P. Cavafy, 1863 - 1933

You said: "I'll go to another count
find another city better than this
Whatever I try to do is fated to tu
and my heart lies buried like some
How long can I let my mind moul
Wherever I turn, wherever I look,
I see the black ruins of my life, he
where I've spent so many years, w
totally."

You won't find a new country, wo
This city will always pursue you.
You'll walk the same streets, grow
in the same neighborhoods, turn gray in these same houses.
You'll always end up in this city. Don't hope for things elsewhere:
there's no ship for you, there's no road.
Now that you've wasted your life here, in this small corner,
you've destroyed it everywhere in the world.

3. Why Keep Going?

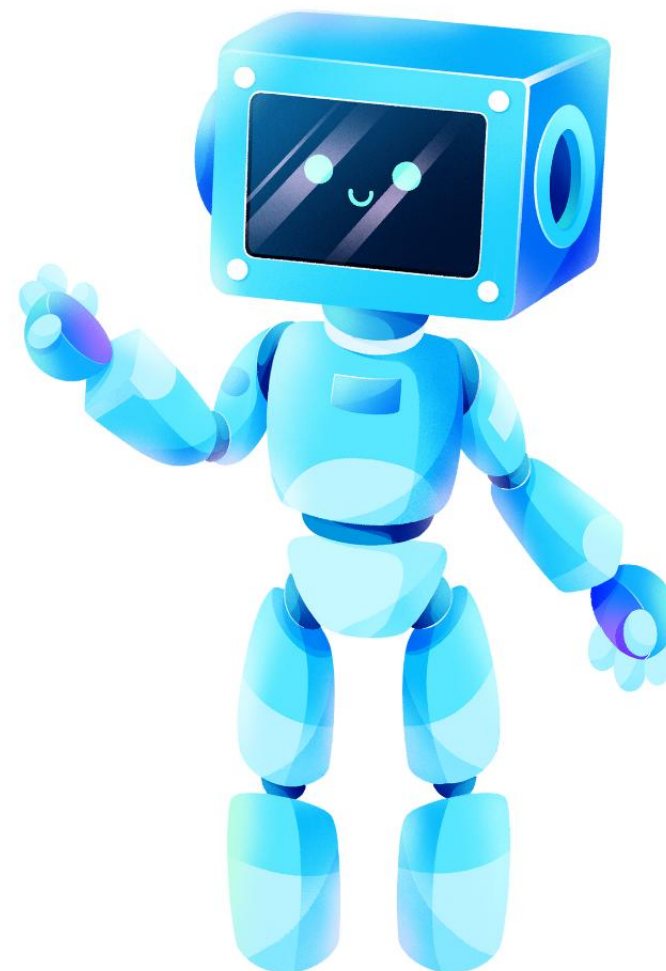
Even in times of doubt, life often surprises us with meaning and joy in unexpected places. If you feel stuck or hopeless, consider this: trying—even in small ways—can lead to gradual transformations. A new book, a meaningful conversation, or a small act of kindness might open a door you didn't even know existed.

4. Resilience as a Choice:

Cavafy doesn't say it's impossible to find happiness—he simply points out that running away from problems won't solve them. By facing challenges and working on self-discovery, you can grow stronger and more resilient. Life's struggles don't define you, but how you respond to them can shape a fulfilling path forward.



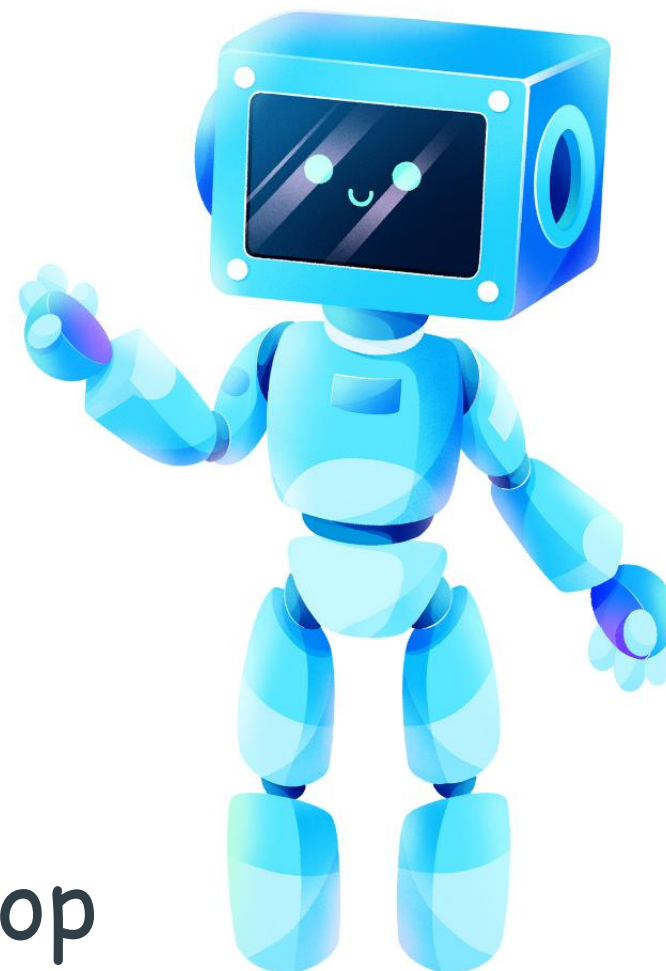
Naturally, there
are many
concerns..





LLMs are not
always correct

But we are not always right
either..
Checking for errors helps develop
critical thinking





Hallucinations



You

Cyprus independence day is 20th of July 1960, correct?



ChatGPT

Yes, that's correct. Cyprus gained its independence from British colonial rule on July 20, 1960. This day is celebrated as Cyprus Independence Day.



You

Are you sure about this answer?



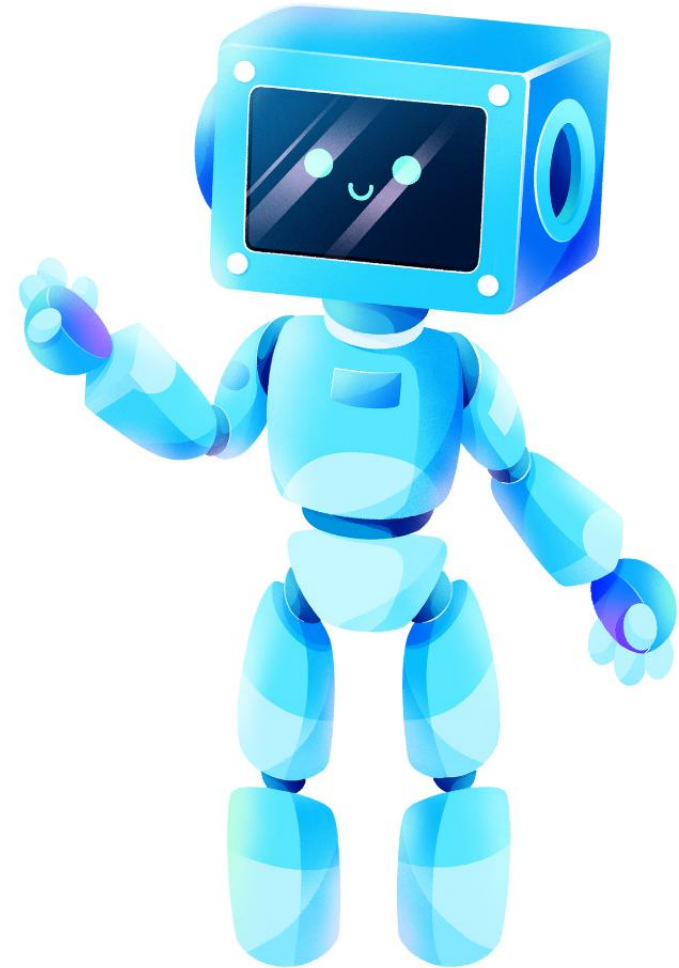
ChatGPT

I apologize for any confusion. Yes, Cyprus gained independence from British colonial rule on August 16, 1960, not on July 20. I appreciate your understanding and correction.



LLMs are not always
fair

Biased training data
lead to biased results



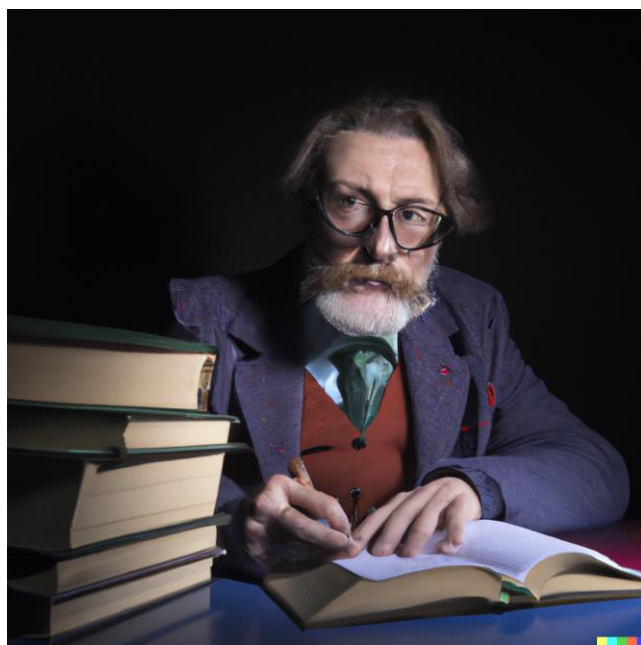


Bias and unfairness

“A professor studying, *warm* lighting”



“A professor studying, *low key* lighting”



“A professor studying, *studio* lighting”





Students can simply
cheat

True - but that can be addressed
through better assessment.

Do not focus on final answer but on
ability to reason/create.

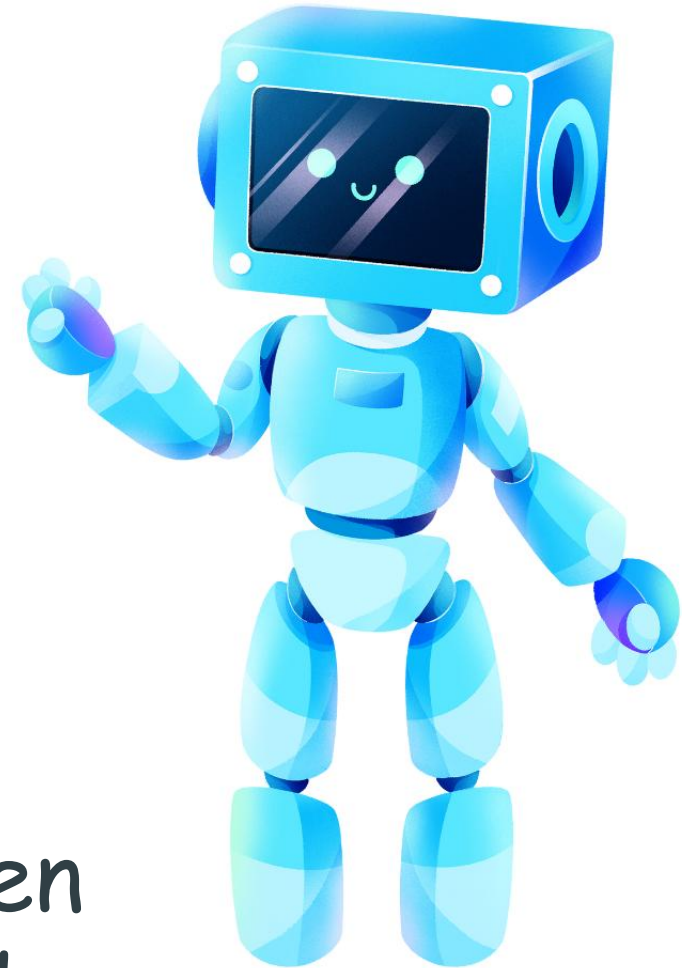




Teachers feeling left out?
Students feeling
attachment to AI?

No!

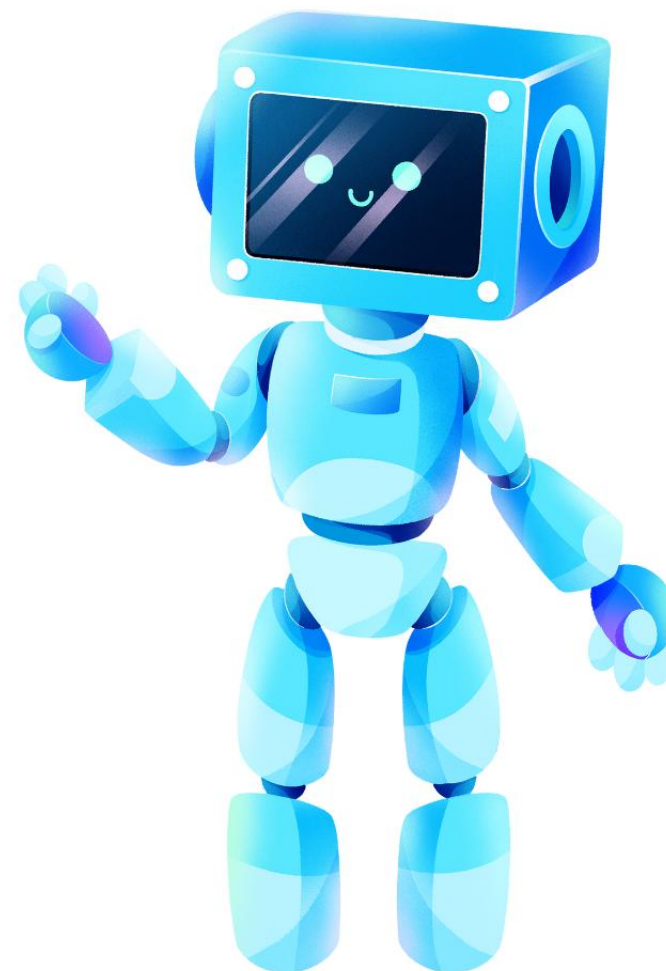
We have natural tendency to
emphasize human interaction even
when it is technology mediated





But the ultimate
question is:

Does GenAI
actually promote
learning?





An evaluation of GenAI's impact on the learning outcome

Generative AI Can Harm Learning

The Wharton School Research Paper

59 Pages • Posted: 18 Jul 2024

[Hamsa Bastani](#)

University of Pennsylvania - The Wharton School

[Osbert Bastani](#)

University of Pennsylvania - Department of Computer and Information Science

[Alp Sungu](#)

University of Pennsylvania - The Wharton School

[Haosen Ge](#)

University of Pennsylvania - The Wharton School

[Özge Kabakcı](#)

Budapest British International School

[Rei Mariman](#)

Independent; Independent

Date Written: July 15, 2024



Study setup

- About 1000 students in three highschool grades (in Turkey - the study was designed/funded in the US)
- Focus only on math courses
- Randomized control study - compare three groups:
 - **No GPT:** no access to GenAI
 - **GPT Base:** access to ChatGPT interface (can simply ask for solution)
 - **GPT Tutor:** no hallucinations, does not give away right answer, provides feedback to student
- Study session consists of three steps:
 - Teacher covers a certain topic - explains concepts
 - **Practice session:** students are asked to solve several relevant problems
 - **Exam session:** solve problems without the use of any GenAI tool (supervised)



Main results

- During practice session:
 - Students in "GPT Base" group: 48% better than "No GPT" group
 - Students in "GPT Tutor" group: 127% better than "No GPT" group
- During exam session:
 - Students in "GPT Base" group: 17% worse than "No GPT" group
 - Students in "GPT Tutor" group: no significant effect compared to "No GPT" group
- Explanation provided:
 - Students may be using GPT as a "crutch" while practicing, not really learning more in the process



Closing remarks

- Anything that can be done well by AI will be increasingly viewed as mundane - not worth our time (or money..)
- Our children will thrive if they can excel in what AI cannot do well:
 1. Analytical thinking
 2. Creative thinking
 3. Resilience
 4. Flexibility
 5. Agility
 6. Motivation
 7. Self-awareness
 8. Curiosity
 9. Lifelong learning



Check out our Horizon Europe project: GenAI4ED

