TESCREAL
and
Regulation of AI/AGI
Feb. 1, 2024 - Brant, Kevin, Torben

Since late 2022, we've been attributing "AI" - unintentionally, perhaps - to be almost exclusively ChatGPT.
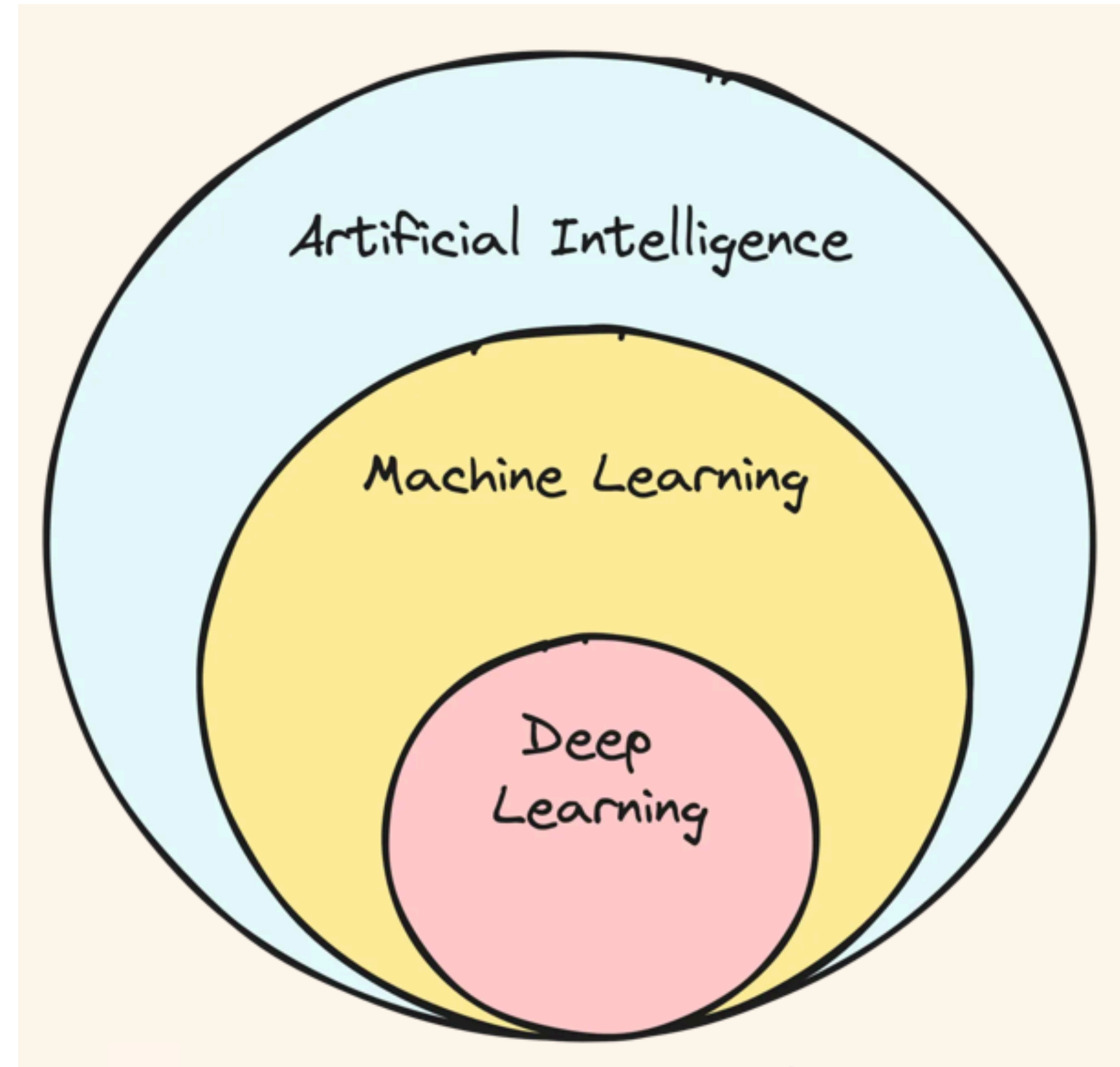
- ChatGTP was introduced (as a free research preview) on Nov. 30, 2022;
- On Feb. 1, 2023, OpenAI introduced ChatGPT Plus and
- On Mar. 1, ChatGPT API was introduced [Application Programming Interface, a set of defined rules that enable different applications to communicate with each other]
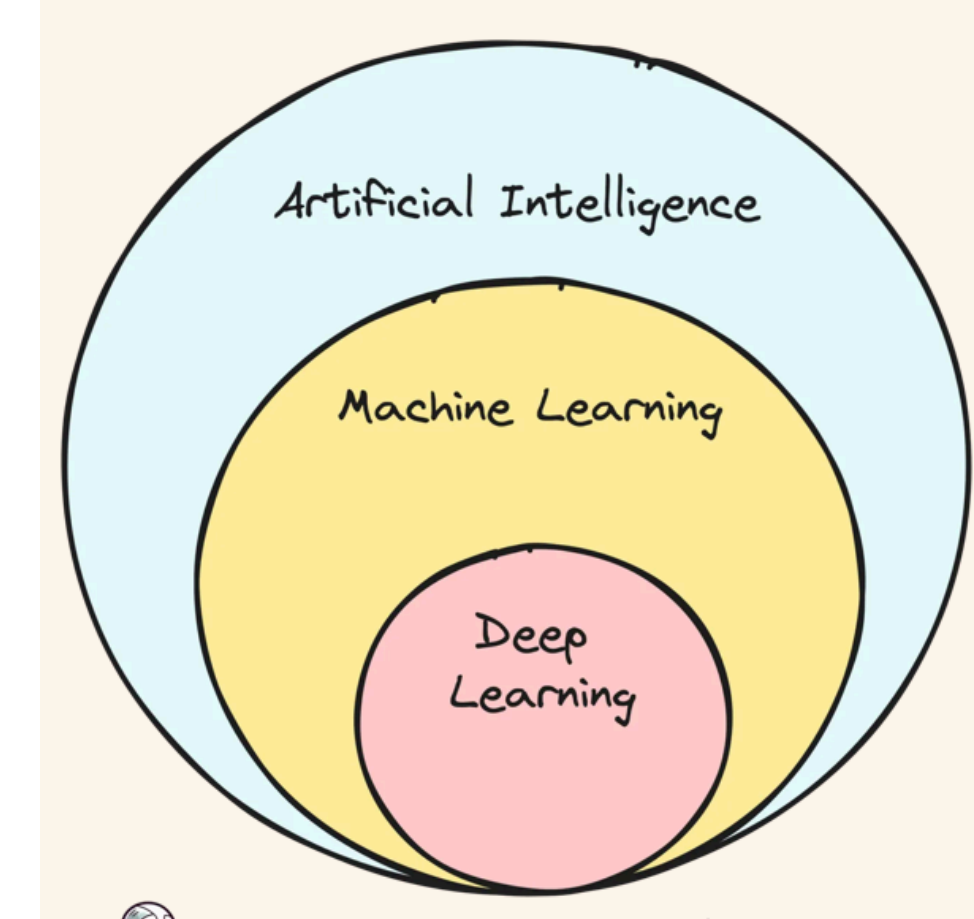- Today, we are on 4th gen. ChatGPT

To follow what is going on, go here and subscribe to a 5 min newsletter every day: https://aitoolreport.beehiiv.com/subscribe?ref=k6hcs6dLI3

But AI/AGI comprises many things. It refers to anything that allows a machine to do something that normally would require human intelligence, like:

1) Machine Learning

2) Natural Language Processing

3) Robotics
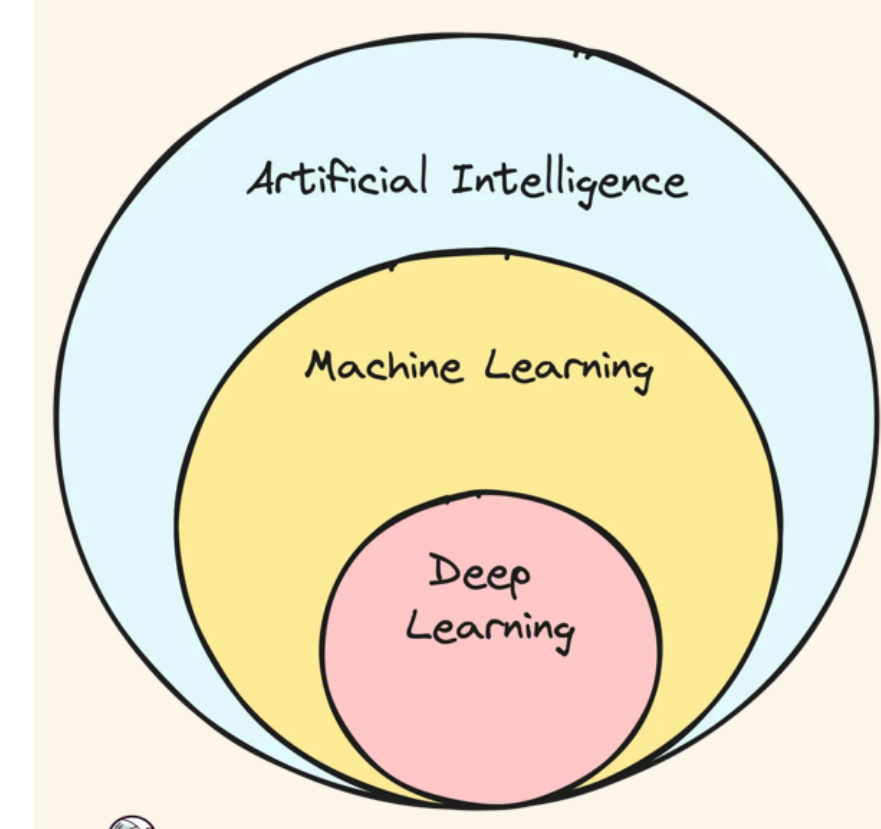
4) Neural Networks

5) Deep learning

1) *Machine Learning*

A large part of AI is machine learning. ML allows the computer to learn and improve itself from existing data, w/o explicitly being programmed for it.

This kind of AI also encapsulates other kinds of AI and their associated algorithms (generally very fancy and advanced statistics).
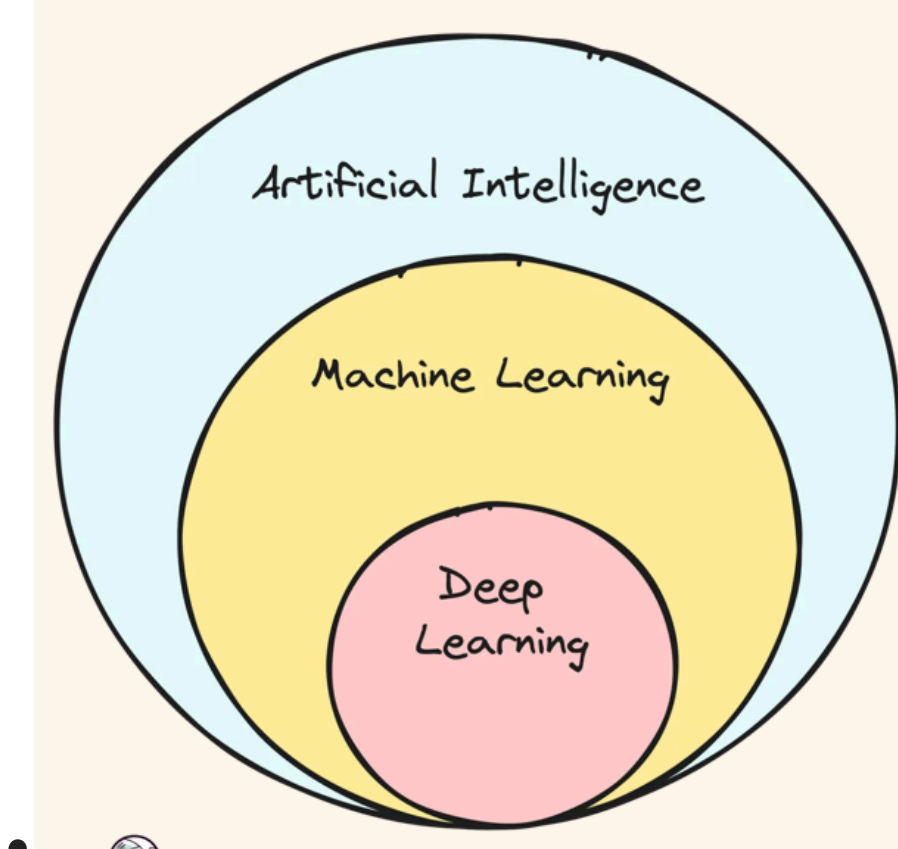
2. *Natural Language Processing*

Natural Language Processing (NLP) involves the development of algorithms that can understand, interpret, and respond to human languages.

NLP is a blend between computational linguistics with statistical, machine learning, and deep learning models.

This branch of AI includes applications such as ChatGPT, Siri, DALL-E/Google Lumiere, Lydia(DeepMind's text-to-music) or DreamTrack (text to a specific artist's voice)
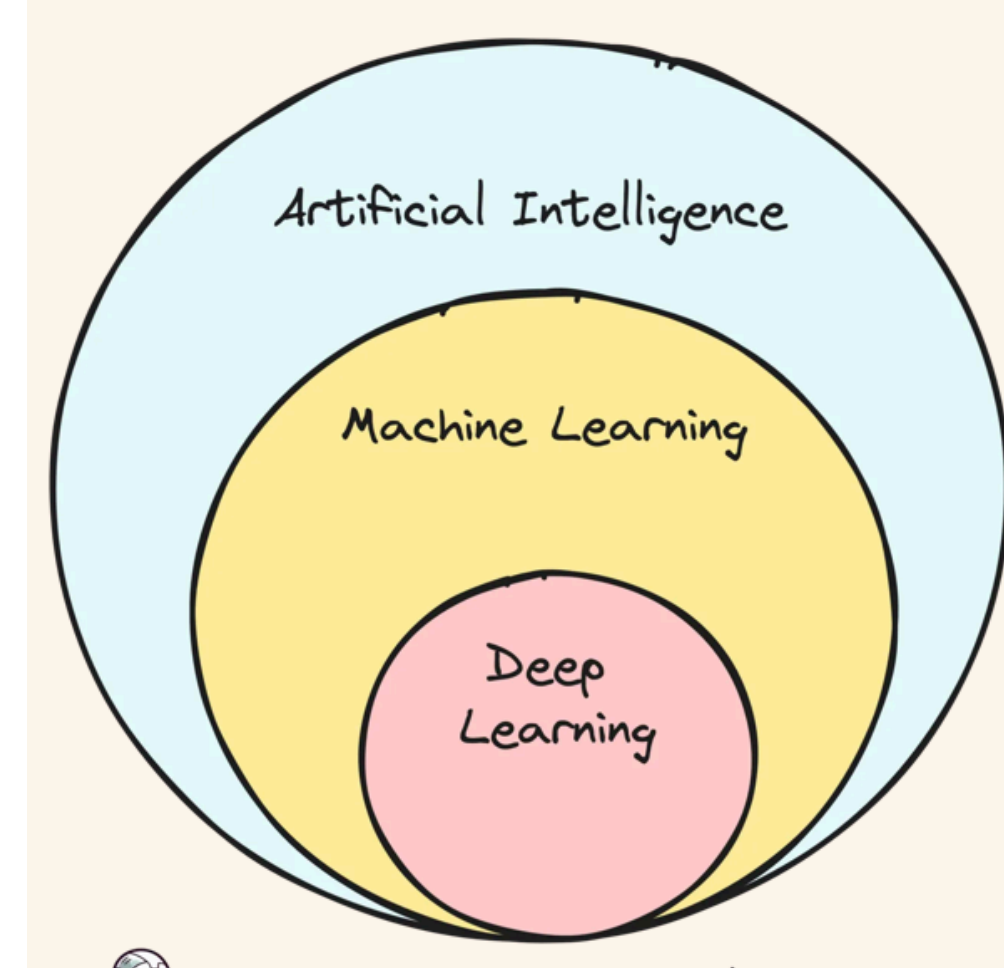
3. Robotics

While a lot of AI is software, there are significant applications in the robotics industry.

Robotics has been around for several years, but we're now seeing an explosion in comparison to previous years due to the growth of the AI industry as a whole.

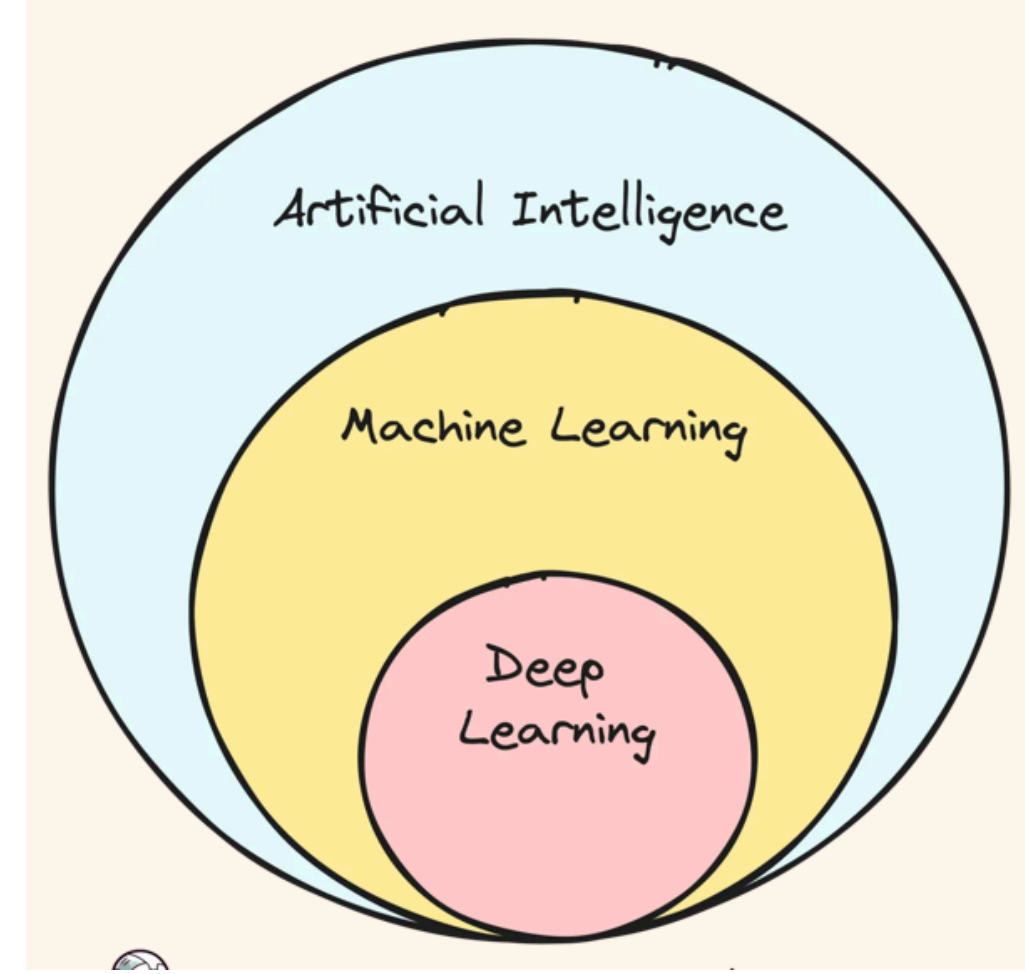Think of this subset of AI as a "hands-on" application of AI.

Artificial Intelligence

Machine Learning

Deep Learning

4) *Neural Networks*

Neural networks are modeled after our brains; this type of AI is designed to learn and make decisions like humans.

The biggest advantages of a neural network compared to other kinds of AI is that it can be trained to recognize patterns in data that are too complex for a traditional algorithms.

Since they're able to recognize complex patterns in data, it allows us to apply it to a variety of advanced applications and fields, such as image and video recognition and predictive analysis.

5. Deep Learning

Deep learning is a subset of neural networks and has a primary application in computer vision, self-driving vehicles, and speech recognition.

A deep learning network generally consists of many layers of processing for classification or predictions.

Deep learning can be scaled across many industries and applications, and is very widely used in today's market to solve some of the most complex problems we face.

*) **TESCREAL** (Acronym for 7 isms that dominate the AI/Technology conversation:

Transhumanism
Extropianism
Singularitanism
Cosmism
Rationalism
Effective Altruism
Longtermism

They are . . .

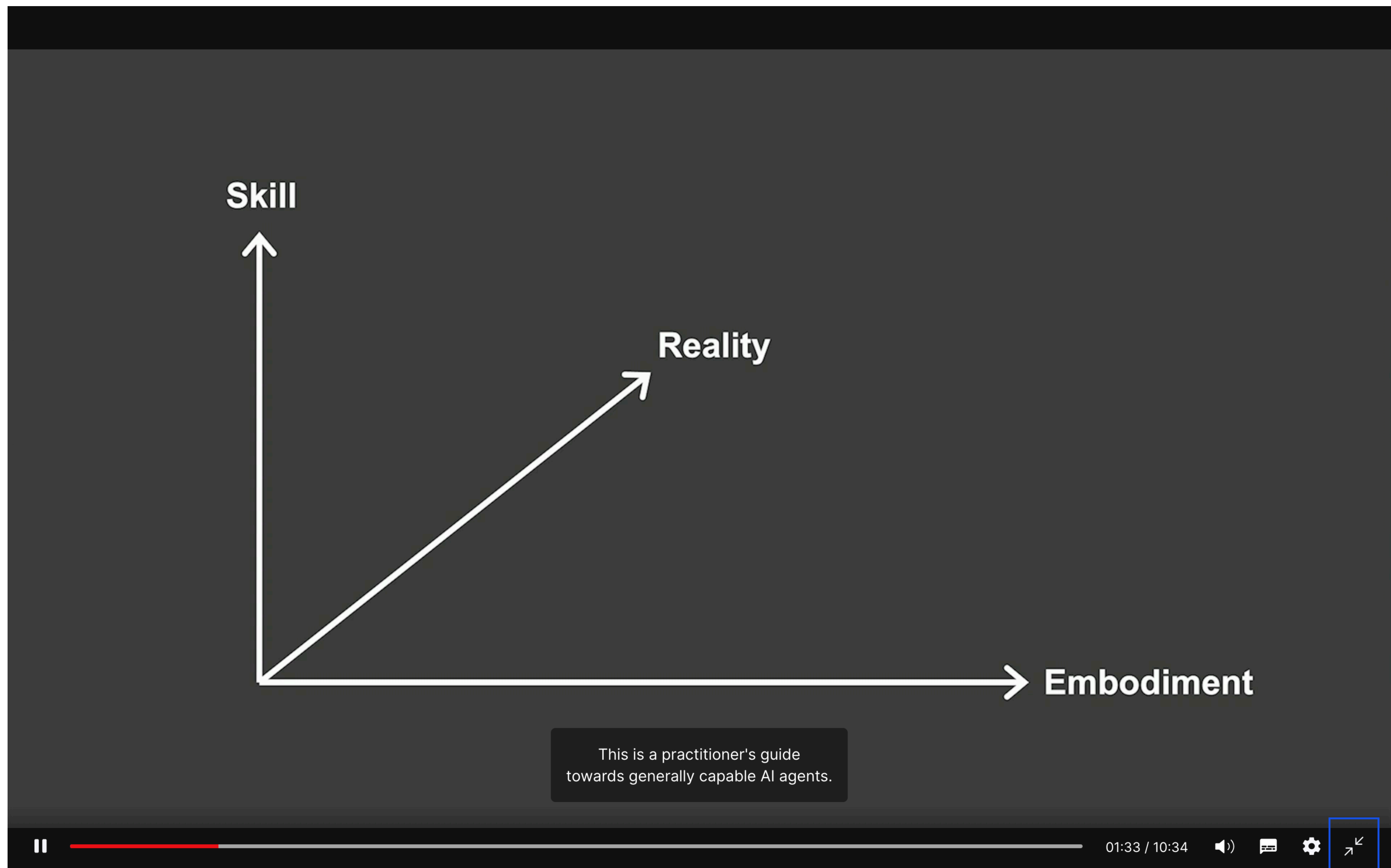**TESCREAL** (Acronym for 7 isms that dominate the AI/Technology conversation:

1. **T**ranshumanism: the belief that humans can use technology to improve themselves, both physically and mentally.

2. **E**xtropianism: a branch of transhumanism that focuses on the idea of progress and believes that the future can be improved through technology.

3. **S**ingularitarianism: the belief that artificial intelligence will eventually become so advanced that it will lead to a significant transformation of human existence, with potential outcomes ranging from utopia to existential risk.

4. **C**osmism: the belief in the potential for humans to explore and colonize space, as well as the resurrection of the dead.

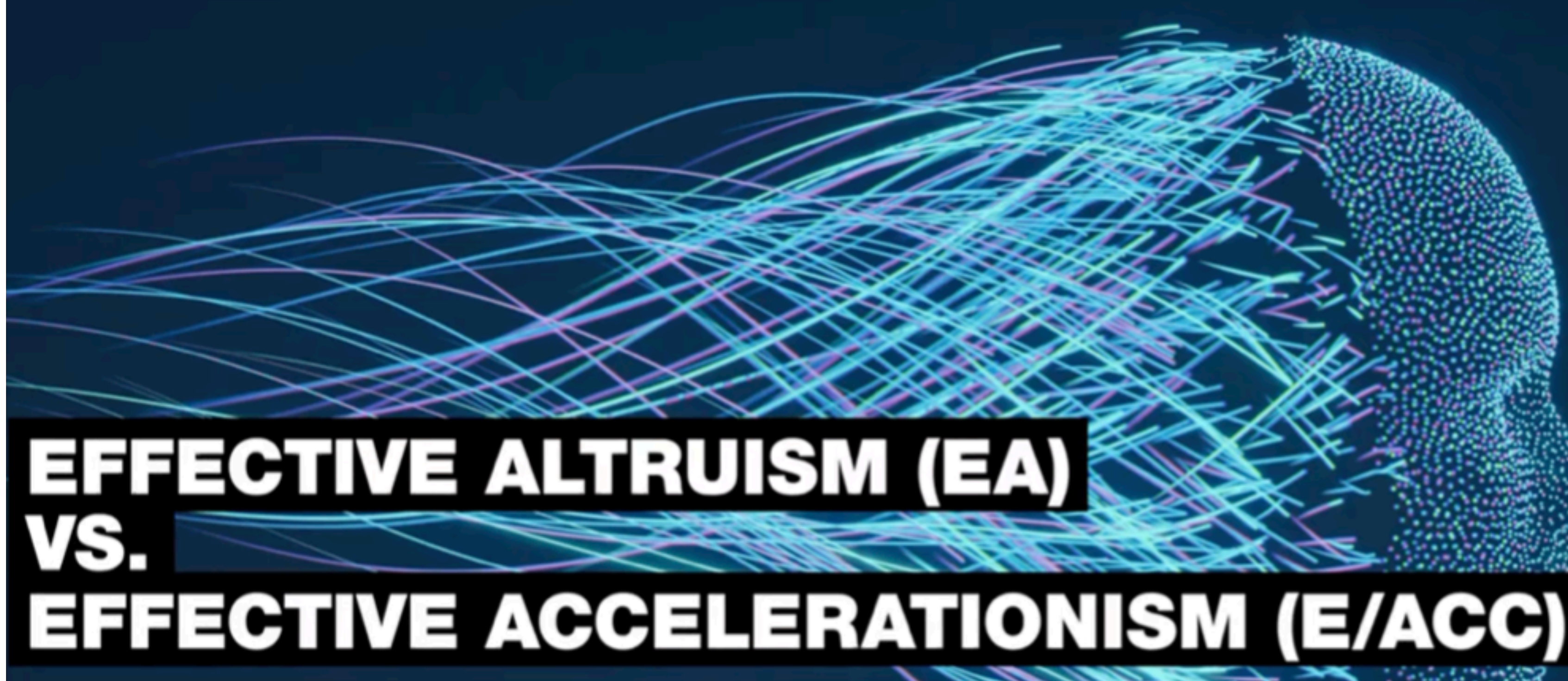**TESCREAL** (Acronym for 7 isms that dominate the AI/Technology conversation:

5. **R**ationalism: the emphasis on evidence-based science and the ethical use of technology to extend human lives.

6. **E**ffective **A**ltruism: a philosophy and social movement that advocates using evidence and reason to determine the most effective ways to benefit others.

7. **L**ongtermism: a focus on the long-term consequences of our actions and the idea of ensuring that the future goes as well as possible

Jim Fan: https://www.ted.com/talks/jim_fan_the_next_grand_challenge_for_ai

EFFECTIVE ALTRUISM (EA) VS. EFFECTIVE ACCELERATIONISM (E/ACC)

Question:
  Should we
- regulate (EA) (perhaps have a moratorium), or
- accelerate (E/ACC)

That was the the core issue of the conflict over Sam Altman's firing and rehiring in OpenAI (in a matter of a few days - the new board had to balance between moving fast or adding guardrails.

EA:

Ian Bremmer: The rapid evolution of AI outpacing government regulation is a concern for global stability and security in 2024 -

The Biden administration will use the Defense Production Act to force the tech industry to inform the Commerce Department when they start new AI projects (training "high powered AI algorithms")

EU Commission is proposing laws of regulation in 2024

UK is working on laws that prohibits AI (being *non-human*) to be patent holders of "inventions"

Arguments for restrictions

- Much like humans have wiped out many species, if we become the less intelligent beings we should expect to be wiped out
  - We wiped out the African black rhinoceros over a horn of no scientific value

- Harms are already being caused by AI.
  - AI being controlled by the tech elite could lead to more social inequality

- Even side effects could be very harmful
  - AIs may be threatened when we try to create new and smarter AI, and would have the knowledge to create and release biological agents

- AI may have many levers to work with
  - AI will be embedded in our military, police force, government, etc

- It's easy to imagine a human might use AI to wreak havoc
  - Absolute power corrupts absolutely

Arguments for restrictions

Recent examples

- Deep fake Joe Biden robocalls urging voters not to vote in the New Hampshire primary

- Deep fake porn of Taylor Swift viewed over 50 million times on first day

- George Carlin estate suing company that used his likeness and voice to produce a comedy show

Self-regulation?

- OpenAI (and Common Sense Media) **encourage the safe and responsible use of ChatGPT**.

- They want top **create family-friendly GPTs (on sale in the GPT store), GPT guidelines, and education materials**—for kids and parents.

- This comes **after OpenAI agreed to support a framework**, which includes a 'nutrition label-like' method of assessing the safety, transparency, ethical use, and impact of AI.

This collaboration comes *after* the FTC launched an investigation last year over claims ChatGPT was breaching data security laws - and after EU authorities have raised similar concerns . . .

Regulation underway

- For months, Italy's data protection authority (Garante) has **been investigating OpenAI over its compliance with GDPR laws**, temporarily suspending ChatGPT in Europe last year.

- While Garante hasn't confirmed what data protection laws ChatGPT has violated this time, the main issue is expected to be **how OpenAI processes personal data to train its models**

- OpenAI **has been given 30 days to defend itself against the allegations**

- Although the consequences of the suspected breach are unknown, confirmed security breaches in the EU **can attract up to €20 million in fines** (up to 4% of global annual turnover).

Arguments for NO restrictions

E/ACC in one sentence:

'Techno-optimist Manifesto'* gives us a glimpse into the thinking that thrives in Silicon Valley:

"AI and unrestricted expansion of new technology, can - together with setting the market forces free - save humanity."

Corporate thinking:
– we need to be 'the first' - or someone else will

*) Marc Andreassen/American entrepreneur, tech-investor, software engineer

"The real problem is not whether machines think but whether men do."
– B.F. Skinner

Food for thought:

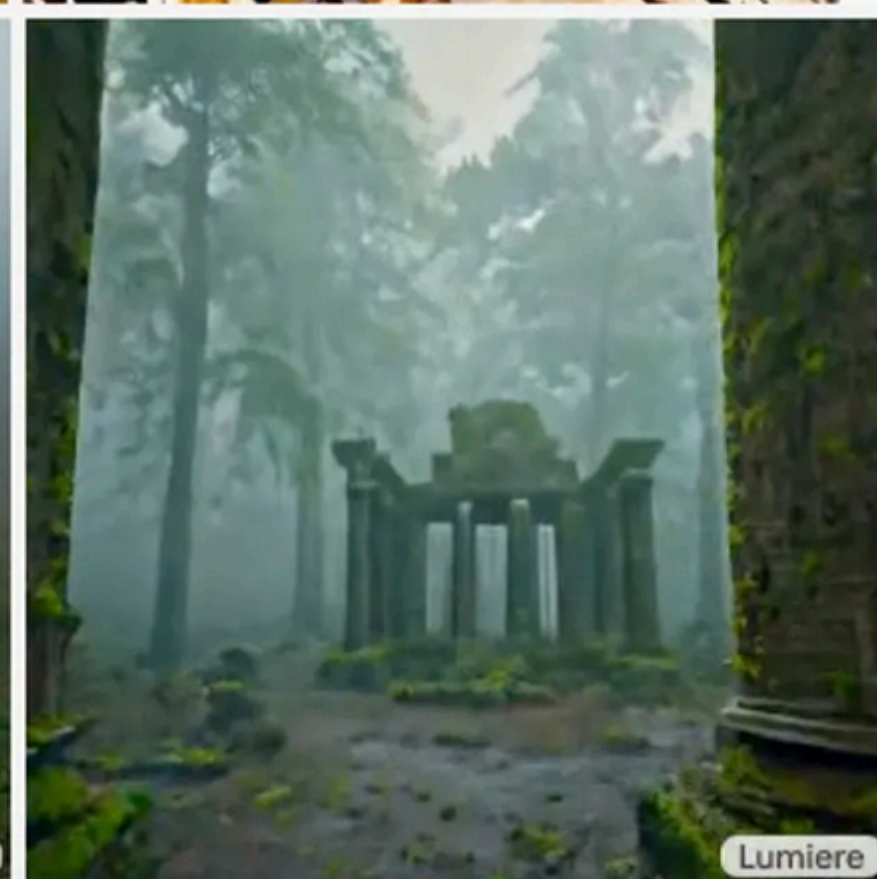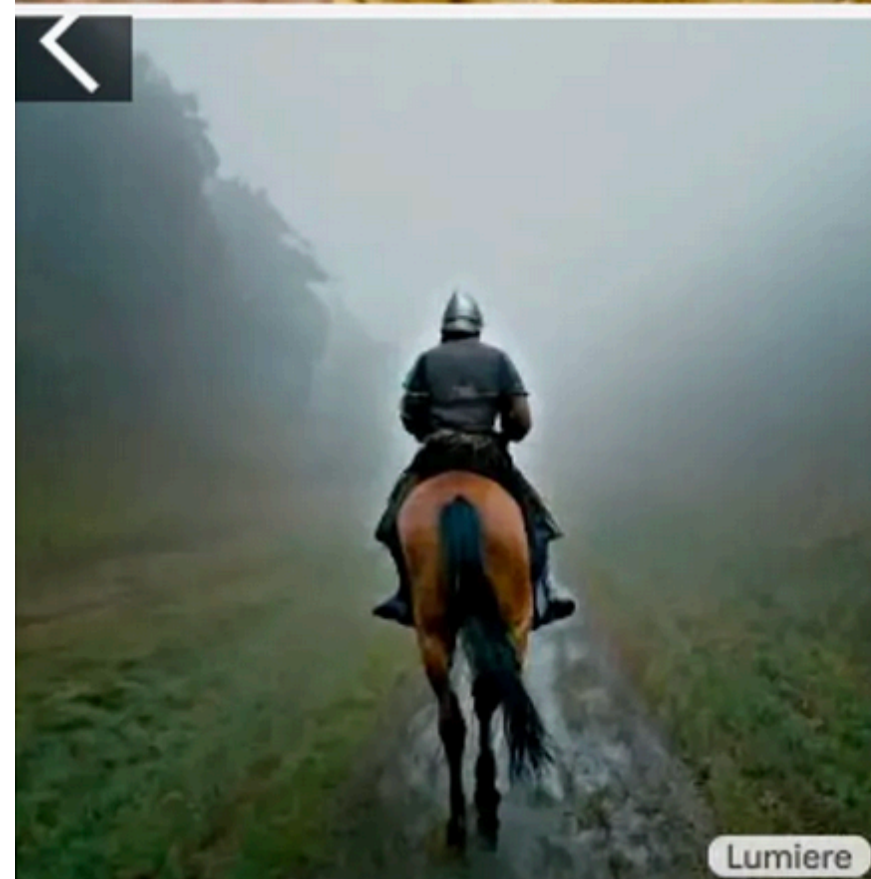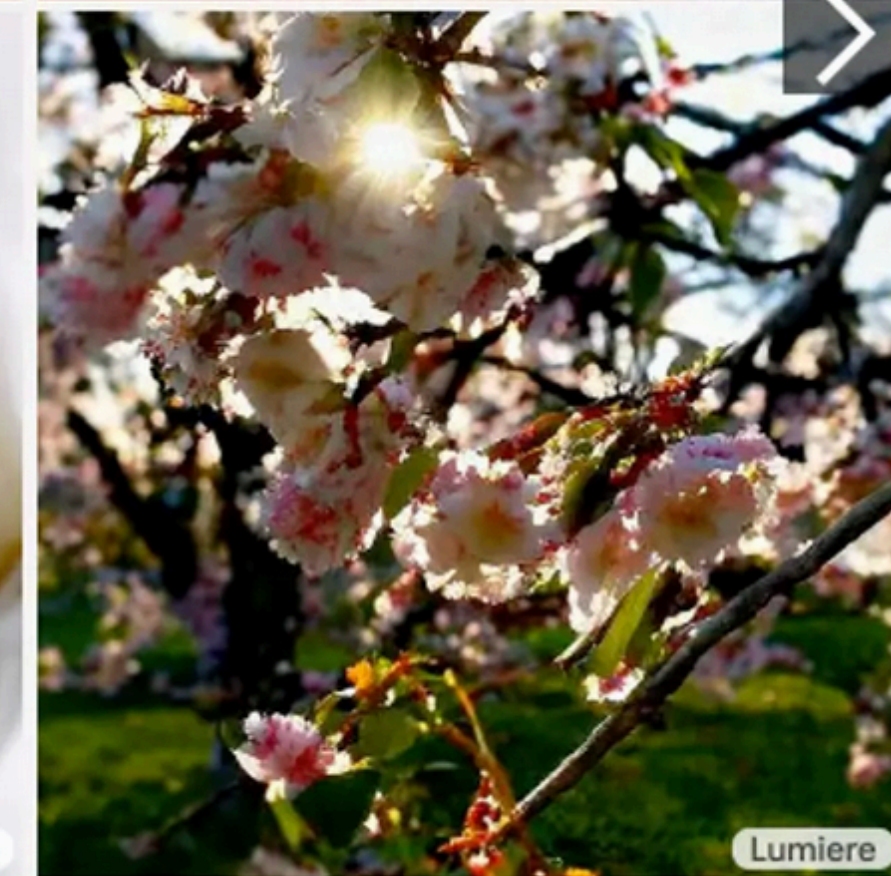– AI won't take your job . . . someone using AI will!

Food for thought:

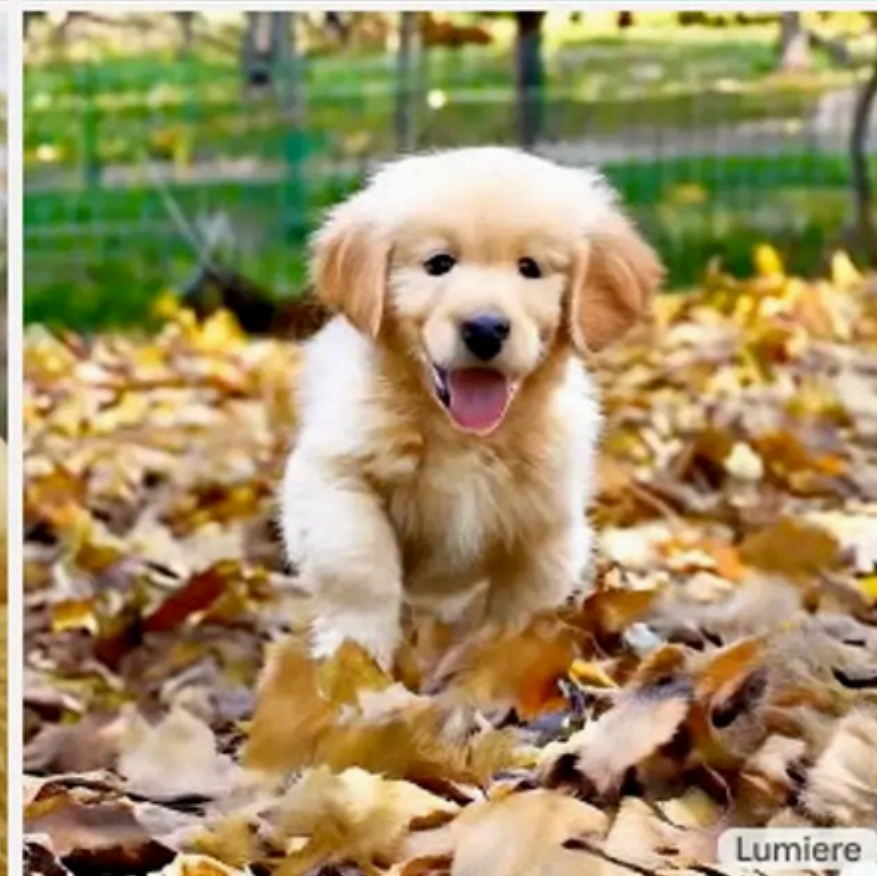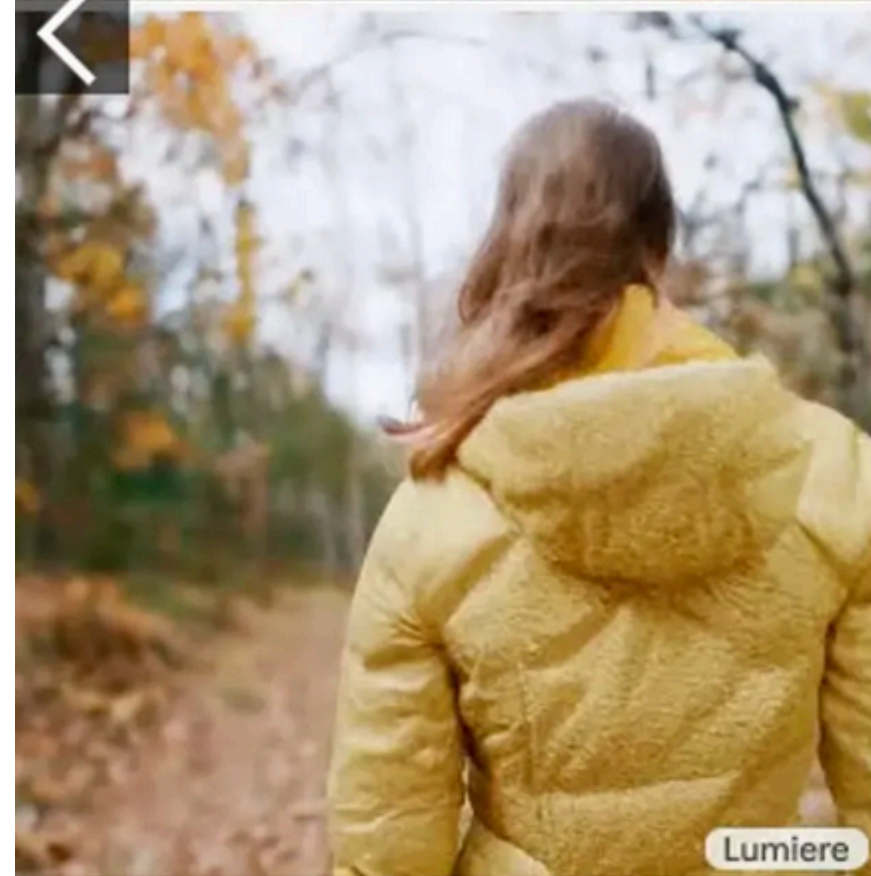- AI won't take your job . . . someone using AI will!

- The best time to level up your AI skill was yesterday;
  The second best time?
  Right now!

What Altman says about ChatGPT:

- 😮 AI systems like OpenAI's ChatGPT are known to "hallucinate," generating nonsensical outputs.
- 🧠 Large language models (LLMs) like GPT are not models of brains, but rather models of language and its patterns.
- 💡 Creative cognition involves a structured, distributed search for solutions to problems.
- 🤖 GPT can generate ideas and evaluate them, but ultimate evaluation and goal-setting still rest with the user.
- 🔄 GPT has the potential to perform adaptive, creative search, but currently falls short in grasping aesthetic goals and fostering dialogue.
- ⚠️ Aligning AI systems with the user's creative goals can be complex and potentially dangerous, as seen in the example of a recipe app suggesting dangerous concoctions.
- 🚦 "Tripping" might be a more accurate term for AI errors of judgment rather than "hallucinating."
- 💡 GPT's errors of judgment and factual inaccuracy fall into a wider problem of groundedness in real-world concerns.
- 🔄 GPT's language model is designed to make plausible predictions, not report accurate information or share the user's goals.
- 👥 Human creativity and attention to accuracy are not mutually exclusive; creative capability should not distract from expectations of accuracy.
- 🌐 GPT's capabilities are being increasingly enhanced with input and output filters, creating a complex assemblage.
- © GPT's use of copyrighted texts raises questions about fair use and plagiarism.
- 🤔 Multiple concerns arise when considering the potential of AI systems like GPT, including creativity, plagiarism, factual accuracy, safety, usability, explainability, energy efficiency, and profit.
- 💪 GPT can be a powerful generatively creative tool, but a clear-headed understanding of its grasp on reality is critical.

Creations in Google's Lumiere

The Biden administration is using the Defense Production Act to inform the Commerce Department when they start new AI projects (training high powered AI algorithms.