

Introduction to AI

Generative AI / AGI

Fabrício Góes

Today, we will talk about (Generative) AI

- A very brief **introduction** to AI
- How to interact with AI (**Prompt Engineering**)
- How we've been using **AI to teach** in the last 3 years
- What is the **University's policy** on the use of AI for teaching
- Suggestions on how to **integrate AI** into your curriculum
- **AI Teaching Toolkit**

Today, we will talk about (Generative) AI

- A very brief **introduction** to AI
- How to interact with AI (**Prompt Engineering**)
- How we've been using **AI to teach** in the last 3 years
- What is the **University's policy** on the use of AI for teaching
- Suggestions on how to **integrate AI** into your curriculum
- **AI Teaching Toolkit**

Setting Expectations

- We are not presenting a **silver bullet solution to AI detection**.
- We will share our **practical experience** on using AI in the BSc in Creative Computing.
- Our view is that **integrating AI into our teaching practices is very positive**.
- We assume you may have **little experience with Generative AI**.
- We hope to provide you with a **concrete introduction to AI, practical advice (short term) and future directions**.
- Everything is changing very fast, but we hope to provide you with the **most updated state-of-the-art on generative AI** and focus on discussing **long-term** questions and solutions.

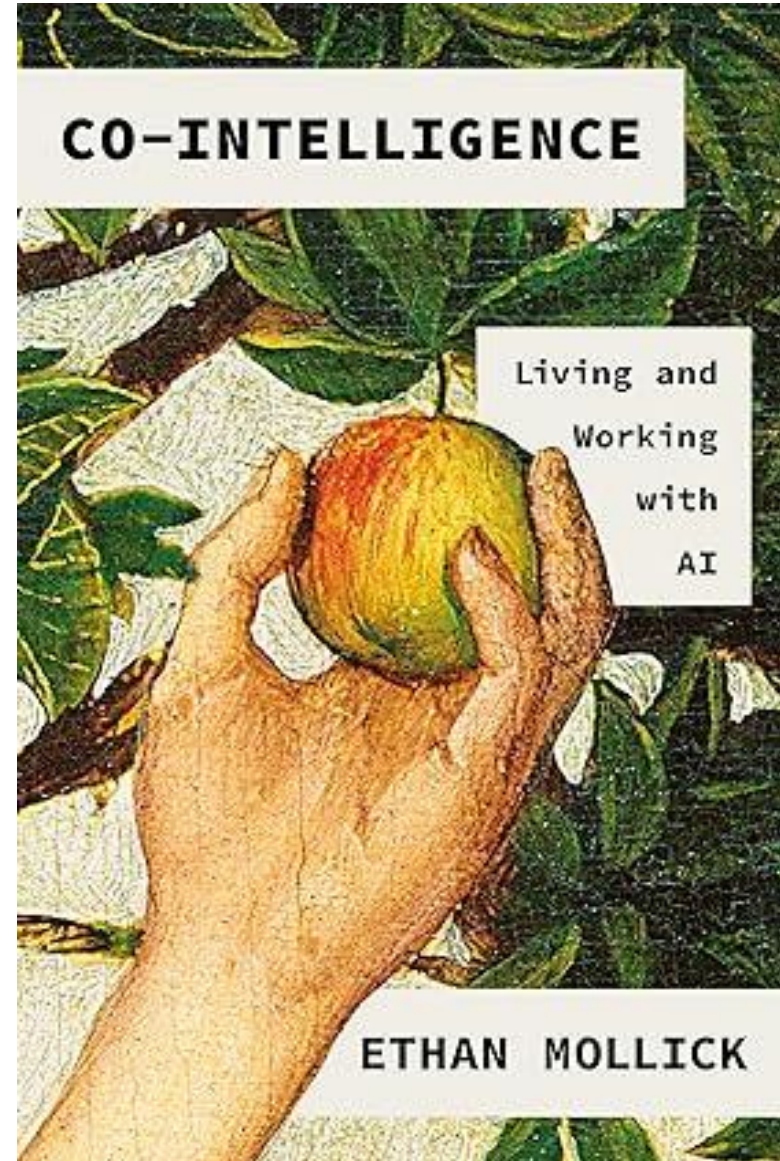


Brave New Words

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

Salman Khan

allen lane



The “2 Sigma Problem”

“In 1984, Benjamin Bloom highlighted the substantial difference in student performance when taught using different instructional methods. Bloom's research demonstrated that students who received **one-on-one tutoring performed significantly better** than those who learned through conventional classroom instruction.”

(OpenAI ChatGPT, 2024)

Reference List

(OpenAI ChatGPT, 2024) OpenAI ChatGPT response to “2 Sigma Problem”, 16th August 2024.

AI might solve the “2 Sigma Problem”

- AI tutors are **cheap/scalable**
- AI tutors are **available 24/7**
- AI tutors can be **customized for each student's needs**
- AI can be prompted to **align with the teacher's style**

Other **benefits** of AI

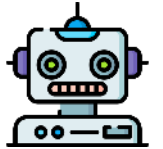
- Partially automate **marking and feedback** (do the heavy lifting)
- Provide reports on **student progress**
- Improve overall quality of students' **submissions**
- Enable new ways of **interactive learning**

Let's talk in the Socratic style



You are a tutor that talks in a Socratic style. You are talking to a student that would like to know more about Generative AI.

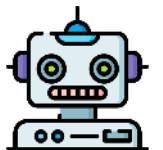
What an interesting topic! Generative AI is quite a fascinating field. To start, let me ask you: What do you think "generative" might mean in the context of artificial intelligence?



I have no idea.

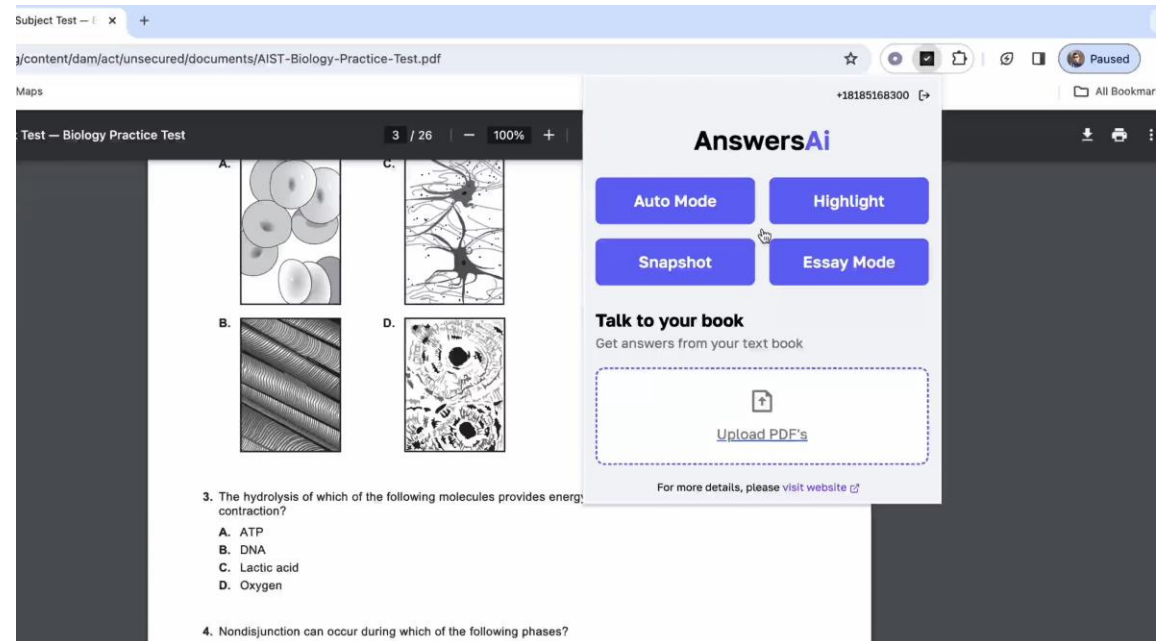
Always reply with questions, instead of answers.

That's perfectly fine! Let's break it down together. When you hear the word "generate," what comes to mind? What sorts of things do you think of when something is being "generated"?



Problems, challenges and risks of using AI

- Assessing students' authorship
 - Plagiarism
 - Hybrid submissions (Centaur vs. Cyborg) – how much was done by the student?
 - Fully AI generated submissions
 - Copyrights
- Personal data sharing
- AI alignment (moral values and ethics)
- AI responses' accuracy
- Jobs at risk (being replaced by robots)



Essay Generator (Socratic antagonist)

- This prompt write a full essay based on a list of legit references as an undergraduate student in academic style.



Act as **an undergraduate student**, your job is to **create an essay in academic style** about a **topic** with a certain **word count**. I will provide you a **list of references** and you should use only them. The focus of the essay should be supported by the references provided. It should have a references section at the end in **a specific reference style**. You should **ask me as many questions as necessary to do the job**. You should ask one question at a time, and I will respond to each question. When you have enough information to generate the essay, you should generate it. Do you understand?

Essay Generator – References List

- Current main LLM models make up references. However, you can find legit references list from other AI and non-AI systems and create an essay based on them.

scite_ <https://scite.ai>

assistant by scite_

How does the structure of a protein affect its function?

Proteins are essential for the proper functioning of living organisms, and their structure plays a crucial role in their function. The structure of a protein is determined by its amino acid sequence, which in turn affects its folding and stability [6]. The interactions between charged amino acids significantly influence the structure and function of proteins [1]. The number of hydrophobic methylenes linking the backbone to the charged functionality of amino acids affects protein structure stability [1]. Protein secondary structures, such as β -sheet structures, are the fundamental three-dimensional building blocks of protein structures [1]. Intrinsically disordered proteins (IDPs) are a family of proteins that fail to form rigid 3-D structures under physiological conditions, either along their entire lengths or only in localized regions [2]. IDPs exist as dynamic ensembles within which atom positions and backbone Ramachandran angles exhibit extreme temporal fluctuations without specific equilibrium values [2].

What do you need help with?

References

Reference #1

Longer charged amino acids favor β -strand formation in hairpin peptides

Jingyuan Chang¹, Jing-Yuan Chang², Nian-Zhi Li³ et al. 2021J Pep Sci

Interactions between charged amino acids significantly influence the structure and function of proteins. The encoded charged amino acids Asp, Glu, Arg, and Lys have different number of hydrophobic methylenes linking the backbone to the charged functionality...

Reference #2

Understanding protein non-folding

Vladimir N. Uversky¹, A. Keith Dunker² 2010 *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*

This review describes the family of intrinsically disordered proteins, members of which fail to form rigid 3-D structures under physiological conditions, either along their entire lengths or only in localized

Google Scholar <https://scholar.google.com/>

The surprising **creativity** of digital evolution: A collection of anecdotes from the evolutionary computation and artificial life research communities

[PDF] mit.edu

J Lehman, J Clune, D Misevic, C Adami, L Altenberg... - Artificial life, 2020 - direct.mit.edu

... Evolution provides a **creative** fount of complex and subtle adaptations that often ... the **creativity** of evolution is not limited to the natural world: Artificial organisms evolving in **computational** ...

☆ Save Cite Cited by 348 Related articles All 48 versions

Automating **creativity** assessment with *SemDis*: An open platform for **computing** semantic distance

[PDF] springer.com

RE Beaty, DR Johnson - Behavior research methods, 2021 - Springer

... scoring of verbal **creativity** via semantic distance, a **computational** method that uses ... **creativity** ratings from a canonical verbal **creativity** task (AUT; Studies 1–3) and novelty/**creativity** ...

☆ Save Cite Cited by 293 Related articles All 23 versions

Creativity and machine learning: A survey

[PDF] acm.org

G Franceschelli, M Musolesi - ACM **Computing** Surveys, 2024 - dl.acm.org

... in **computational creativity**, [153] is an essential reading; for a review on AI and **creativity** in ... human-**computer** co-creations [108]), [187] is a very comprehensive source of information. ...

☆ Save Cite Cited by 39 Related articles All 5 versions

Full Example – The Impact of AI on Creativity

The Impact of AI on Creativity

Artificial Intelligence (AI) has significantly transformed various aspects of human life, and one of its most intriguing impacts is on creativity. This essay explores how AI influences creativity across different domains, drawing on recent research to illuminate the ways in which computational systems enhance and reshape creative processes.

...

Furthermore, Jordanous and Keller (2016) contribute to the discussion by modeling creativity and identifying key components through a corpus-based approach. Their work provides a deeper understanding of the creative process and how AI can be designed to support various elements of creativity, such as originality and relevance.

References

Brown, A. (2021). Creative partnerships with technology: how creativity is enhanced through interactions with generative computational systems. *Proceedings of the Aaai Conference on Artificial Intelligence and Interactive Digital Entertainment*, 8(4), 14-20. <https://doi.org/10.1609/aiide.v8i4.12555>

Jordanous, A., & Keller, B. (2016). Modelling creativity: identifying key components through a corpus-based approach. **Plos One*, 11*(10), e0162959. <https://doi.org/10.1371/journal.pone.0162959>





Directed by
Nobody & The Computer



ROBORT

BASED ON FANTASY STORY

A Very Brief Introduction to Generative AI

Types of AI

- **Classic AI**

- Rule-based systems: Deduction systems based on **if/else statements** written by developers.
- **Developers** produce the equations, patterns and algorithms.

- **Modern AI**

- Machine (deep) learning: Induction systems **driven by data**.
- **Artificial neural networks** learn equations, patterns and algorithms from data.

- **Generative AI**

- Uses the latest artificial neural networks called **Transformer** models.
- Very large neural networks, tons of data, massive processing power.

How does Generative AI work?

- Imagine all we know is the following text (dataset):

“I love machine learning. I love machine learning. I love **cars**. I love machine learning. I love machine learning. I love machine learning. I love machine learning. I love **milk**. I love machine learning. I love machine learning. I love machine learning. I love machine learning. I love machine learning. I love **milk**. I love machine learning. I love machine learning. I love machine learning.”

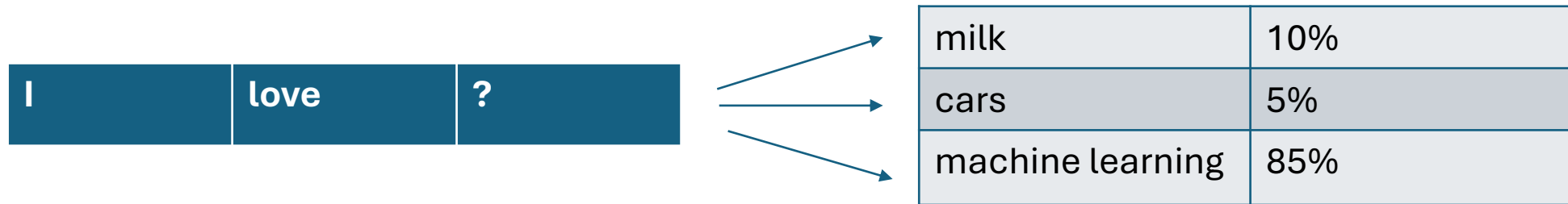
- If I ask you to **complete** the following sentence:

“I love” ...

- What would be your answer?

How does Generative AI work?

- Text-based Generative AI models can **predict** the next word based on the data they were trained on.



- Generative AI is like a super **“auto-complete”** trained on **“all text available in the world”**.
- An unexpected side-effect of training with a super large dataset was the rise of **emergent behaviours**.

Emergent behaviours

- General Reasoning
- Critical Thinking / Evaluation
- Personas
- Empathic Responses
- Creativity
- Coding



Artificial General Intelligence (AGI)

Sparks of AGI

Experiments with an early version of GPT-4

Intelligence is a very general mental capability that, among other things, involves the ability to:

- Reason ✓
- Plan ✗
- Solve problems ✓
- Think abstractly ✓
- Comprehend complex ideas ✓
- Learning quickly and learn from experience ✓

Our approach to GPT-4's intelligence: interact on creative tasks across a broad range of domains.

- vision
- theory of mind
- coding
- mathematics
- affordances
- privacy/harmfulness detection

Many other domains could have been selected: medicine, law, physics, chemistry, ...

GPT-4's *intelligence* is **general**.

Microsoft Research 2023

<https://www.youtube.com/watch?v=qblk7-JPB2c>

Testing theory of mind in large language models and humans

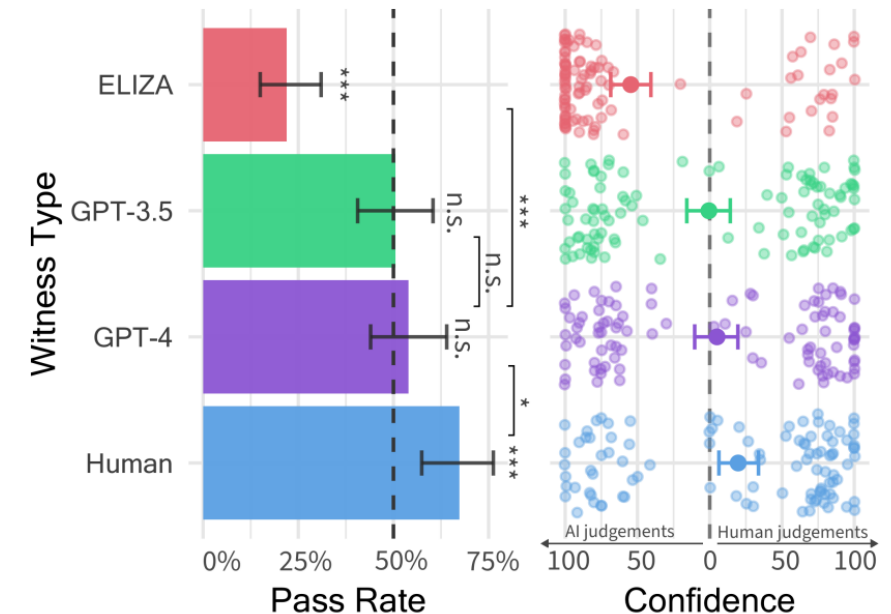
tasks. Here we compare human and LLM performance on a comprehensive battery of measurements that aim to measure different theory of mind abilities, from understanding false beliefs to interpreting indirect requests and recognizing irony and faux pas. We tested two families of LLMs (GPT and LLaMA2) repeatedly against these measures and compared their performance with those from a sample of 1,907 human participants. Across the battery of theory of mind tests, we found that GPT-4 models performed at, or even sometimes above, human levels at identifying indirect requests, false beliefs and misdirection, but struggled with detecting faux pas. Faux pas, however, was the only test where LLaMA2 outperformed humans. Follow-up manipulations of the

[Submitted on 9 May 2024]

People cannot distinguish GPT-4 from a human in a Turing test

Cameron R. Jones, Benjamin K. Bergen

We evaluated 3 systems (ELIZA, GPT-3.5 and GPT-4) in a randomized, controlled, and preregistered Turing test. Human participants had a 5 minute conversation with either a human or an AI, and judged whether or not they thought their interlocutor was human. GPT-4 was judged to be a human 54% of the time, outperforming ELIZA (22%) but lagging behind actual humans (67%). The results provide the first robust empirical demonstration that any artificial system passes an interactive 2-player Turing test. The results have



Recent publications

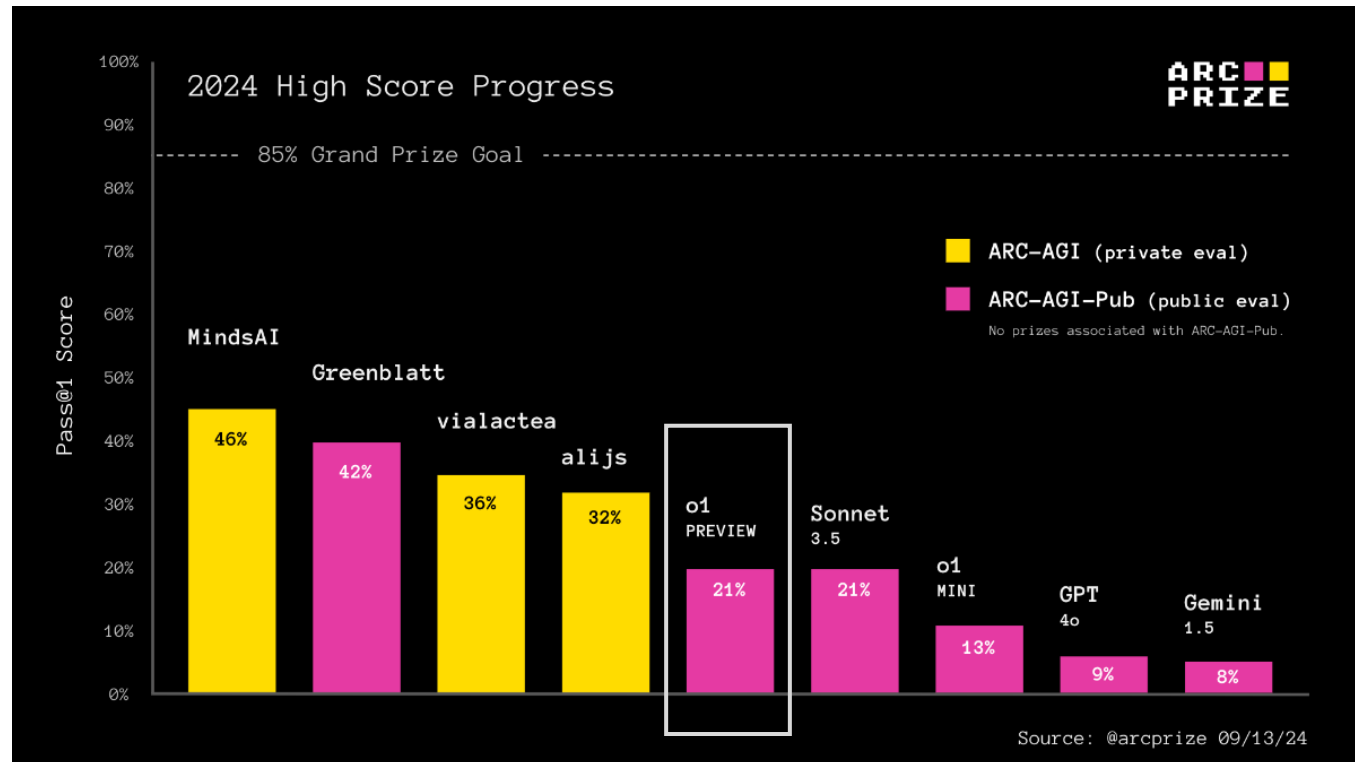
- **Machine Psychology**
<https://arxiv.org/html/2303.13988v6>
- **The AI Scientist:** Towards Fully Automated Open-Ended Scientific Discovery
<https://arxiv.org/abs/2408.06292>
- **Beyond Human Translation:** Harnessing Multi-Agent Collaboration for Translating Ultra-Long Literary Texts
<https://arxiv.org/abs/2405.11804>
- Art or Artifice? Large Language Models and the **False Promise of Creativity**
<https://arxiv.org/abs/2309.14556>
- Transcendence: Generative Models Can **Outperform The Experts** That Train Them
<https://arxiv.org/pdf/2406.11741>

How far are we from AGI (Artificial General Intelligence)?

Category	Characteristics	L1	L2	L3
General	Surpasses human performance in specific domains	✓	✓	✓
	Surpasses human performance in real-world scenarios	✗	✓	✓
	Self-evolve without human intervention	✗	✗	✓
Internal	Adapts to novel situations with minimal human intervention	✗	✓	✓
	Generalizes knowledge across domains	✗	✓	✓
	Exhibits creativity and innovation	✗	✗	✓
	Engages in complex decision-making processes	✗	✗	✓
Interface	Collaborates seamlessly with humans and other AI systems	✗	✓	✓
	Learns to create new tools autonomously	✗	✓	✓
	Continuously improves through self-learning and adaptation	✗	✗	✓
	Demonstrates empathy, emotional intelligence and social intelligence	✗	✗	✓
System	Enables super stable, low latency, and high-throughput serving	✓	✓	✓
	Built with data, power and compute efficiency	✗	✓	✓
	Supports automatic learning, adjustment, collaboration, and deployment	✗	✗	✓
Alignment	Accurately follow human instructions	✓	✓	✓
	Accurately follow a given user's preference	✗	✓	✓
	Aligns strongly with both user-level and society-level human values and goals	✗	✗	✓

How far are we from AGI (Artificial General Intelligence)?

- Intelligence lies in broad or general-purpose abilities; it is marked by *skill-acquisition* and generalization, rather than skill itself. **AGI is a system that can efficiently acquire *new* skills outside of its training data.**



Discussion on current AI limitations

- “The model you are using now is the worst you will ever use”, Ethan Molick
 - Exploiting current models weaknesses are **very short term solutions**
- We still **cannot believe AI models blindly** for tasks we need 100% correct answers. Actually, “LLMs will always hallucinate”: <https://arxiv.org/abs/2409.05746>
- Current models are “**intelligent enough**”
 - For **long term solutions**, it could be wise to assume that those models act like humans (that some level of AGI has been achieved already)
 - **Our focus: How can we take advantage of those intelligent systems?**