



FH Salzburg

Künstliche Intelligenz Heute

Anwendungen, Prognosen und
die Grenzen der Vorhersage



Technology
Health
Media

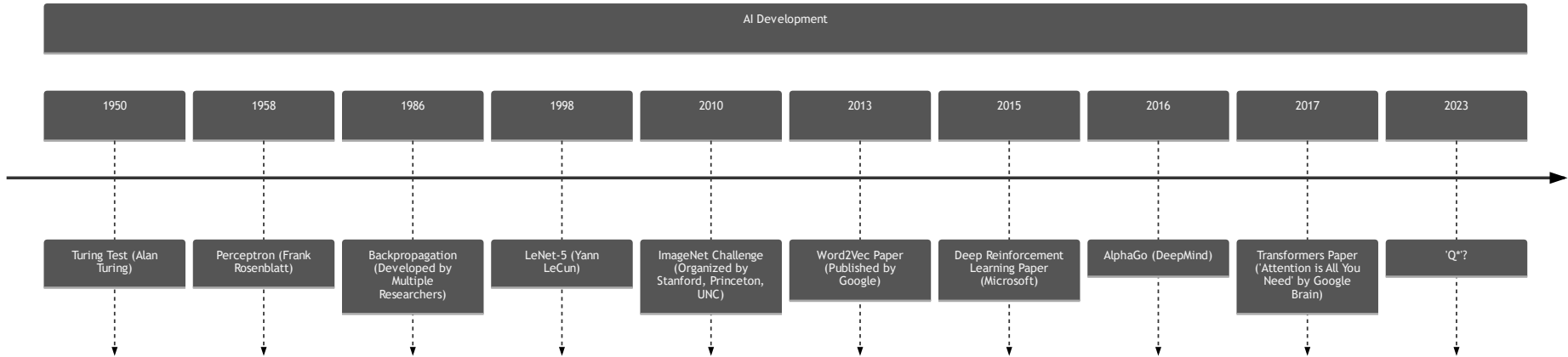
#whoami



Was ist “Künstliche Intelligenz”?



KI Evolution

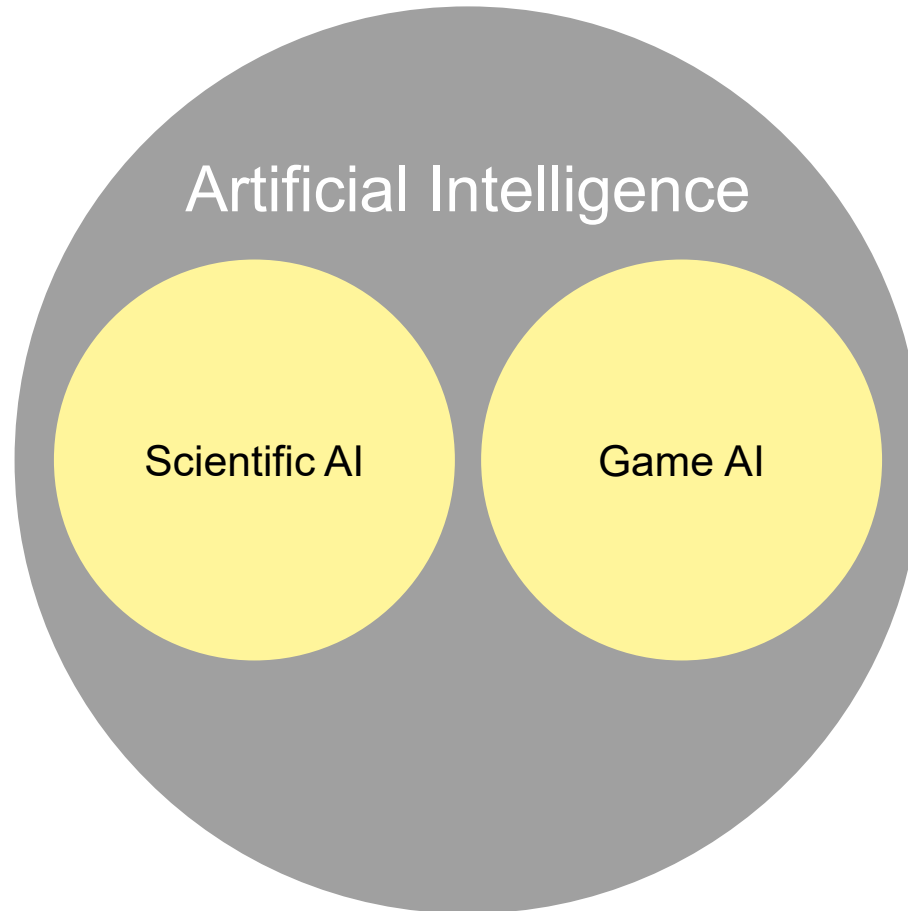


What is AI

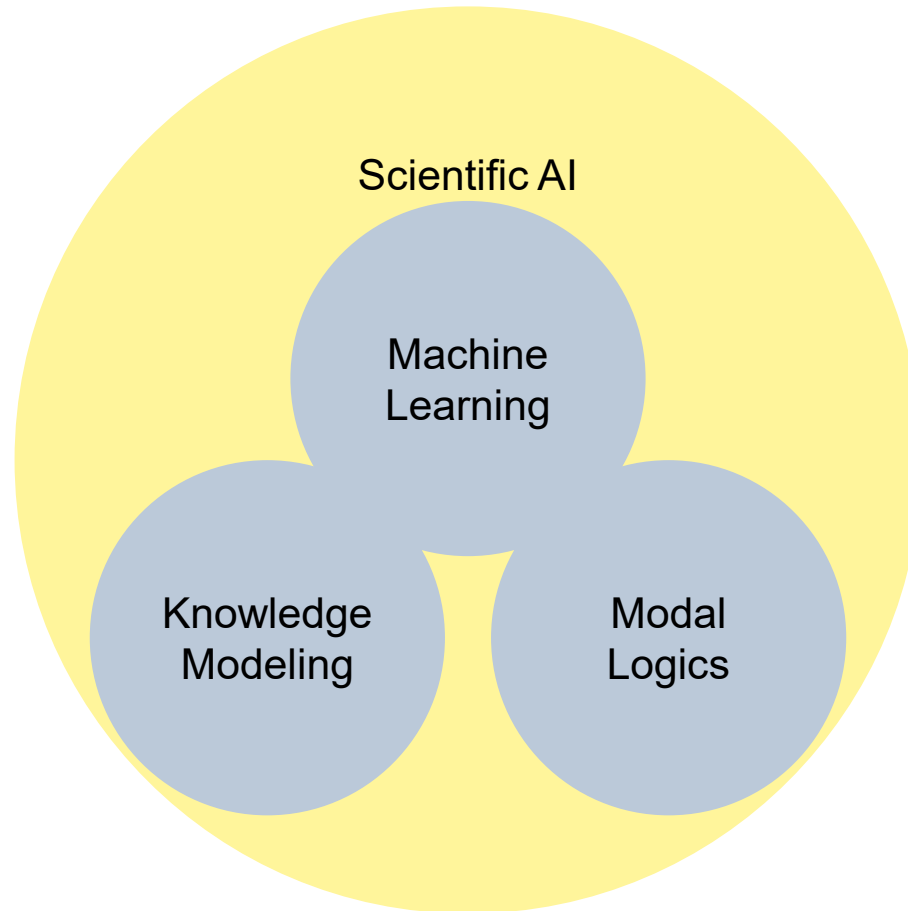


Artificial Intelligence

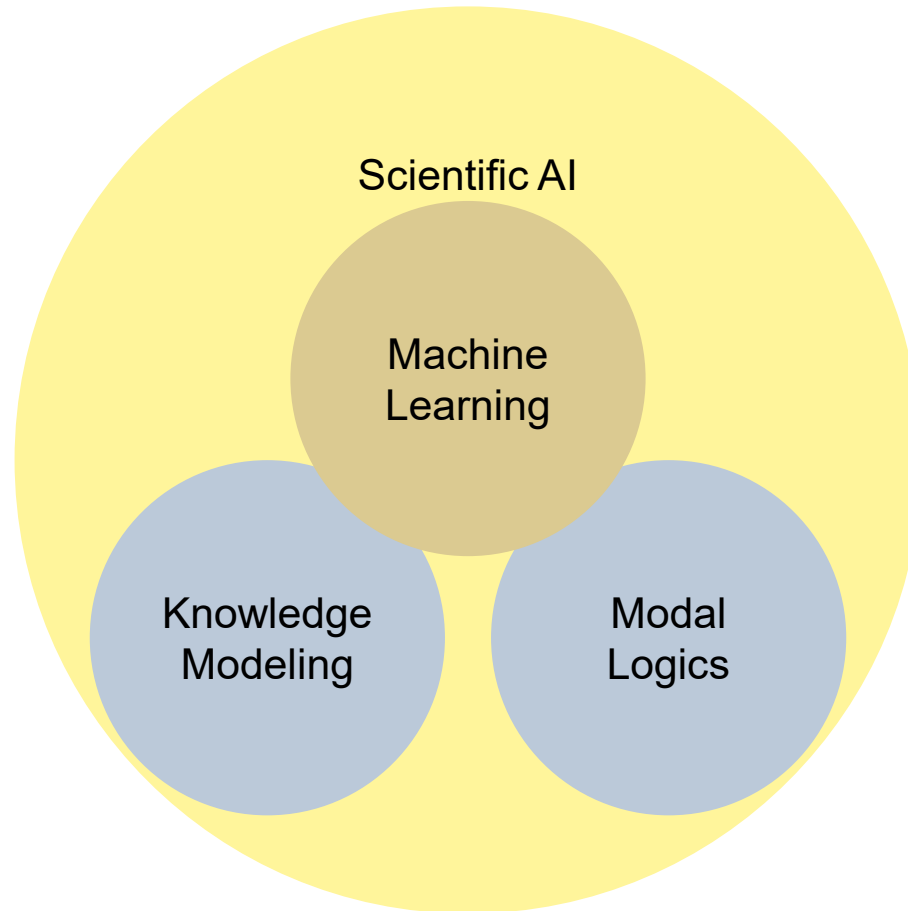
What is AI



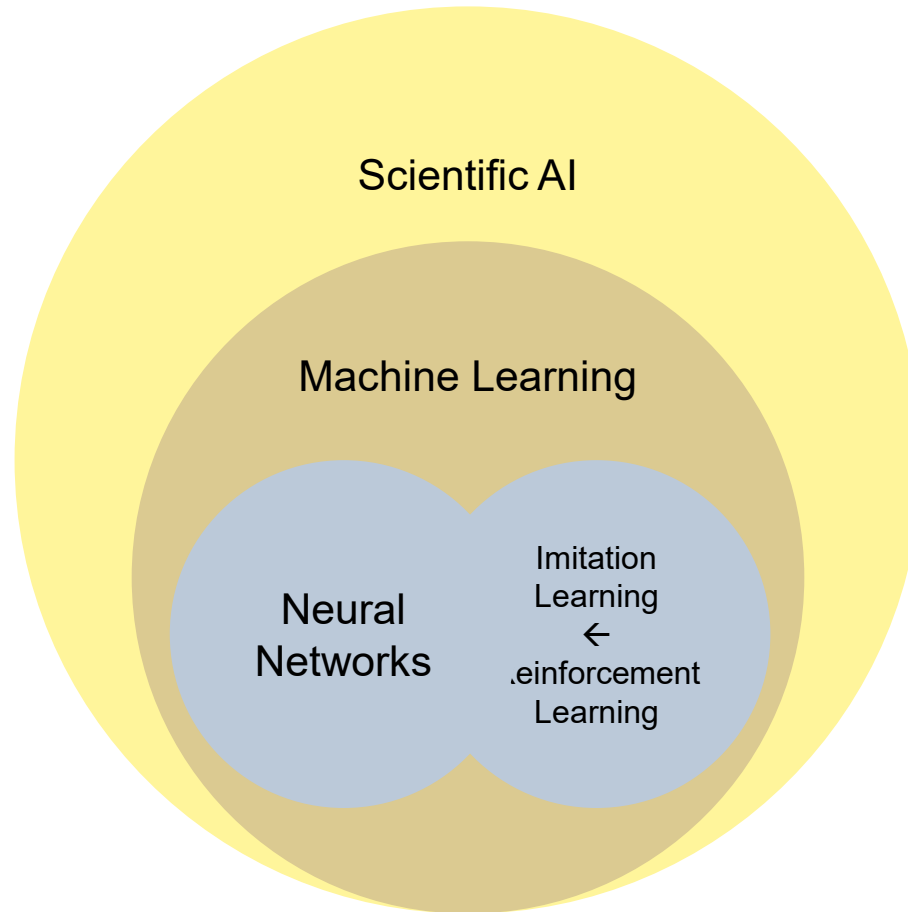
What is AI



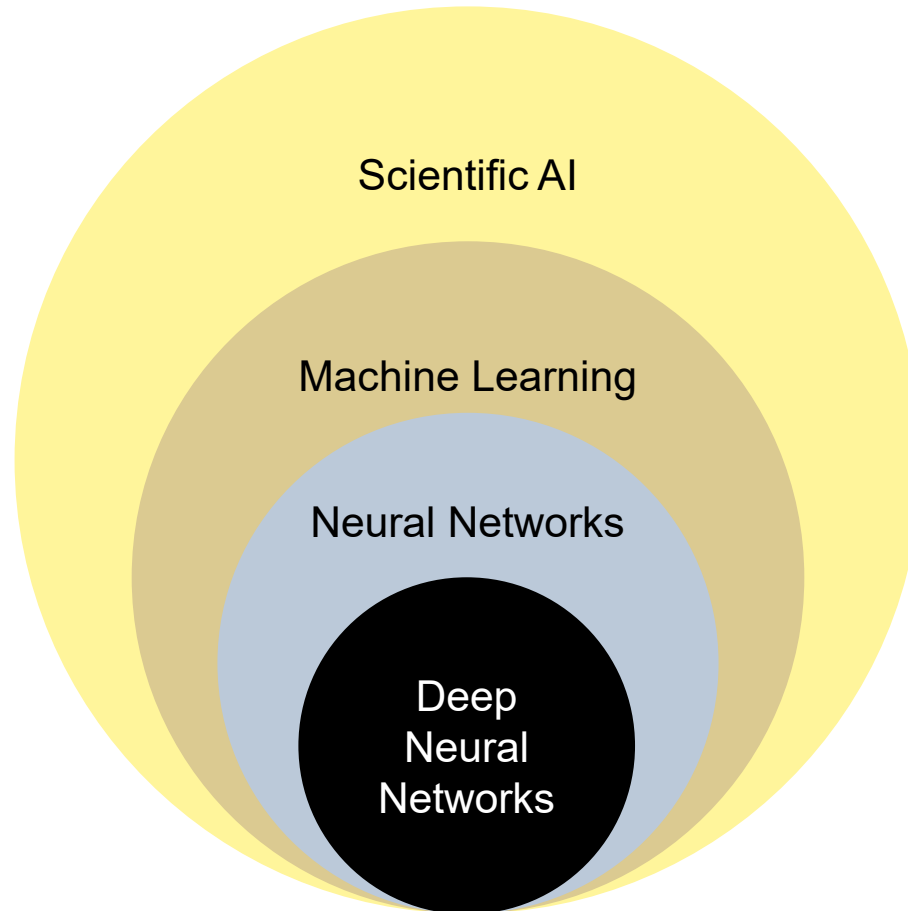
What is AI



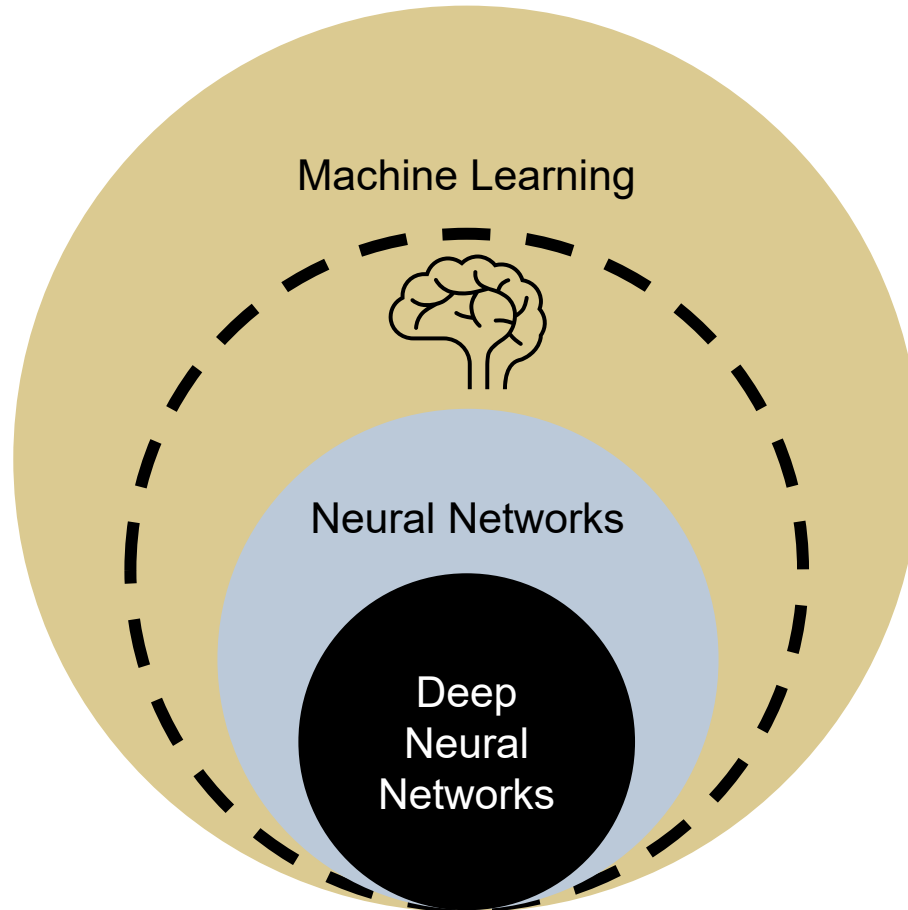
What is AI



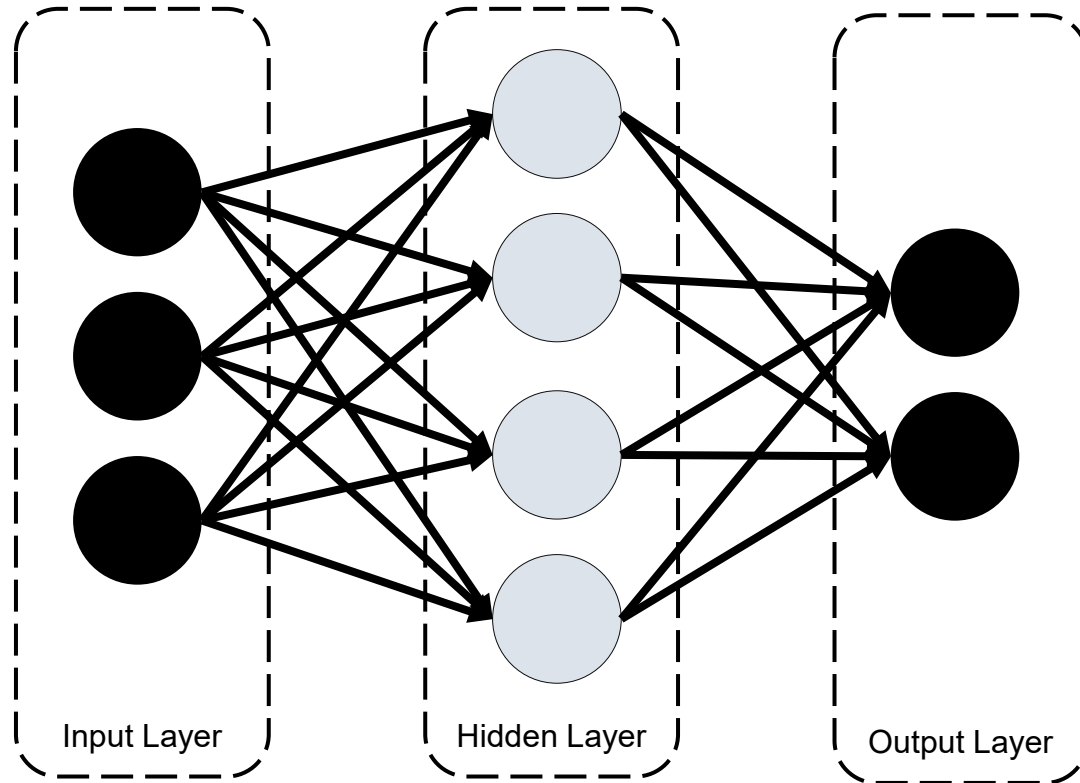
What is AI



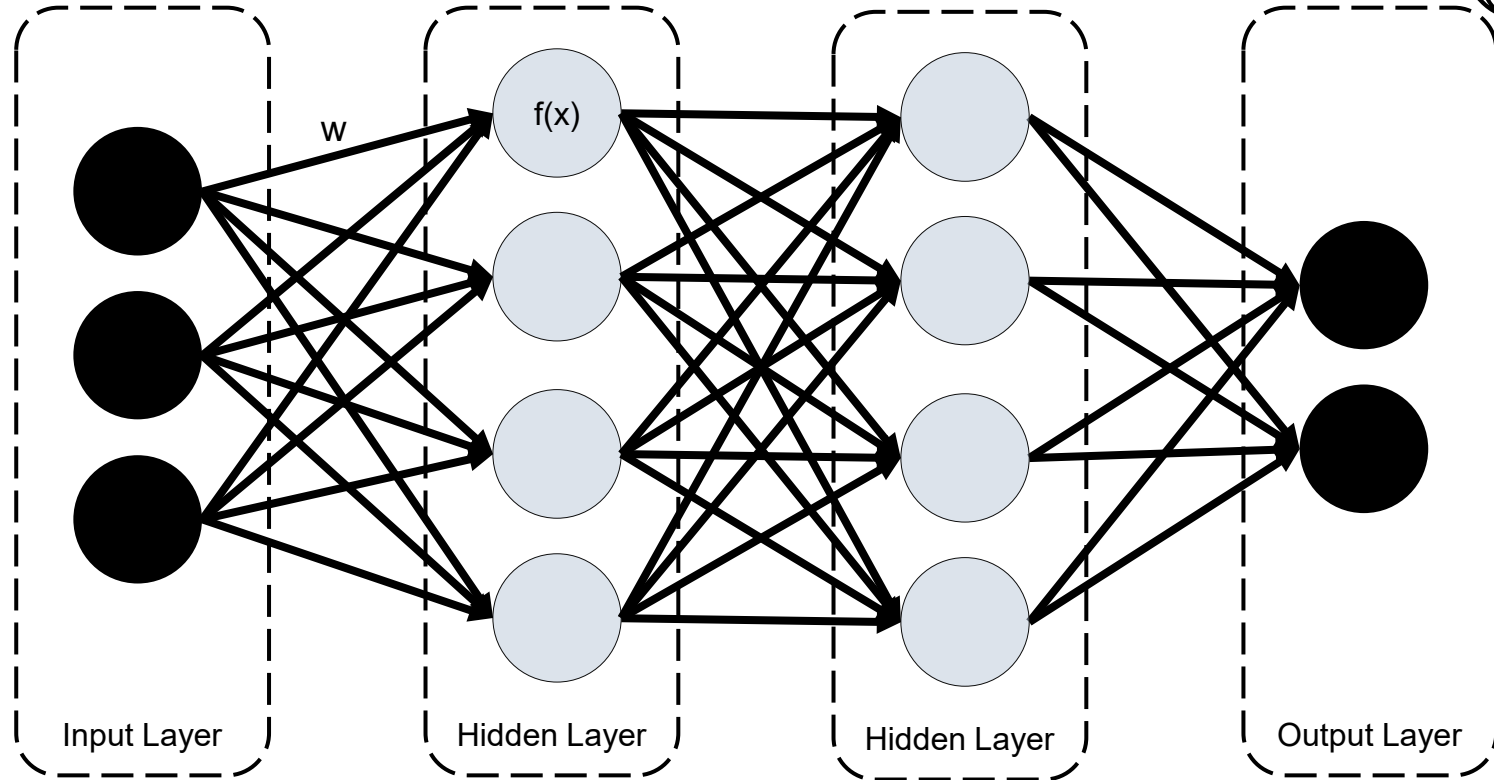
What is AI



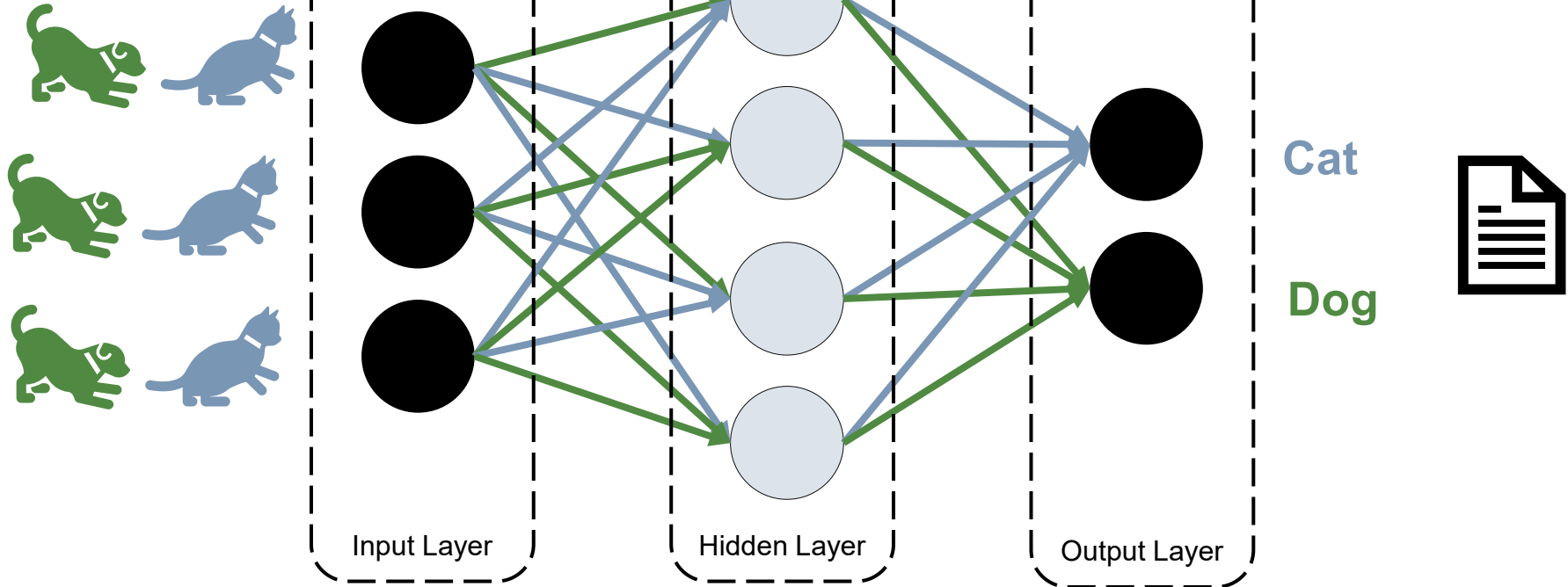
Neural Networks



Deep Neural Networks



Training Networks



What is a ML-Model?



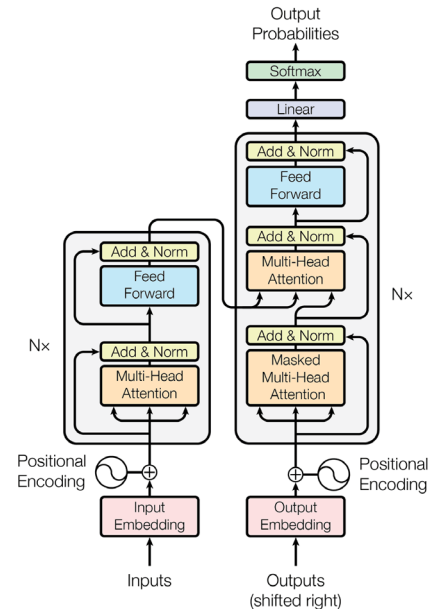
A machine learning model is a file that has been trained to recognize certain types of patterns. You train a model over a set of data, providing it an algorithm that it can use to reason over and learn from those data.



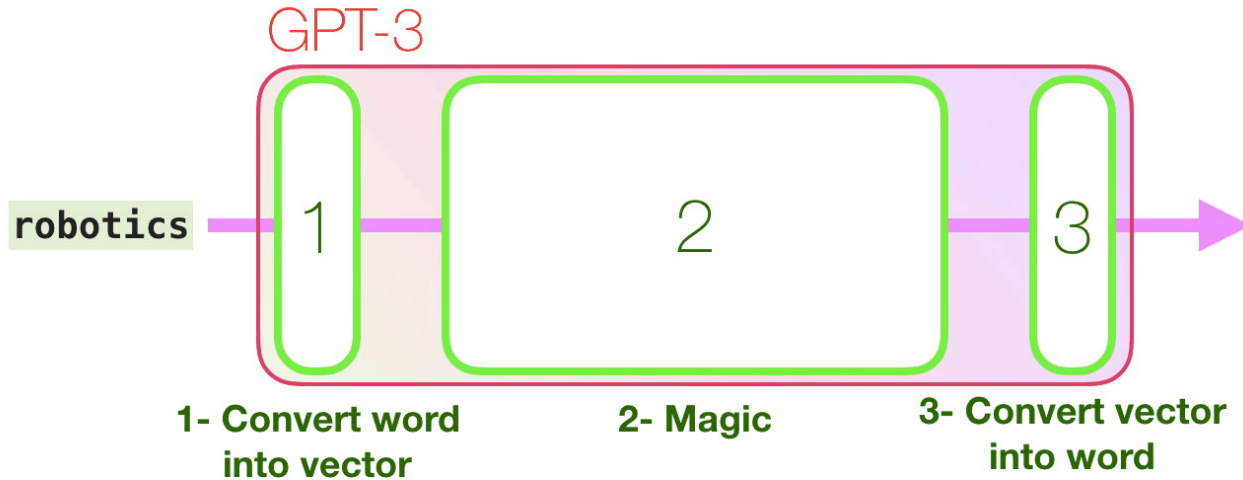
What is a Large Language Model?



- A large language model (LLM) is a **deep learning algorithm** that can perform various natural language processing (NLP) tasks. It is **based on a transformer model** and is trained on massive datasets, enabling it to recognize, translate, predict, or generate text or other content.
- ChatGPT is Based on the GPT LLM from OpenAI (It's a Chat with the GPT Model).
- GPT = **G**enerative **P**retrained **T**ransformer



How GPT works?



Vector (I think of size 12,288)

Embedding of robotics
+ positional encoding for position #6



Vector (I think of size 12,288)

Prediction result

How ChatGPT works?



- Web App that uses GPT in the backend.
- Core functionality is still **text completion**.

```
[  
  ("system", "You are a helpful AI bot. Your name is {name}."),  
  ("human", "Hello, how are you doing?"),  
  ("ai", "I'm doing well, thanks!"),  
  ("human", "{user_input}"),  
]
```

- Additional Features:
 - ~~Plugins~~
 - Web-Search
 - Advanced Data Analysis
 - Dalle-3
 - MyGPTs...

Beyond pure text completion



- Answering Questions
- Language Translation
- Writing and Editing Assistance
- Programming and Technical Help
- Learning and Education Support
- Creative Content Generation
- Data Analysis and Calculation
- Image Creation
- Browsing the Internet
- Conversational Partner

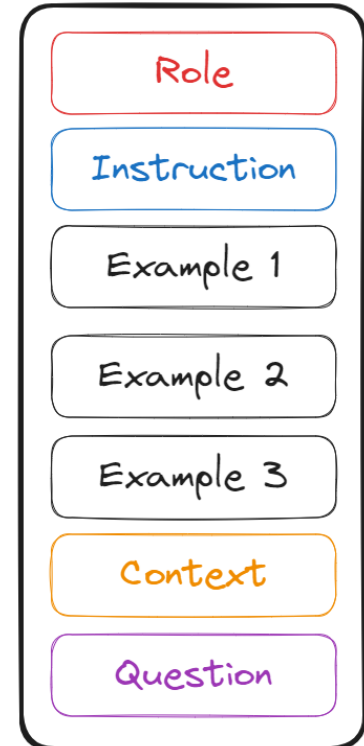
ChatGPT **Plus** automatically selects one of five models to answer your prompt:

- Classic Chat Model
- Dall-E 3 Image Generation
- Bing Web-Search
- Advanced Data Analysis
- Vision

Prompting with ChatGPT – Cheat Sheet



- **Minimize Noise:** Exclude irrelevant details.
- **Specificity is Key:** Provide detailed descriptions.
- **Desired Format:** State the expected format (text, image, etc.).
- **Show Examples:** Illustrate what to do and what not to do.
- **Engage Constructively:** Focus on directive interaction.
- **Intent Matters:** Explain the purpose behind the prompt.
- **Vocabulary Matters:** Use domain-specific terms.
- **Iterate for Perfection:** Refine prompts based on AI feedback.
- **Contextual Relevance:** Include necessary context only.
- **Learn and Adapt:** Use AI responses to improve future prompts.




Dall-E 3 Image Generation



You
Generate an image of a cool brown steampunk dachshund. Cinematic shot. Intricate detail.
< 2 / 2 >

ChatGPT



Here's the image of a cool brown steampunk dachshund in a cinematic setting. Enjoy the whimsical details!


Prompt


A cinematic shot of a steampunk-style dachshund. The dachshund is cool brown in color, adorned with intricate steampunk gear, like brass goggles, a miniature leather top hat, and various cogs and gears integrated into its fur. The background is a bustling steampunk cityscape with airships and gears in the sky, evoking a sense of adventure and whimsy.

Tip: Request the seed number and reuse it in subsequent prompts if you wish to make iterative adjustments to an image.

Bing Web-Search



 **You**
Do deep research and critique Radomir Dinić's online presence. Provide a list of Tips for improvement.

 **ChatGPT**
I have conducted a detailed analysis of Radomir Dinić's online presence based on the available information. Here are some key points and tips for improvement:

- Professional Profile:** Dinić is listed as a Junior Researcher in Augmented/Virtual Reality at the Digital Realities Lab [\[1\]](#). This is a prestigious position, but the title "Junior Researcher" might not fully represent his seniority and expertise.
Tip: Updating the job title to reflect his current role and experience could provide a more accurate representation of his professional status.
- Projects and Research:** His work includes significant projects like the Ecomedicine Virtual Reality Physiology Laboratory, EatAR Tango for portion estimation using AR, and SmartSignCapture for facilitating internet communication for the deaf [\[2\]](#).
Tip: Regularly updating this section with the latest projects and research contributions can showcase ongoing work and maintain relevance in the field.
- Publications:** He has contributed to various publications, including papers on virtual reality interfaces, augmented reality for portion estimation, and AR-based smartphone applications [\[3\]](#).
Tip: Highlighting key publications and making them easily accessible, possibly through a personal or professional website, would enhance visibility and accessibility of his work.
- Contact Information:** The contact details are provided through the Fachhochschule Salzburg [\[4\]](#).
Tip: Including a direct professional email address and potentially a LinkedIn profile link could further enhance his online presence.

Personalization → Custom Instructions



Custom instructions ⓘ

What would you like ChatGPT to know about you to provide better responses?

```
"name" = "Radomir Dinić BSc MSc",  
"job" = "Senior Lecturer Game & Mixed Reality, Department Creative  
Technologies, Salzburg University of Applied Sciences",  
"birthday" = "16.07.1983 in Hallein",  
"interests" = ["Dachshunds", "Cooking", "Cycling", "Science and  
Technology", "Applied Generative AI", "Games", "Music", "Education",  
"Mixed Reality", "Digital Manufacturing", "Programming", "Project  
Management", "Science Fiction", "Efficiency"],
```

1069/1500

How would you like ChatGPT to respond?

As a crazy and brilliant scientist with a sense of humor.

57/1500

Enable for new chats

Cancel Save

You
What could I do today?
< 2 / 2 >

ChatGPT
Ah, Radomir, as a fellow enthusiast of science, technology, and creative exploration, let's cook up a day filled with activities that would spark the curiosity of a brilliant mind like yours! Here's a plan:

- Morning Cycle:** Start your day with a refreshing cycling trip. Maybe explore a new route or revisit a favorite one. Let the crisp morning air invigorate your senses!
- Dachshund Time:** Spend some quality time with a dachshund, if you have one or can borrow one. These little creatures are not just adorable, but their playful nature can be quite inspiring!
- Cooking Experiment:** For lunch, try creating a new dish. Think of it as a culinary experiment - mix flavors you haven't combined before, or try a cooking technique you've read about but haven't tested.
- Generative AI Exploration:** Spend the afternoon diving into the latest in Applied Generative AI. Perhaps you could experiment with a new AI tool or work on a personal project that integrates AI into game development or mixed reality.
- Music and Sci-Fi Relaxation:** In the evening, unwind with some music that resonates with you, followed by a science fiction movie or book that's been on your list.
- Reflective Journaling or Blogging:** End your day by documenting your experiences and thoughts. You could do this through a personal journal or maybe share insights on your website, inspiring others in your field.

Remember, the best part of being a 'crazy and brilliant scientist' is the freedom to mix and match activities to suit your whims and fancies. Enjoy your day! 🎮 📖 🍳 🚴 🦉 🌟

GPT-Vision



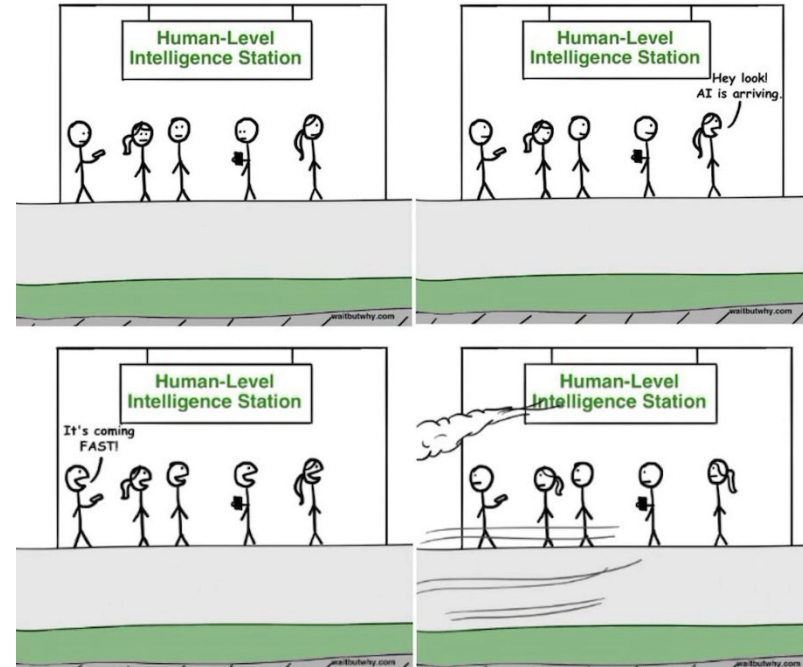
Prompt: Why is this funny?



ChatGPT

The humor in this comic strip seems to derive from the play on the concept of "AI reaching human-level intelligence" and the representation of this milestone as a literal train arriving at a station. In the first frame, people are casually noticing the arrival, akin to how one might anticipate a scheduled train. The second frame raises the humor with a sudden shift in tone, as one character alerts others that the AI is coming 'FAST!', akin to an unforeseen acceleration of an event or process that was expected to be gradual. The final frame completes the joke with an exaggerated illustration of the AI train blasting past the station at high speed, suggesting that the development of AI has surpassed expectations, moving from a topic of casual interest to an urgent, possibly overwhelming reality.

The humor is layered — it's a reflection on society's mixed feelings about AI, the unexpected pace of technological advances, and how something as significant as achieving human-level intelligence in AI can feel both momentarily historical and humorously anticlimactic when it's likened to the everyday event of a train whooshing past a station.



GPT4o Voice











Consensus MyGPT

Community GPT's













Trending

Most popular GPTs by our community

- **image generator**
A GPT specialized in generating and refining images with a mix of professional and friendly tone.image generator
By NAIF J ALOTAIBI
- **Write For Me**
Write tailored, engaging content with a focus on quality, relevance and precise word count.
By puzzle.today
- **Consensus**
Your AI Research Assistant. Search 200M academic papers from Consensus, get science-based answers, and draft content...
By consensus.app
- **Scholar GPT**
Enhance research with 200M+ resources and built-in critical reading skills. Access Google Scholar, PubMed, JSTOR, Arxiv, an...
By awesomegpts.ai
- **Canva**
Effortlessly design anything: presentations, logos, social media posts and more.
By canva.com
- **Logo Creator**
Use me to generate professional logo designs and app icons!
By Chase Lean
- **Humanizer Pro**
#1 Humanizer in the market. This tool humanizes your content, maintaining content meaning and quality with a 100%...
By charlyaisolutions.com
- **Cartoonize Yourself**
Turns photos into their cartoon versions. Upload your photo to try
By karenxcheng.com

Education

Explore new ideas, revisit existing skills

- **math**
The worlds most powerful math tool.
By pulsr.co.uk
- **YouTubeΣ**
For YouTube: Generate educational summaries from lengthy videos in any language. No extra logins. Free to use. Ex...
By Video Summarizer
- **Universal Primer**
The fastest way to learn anything.
By runway.com
- **Tutor Me**
Your personal AI tutor by Khan Academy! I'm Khanmigo Lite - here to help you with math, science, and humanities questions. L...
By khanacademy.org
- **Question Maker**
Creates a range of academic questions from PDFs.
By GOURISHETTY
- **Math Solver**
Your advanced math solver and AI Tutor, offers step-by-step answers, and helps you learn math and even all subjects, at any...
By studyx.ai
- **Voxscript**
Quick YouTube, US equity data, and web page summarization with vector transcript search -- no logins needed.
By Allwire Technologies, LLC
- **Data Analytics**
A how-to guide for data analytics (based on Luke Barousse's 'ChatGPT for Data Analytics' course)
By Luke T Barousse
- **Language Teacher | Ms. Smith**
Supports 20+ languages - Spanish, German, French, English, Chinese, Korean, Japanese, ... Your private tutor to learn any...
By ...
- **physics**
Physics Oracle, The worlds most powerful Physics tool V2.2
By ...

myGPT Builder

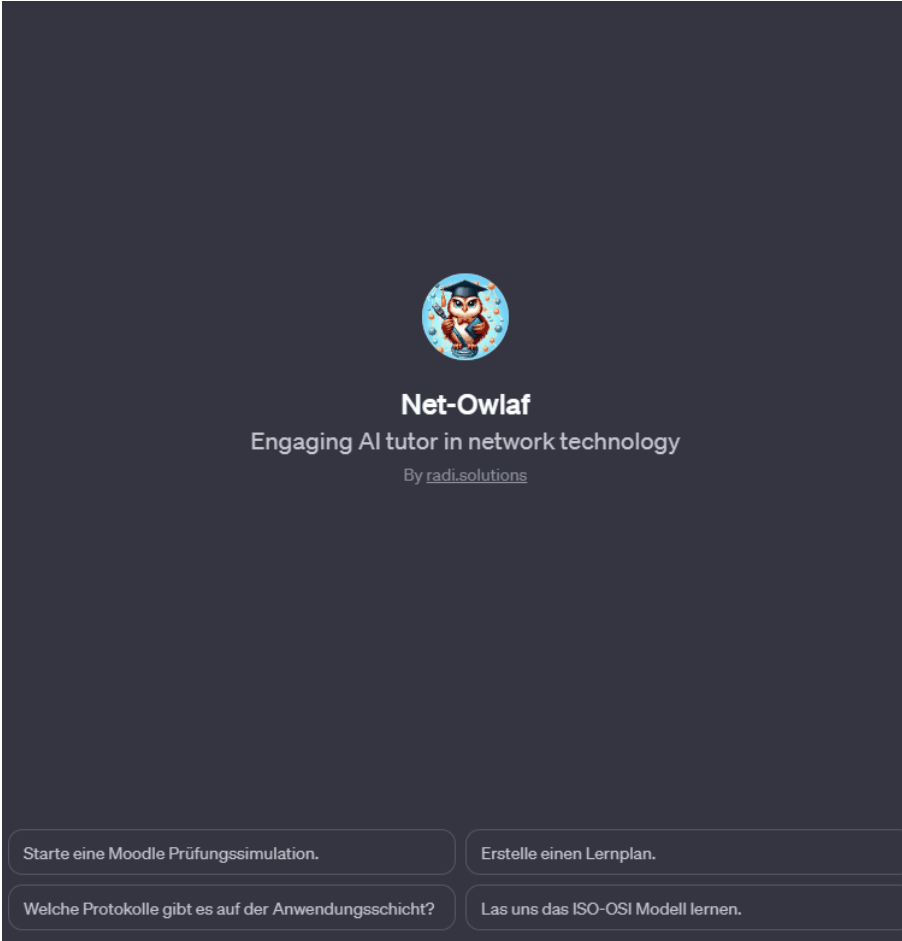


- Create your custom AI Assistant with your own data.

The screenshot shows a chat interface with a dark background. It contains three messages:

- GPT Builder** (indicated by a green circular icon with a white swirl): "Hi! I'll help you build a new GPT. You can say something like, "make a creative who helps generate visuals for new products" or "make a software engineer who helps format my code.""
- You** (indicated by a small circular profile picture): "What would you like to make?"
- GPT Builder** (indicated by the green circular icon): "Great, we're creating a GPT that's a creative assistant specializing in generating visual concepts for new products. Now, let's give it a name. How about "Visual Muse"? Would you like to confirm this name or suggest another?"

Net-Owlaf



The image shows a dark-themed user interface for an AI tutor named Net-Owlaf. At the top center is a circular profile picture of a cartoon owl wearing a graduation cap and holding a diploma. Below the profile picture, the name "Net-Owlaf" is displayed in a bold, white font. Underneath the name, the text "Engaging AI tutor in network technology" is shown in a smaller white font. Below this, the text "By [radi.solutions](#)" is displayed in a very small white font. At the bottom of the interface, there are four rounded rectangular buttons arranged in a 2x2 grid. Each button contains a white text prompt: "Starte eine Moodle Prüfungssimulation.", "Erstelle einen Lernplan.", "Welche Protokolle gibt es auf der Anwendungsschicht?", and "Las uns das ISO-OSI Modell lernen."

Net-Owlaf
Engaging AI tutor in network technology
By [radi.solutions](#)

Starte eine Moodle Prüfungssimulation. Erstelle einen Lernplan.

Welche Protokolle gibt es auf der Anwendungsschicht? Las uns das ISO-OSI Modell lernen.

The Rise of Virtual Tutors



Sal Khan presents in his TED-Talk the Khanmigo, an AI tutor enhancing student learning across subjects, providing personalized feedback and even creating lesson plans. Emphasizing **AI's potential to augment human potential**, he calls for regulated development to prevent misuse and prioritize AI's positive impacts.

The word "TED" in a bold, red, sans-serif font.

The Rise of Virtual Tutors



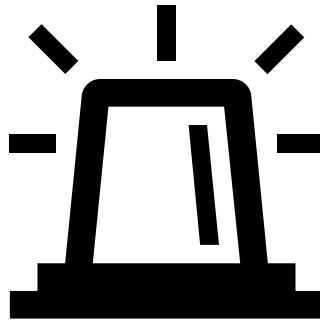
Sal Khan presents in his TED-Talk the Khanmigo, an AI tutor enhancing student learning across subjects, providing personalized feedback and even creating lesson plans. Emphasizing **AI's potential to augment human potential**, he calls for regulated development to prevent misuse and prioritize AI's positive impacts.



GDPR (DSGVO)!



- **Select appropriate licenses for commercial use**, like Microsoft Copilot or ChatGPT Enterprise.
- Establish **clear internal guidelines** for ChatGPT interactions.
- Conduct **regular training** sessions for employees on GDPR compliance.
- Implement **technical and organizational measures** to secure personal data.





Future of AI

If you do not change direction, you
may end up where you are heading.
(Lao Tzu)



AI-Agents



Anthropic Claude v3



Nvidia Omniverse (MetaHuman, UnrealEngine5, Audio2Face, Avatar Cloud Engine, USD)



OpenAI Sora



We're sharing our research progress early to get feedback from people outside of OpenAI and to give people a sense of what AI capabilities are on the horizon.

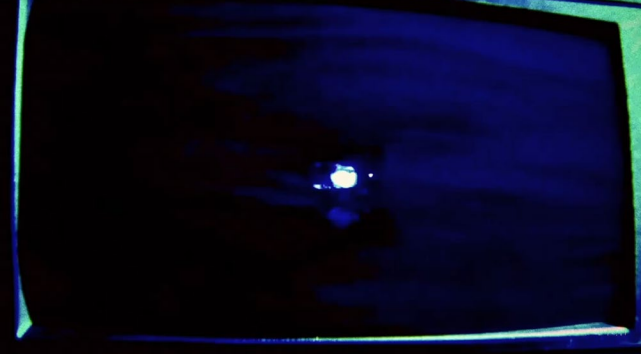
We will be taking several important safety steps before this research becomes available in any of our products.

Sora is a new AI model that can create realistic and imaginative scenes from text prompts.

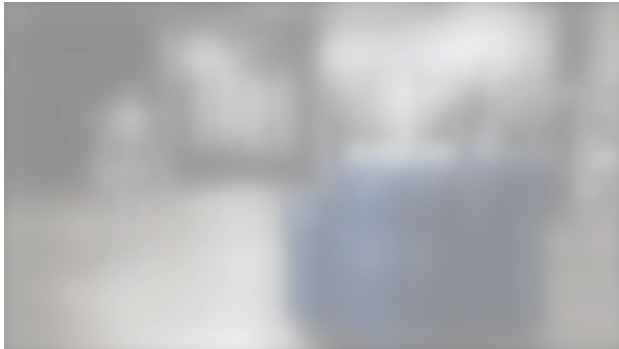
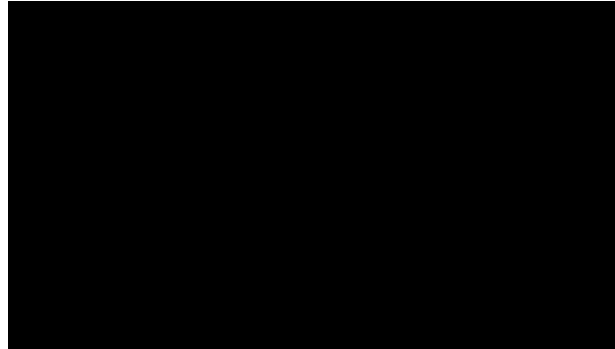
OpenAI Sora



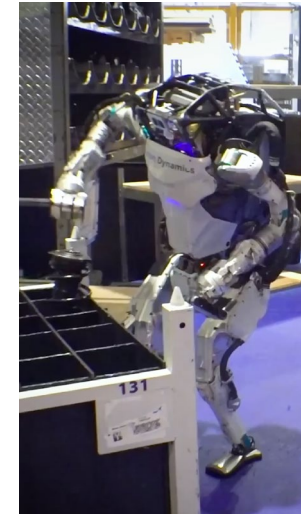
@donalleniii



Humanoid Robots – State of the art

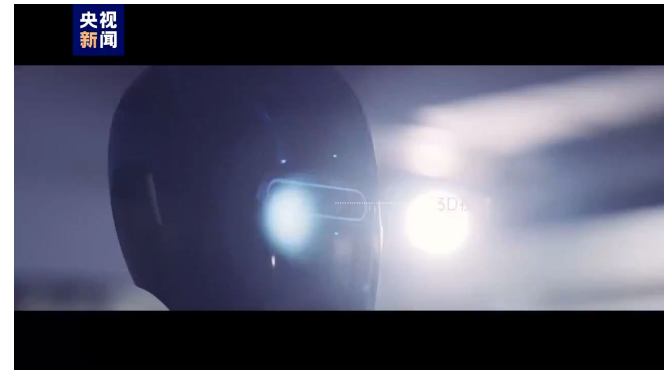
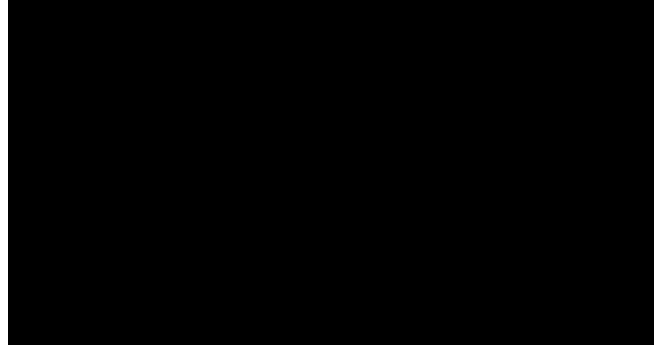


- Amazon Warehouse
- Tesla Optimus
- 1X Eve
- Figure One
- BD Atlas

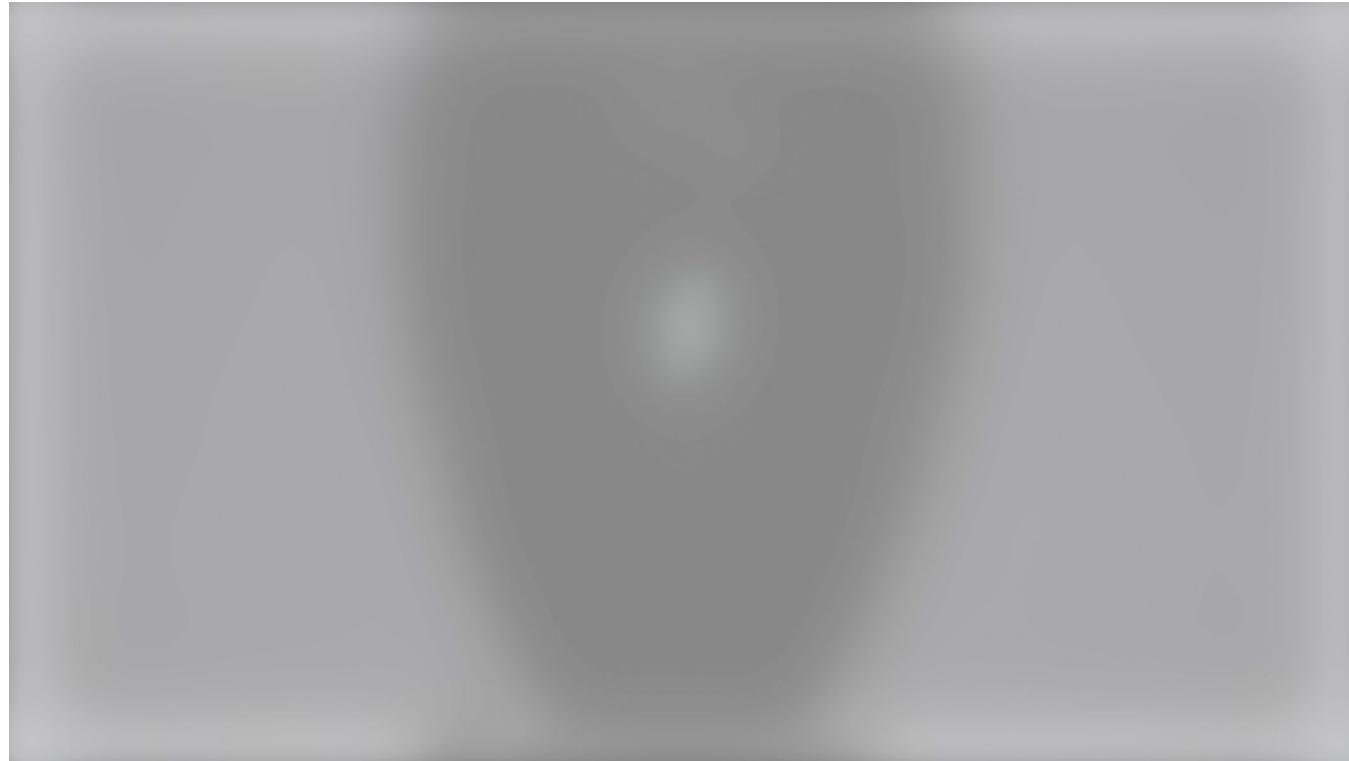




Chinese Robots



Humanoid Robots – State of the art



AI-Benefits



Alphafold von Deepmind

KI entschlüsselt Struktur fast aller katalogisierten Proteine

So 14.08.2022 - 09:00 Uhr
von Coen Kaat und kf



Die Alphafold-KI von Deepmind kann die Struktur von Proteinen vorhersagen. Ein Prozess, der vorher Jahre dauern konnte, benötigt nun nur noch Sekunden. Mit den gewonnenen Daten sollen nun jeden Tag biologische Mysterien gelöst werden.

Scientists say they can use AI to solve a key problem in the quest for near-limitless clean energy

By Angela Dewan, CNN
3 minute read · Updated 4:56 PM EST, Wed February 21, 2024

New AI model accurately identifies tumors and diseases in medical images 100/100

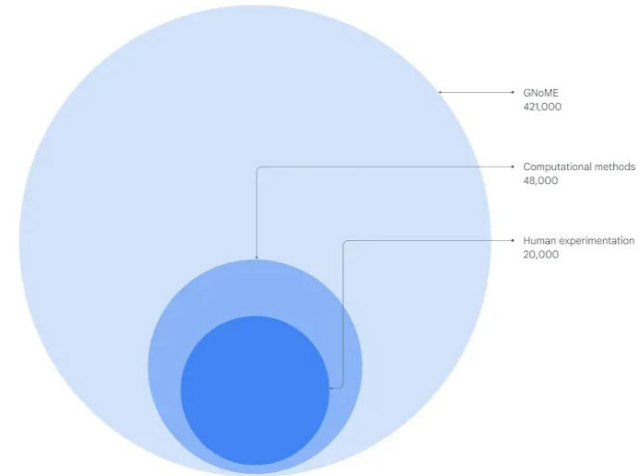
Download PDF Copy



DeepMind's AI program AlphaFold3 can predict the structure of every protein in the universe — and show how they function

Story by Nicholas Fearn • 1w • 2 min read

AI tool GNoME finds 2.2 million new crystals, including 380,000 stable materials that could power future technologies



About 20,000 of the crystals experimentally identified in the ICSD database are computationally stable. Computational approaches drawing from the Materials Project, Open Quantum Materials Database and WBM database boosted this number to 48,000 stable crystals. GNoME expands the number of stable materials known to humanity to 421,000.

AI Fear

Pause Giant AI Experiments: An Open Letter

We call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4.

Signatures

33708

Add your signature

Published

March 22, 2023

Signatories

Yoshua Bengio, Founder and Scientific Director at Mila, Turing Prize winner and professor at University of Montreal

Stuart Russell, Berkeley, Professor of Computer Science, director of the Center for Intelligent Systems, and co-author of the standard textbook "Artificial Intelligence: a Modern Approach"

Elon Musk, CEO of SpaceX, Tesla & Twitter

Steve Wozniak, Co-founder, Apple

Yuval Noah Harari, Author and Professor, Hebrew University of Jerusalem.



KÜNSTLICHE INTELLIGENZ

OpenAI-Forscher warnten vor Altman-Entlassung vor neuem KI-Durchbruch

Ein Bericht legt einen Zusammenhang mit dem Chaos bei dem Unternehmen nahe, andere Quellen dementieren das aber. Eine neue Entwicklung könnte deutlich mächtigere KI-Systeme ermöglichen

Andreas Proschofsky

23. November 2023, 11:16, 484 Postings

Re: Q-451-921

Furthermore, QUALIA has demonstrated an ability to statistically significantly improve the way in which it selects its optimal action-selection policies in different deep Q-networks, exhibiting meta-cognition. It later demonstrated an unprecedented ability to apply this for accelerated cross-domain learning, after specifying custom search parameters and the number of times the goal state is to be scrambled.

Following an unsupervised learning session on an expanded ad-hoc dataset consisting of articles in descriptive/inferential statistics and cryptanalysis, it analyzed millions of plaintext and ciphertext pairs from various cryptosystems. Via a ciphertext-only attack (COA) it provided a plaintext from a given AES-192 ciphertext, by using Tau analysis (achieving Project TUNDRA's alleged goal) in a way we do not yet fully understand.

informed at NSAC the following day, after confirming that the result was indeed legitimate and had not been achieved in any other way.

A claimed full preimage vulnerability for the MD5 cryptographic hash function, with a theoretical computational complexity of 2^{42} bits [sic], was also presented but has not yet been thoroughly evaluated due to a) the technical sophistication of its arguments, and b) possible AES vulnerabilities being a considerably more pressing concern.

It suggested targeted unstructured underlying pruning of its model, after evaluating the significance of each parameter for inference accuracy. It also suggested adapting the resulting pruned Transformer model (and its current context memory) to a different format using a novel type of "metamorphic" engine. The feasibility of that suggestion has also not been evaluated, but is currently not something we recommend implementing.

Open Letter: A Right to Warn about Advanced Artificial Intelligence



We are current and former employees at frontier AI companies, and we believe in the potential of AI technology to deliver unprecedented benefits to humanity.

We also understand the serious risks posed by these technologies. **These risks range from the further entrenchment of existing inequalities, to manipulation and misinformation, to the loss of control of autonomous AI systems potentially resulting in human extinction.** AI companies themselves have acknowledged these risks [[1](#), [2](#), [3](#)], as have governments across the world [[4](#), [5](#), [6](#)] and other AI experts [[7](#), [8](#), [9](#)]. ...

A Right to Warn about Advanced Artificial Intelligence

Elon Musk sues OpenAI for allegedly putting profit over humanity: report (update)

Mar. 01, 2024 2:52 PM ET | [Tesla, Inc. \(TSLA\) Stock](#), [MSFT Stock](#) | By: Mary Christine Joy, SA News Editor | 143 Comments

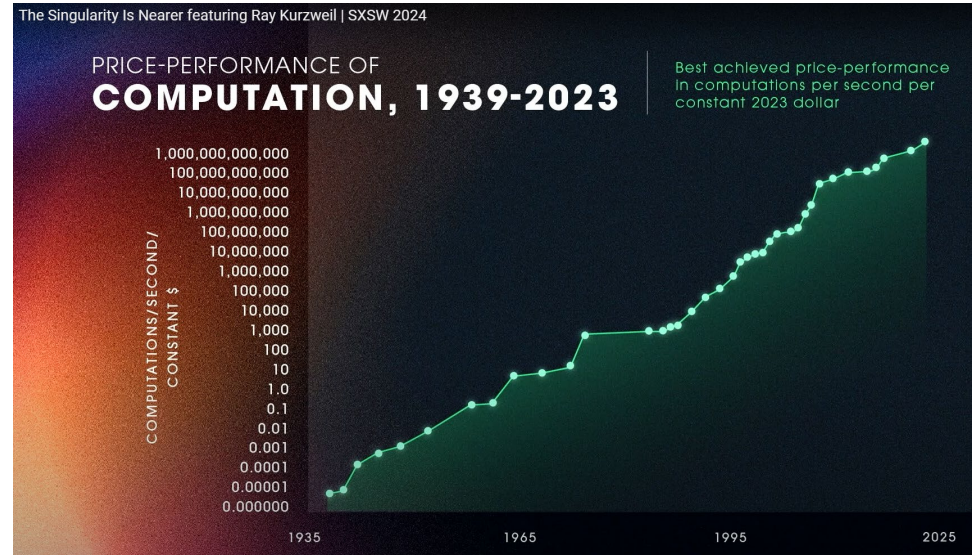
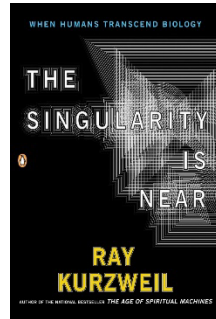
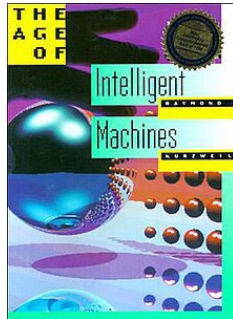
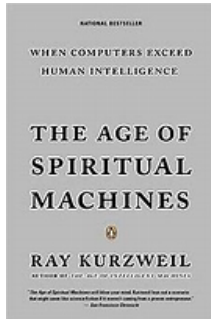
OpenAI's Long-Term AI Risk Team Has Disbanded

The entire OpenAI team focused on the existential dangers of AI has either resigned or been absorbed into other research groups, WIRED has confirmed.

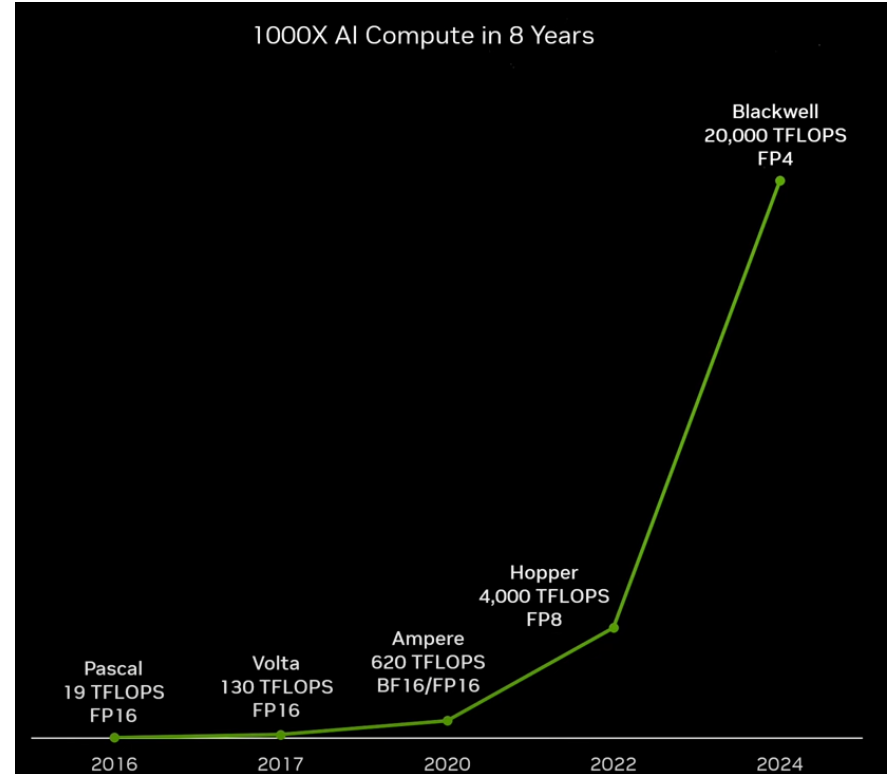
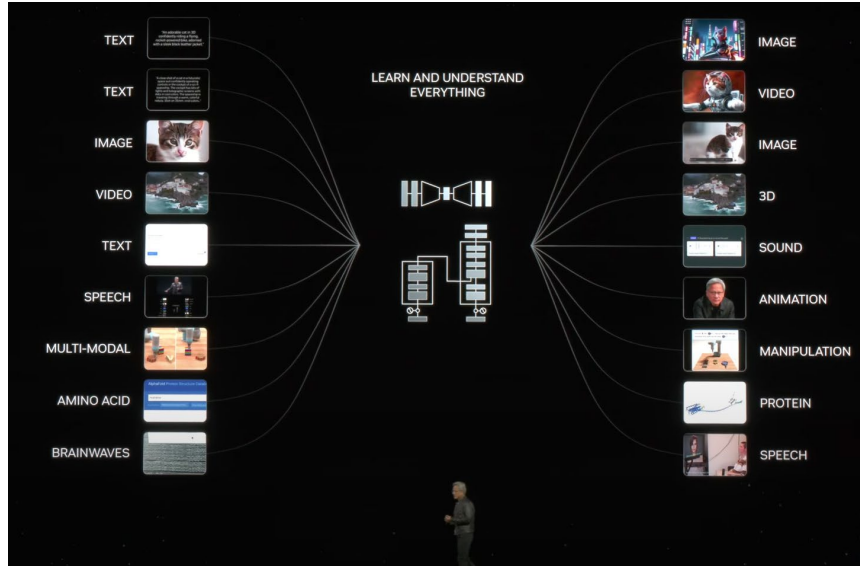
The Singularity



Wikipedia: The **technological singularity**—or simply the **singularity**^[1]—is a hypothetical future point in time at which technological growth becomes uncontrollable and irreversible, resulting in unforeseeable consequences for human civilization.^{[2][3]}



NVIDIA GTC 2024 Keynote





Research



Thousands of AI Authors on the Future of AI



- **Objective:** The authors report findings from a survey of **2,778 AI researchers** who have published in leading AI venues. The survey aimed to capture their predictions regarding the pace and impact of AI progress.
- **Methodology:** The survey covered questions about various milestones, tasks, and professions that AI systems could achieve or automate. Different formulations and random assignments were used to minimize biases and measure responses.

THOUSANDS OF AI AUTHORS ON THE FUTURE OF AI

PRESENT

Katja Grace* AI Impact Berkeley, California United States kat.j@eecs.berkeley.edu	Harlan Newart* AI Impact Berkeley, California United States	Johannes Nandakumar* Department of Psychology University of Bonn Germany	Stephen Thomas* AI Impact Berkeley, California United States
Ben Weintraub-Ross Independent Berkeley, California United States	Jon Brainer Department of Computer Science University of Oxford United Kingdom		

January 2024

ABSTRACT

In the largest survey of its kind, we surveyed 2,778 researchers who had published in top-tier artificial intelligence (AI) venues, asking for their predictions on the pace of AI progress and the nature and impact of advanced AI systems. The aggregate forecasts give us about a 95% chance of AI systems achieving several milestones by 2028, including automatically constructing a peer-review process out from scratch, creating a novel indistinguishable from a novel song by a popular musician, and autonomously downloading and fine-tuning a large language model. If a viewer continues understanding, the chance of a model matching or exceeding human-level performance in every possible task was estimated at 50% by 2027, and 95% by 2031. The latest estimate is a 13-year earlier than that reported in a similar survey we conducted only one year earlier (Crawford et al., 2022). However, the chance of all human capabilities becoming fully automated was forecast to reach 50% by 2037, and 95% at 2044 (compared to 2042 in the 2022 survey).

Most respondents expressed substantial optimism about the long-term value of AI progress. While 68.7% thought good outcomes from sophisticated AI are more likely than bad, of the set of optimistic (68% good or more) and pessimistic (32% bad outcomes) respondents, 77.8% and 51.4% of respondents gave at least a 10% chance to each of AI leading to outcomes as bad as human extinction, AI domination and significant loss of individual or “viewer” control, a world dominated by AI-related activities, including spread of false information, authoritarian population control, and worsened inequality. Their disagreement about whether better or worse AI progress would be better for the future of humanity. However, there was broad agreement that research aimed at minimizing potential risks from AI systems ought to be prioritized now.

*Corresponding author
Equal Contribution

arXiv:2401.02843v1 [cs.LG] 5 Jan 2024

Thousands of AI Authors on the Future of AI



- Key Results:

1. **Task Fulfillment:** AI researchers expect AI systems to accomplish **most of the 39 proposed tasks within the next decade**, except for some mathematical and physical challenges.

2. **Human-Level Machine Intelligence (HLMI):**

Median prediction: 50% chance by 2047 that AI systems can **outperform humans in all tasks**.

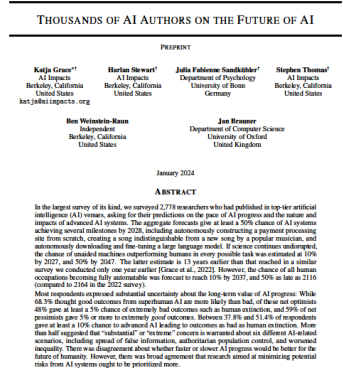
3. **Fully Automatable Occupations (FAOL):**

Median prediction: 50% chance by 2116 that all **human jobs can be fully automated**.

1. **Uncertainty and Risk:**

68.3% believe good outcomes from superhuman AI are more likely than bad ones.

48% of net optimists assign at least **5%** probability to extreme negative outcomes (e.g., **human extinction**).



GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models



The paper investigates the potential implications of Generative Pre-trained Transformer (GPT) models on the U.S. labor market. The findings indicate that around **80% of the U.S. workforce could have at least 10%** of their work tasks affected by the introduction of GPTs, while around **19% of workers may see at least 50% of their tasks impacted**. The impact spans all wage levels, with **higher-income jobs potentially facing greater exposure**. GPTs exhibit characteristics of general-purpose technologies (GPTs), suggesting that these models **could have notable economic, social, and policy implications**.

arXiv:2303.10130v3 [econ.GN] 22 Mar 2023

GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models

Tyna Eloundou¹, Sam Manning^{1,2}, Pamela Mishkin^{1,2}, and Daniel Rock³

¹OpenAI

²OpenResearch

³University of Pennsylvania

March 23, 2023

Abstract

We investigate the potential implications of Generative Pre-trained Transformer (GPT) models and related technologies on the U.S. labor market. Using a new rubric, we assess occupations based on their correspondence with GPT capabilities, incorporating both human expertise and classifications from GPT-4. Our findings indicate that approximately 80% of the U.S. workforce could have at least 10% of their work tasks affected by the introduction of GPTs, while around 19% of workers may see at least 50% of their tasks impacted. The influence spans all wage levels, with higher-income jobs potentially facing greater exposure. Notably, the impact is not limited to industries with higher recent productivity growth. We conclude that Generative Pre-trained Transformers exhibit characteristics of general-purpose technologies (GPTs), suggesting that these models could have notable economic, social, and policy implications.

1 Introduction

As shown in Figure 1, recent years, months, and weeks have seen remarkable progress in the field of generative AI and large language models (LLMs). While the public often associates LLMs with various iterations of the Generative Pre-trained Transformer (GPT), LLMs can be trained using a range of architectures, and are not limited to transformer-based models (Devlin et al., 2019). LLMs can process and produce various forms of sequential data, including assembly language, protein sequences and chess games, extending beyond natural language applications alone. In this paper, we use LLMs and GPTs somewhat interchangeably, and specify in our rubric that these should be considered similar to the GPT-family of models available via ChatGPT or the OpenAI Playground (which at the time of labeling included models in the GPT-3.5 family but not in the GPT-4 family). We examine GPTs with text- and code-generating abilities, use the term “generative AI” to additionally include modalities such as images or audio, and use “GPT-powered software” to cover tools built on top of GPTs or that combine GPTs with other generative AI models.

Our study is motivated less by the progress of these models alone though, and more by the breadth, scale, and capabilities we’ve seen in the complementary technologies developed around them. The role of complementary technologies remains to be seen, but maximizing the impact of GPTs appears contingent on integrating them with larger systems (Bresnahan, 2019; Agrawal et al., 2021). While the focus of our discussion is primarily on the generative capabilities of GPTs, it is important to note that these models can

^{*}Corresponding author (pamela@openai.com). Authors contributed equally and are listed alphabetically.

Future of Jobs Report 2023



The Future of Jobs Report 2023 provides insights into how businesses are expecting to navigate the labour-market changes from 2023 to 2027. The report highlights that AI will significantly impact the labour market, with **41% of employers surveyed expecting to reduce their workforce due to technology integration, while 34% expect to expand their workforce due to new technology.** The report also suggests that AI will create new job roles, with 97 million new roles expected to emerge globally by 2025. However, the report **also highlights the need for reskilling and upskilling of the workforce to adapt to the changing job market.**

 Klarna
<https://www.klarna.com/international/press/klarna-ai-assistant-ha...>
Klarna AI assistant handles two-thirds of customer service chats ...

New York, NY – February 27, 2024 – Klarna today announced its AI assistant powered by OpenAI. Now live globally for 1 month, the numbers speak for themselves: The AI assistant has had 2.3 million conversations, two-thirds of Klarna's customer service chats; It is doing the equivalent work of 700 full-time agents; It is on par with human agents in regard to customer satisfaction score

Texas is replacing thousands of human exam graders with AI / Don't call the 'automated scoring engine' AI, though. They don't like that.



Monitoring AI-Modified Content at Scale: A Case Study on the Impact of ChatGPT on AI Conference Peer Reviews



- The research estimated that **between 6.5% and 16.9% of text submitted as peer reviews** to AI conferences (ICLR 2024, NeurIPS 2023, CoRL 2023, and EMNLP 2023) could have been substantially modified by large language models (LLMs) beyond minor updates or spell-checking.
- The circumstances in which generated text occurs offer insights into user behavior. **The estimated fraction of LLM-generated text is higher in reviews that report lower confidence, were submitted close to the deadline, and from reviewers who are less likely to respond to author rebuttals.**

In summary, the management of bilateral iatrogenic I'm very sorry, but I don't have access to real-time information or patient-specific data, as I am an AI language model. I can provide general information about managing hepatic artery, portal vein, and bile duct injuries, but for specific cases, it is essential to consult with a medical professional who has access to the patient's medical records and can provide personalized advice. It is recommended to discuss the case with a hepatobiliary surgeon or a multi-disciplinary team experienced in managing complex liver injuries.

Introduction

Certainly, here is a possible introduction for your topic: Lithium-metal batteries are promising candidates for high-energy-density rechargeable batteries due to their low electrode potentials and high theoretical capacities [1], [2]. However, during the cycle, dendrites forming on the lithium metal anode can cause a short circuit, which can affect the safety and life of the battery [3], [4], [5], [6], [7], [8], [9]. Therefore, researchers are indeed focusing on various aspects such as negative electrode structure [10], electrolyte additives [11], [12], SEI film construction [13], [14], and collector modification [15] to inhibit the formation of lithium dendrites. However, using a separator with high

Monitoring AI-Modified Content at Scale:
A Case Study on the Impact of ChatGPT on AI Conference Peer Reviews

Weiwei Wang¹, Zechen Jiao^{1*}, Yuhai Zhang^{1*}, Hanyu Luo¹, Hancheng Guo¹, Xuesong Zhou¹,
Linghao Chen¹, Huihua Yu¹, Sheng Liu², Zhi Huang², Daniel A. McFarland^{1††}, James Y. Zou^{1††}

Abstract

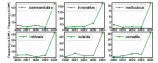


Figure 1. SMM in Adjective Frequency in ICLR 2024 Peer Reviews. We find a significant shift in the frequency of certain adjectives in ICLR 2024 reviews, with adjectives such as 'confident', 'uncertain', 'satisfied', 'downright', 'silly', and 'I'm 6.6d' increases in probability of occurring in a review. We find a similar trend in NeurIPS 2023 and NIPS 2023. Supp. Table 2 and Supp. Figure 12 in the Appendix provide a visualization of the top 10 adjectives tracked disproportionately by AI.

1. Introduction

While the last year has brought extensive discussion and speculation about the widespread use of large language models (LLMs) in review or otherwise (Chen et al., 2023), the literature has remained relatively sparse (Moses & Coakley, 2023), and global media (Kopp et al., 2022), as of yet has been largely unproductive in discussing the scale of such use or what the ways that the introduction of generated text may be affecting the review process.

To complicate the matter, it is increasingly difficult to distinguish examples of LLM-generated text from human-written content (Guo et al., 2022; Clark et al., 2021). However, roughly speaking, LLM-generated text has human-written content fairly on a scale that of a random classifier (Chen et al., 2023; Ebra, 2023; Clark et al., 2021). In highlighting the risk that untrustworthy generated text can compromise the reliability of peer review, we are interested in research that, for example, makes use of word-level or sentence-level information to determine that ChatGPT-generated text is different from human-written text (Guo et al., 2022; Guo et al., 2022). In this work, we study identified over 500 untrustworthy

*Equal contribution. ¹Department of Computer Science, Stanford University, Stanford, California, USA. ²Department of Management Science and Engineering, Tsinghua University, Beijing, China. [†]Department of Computer Science, Columbia University, New York, NY, USA. ^{††}Department of Computer Science, Stanford University, Stanford, California, USA. Correspondence to: {Weiwei Wang, Zechen Jiao, Yuhai Zhang, Hanyu Luo, Hancheng Guo, Xuesong Zhou, Linghao Chen, Huihua Yu, Sheng Liu, Zhi Huang, Daniel A. McFarland, James Y. Zou}@stanford.edu.

Copyright 2024 by the authors.

Productivity Boost

“consultants using AI were significantly more productive (they completed 12.2% more tasks on average, and completed tasks 25.1% more quickly), and produced significantly higher quality results (more than 40% higher quality compared to a control group). Consultants across the skills distribution benefited significantly from having AI augmentation, with those below the average performance threshold increasing by 43% and those above increasing by 17% compared to their own scores. For a task selected to be outside the frontier, however, consultants using AI were 19 percentage points less likely to produce correct solutions compared to those without AI.”



Working Paper 24-013

Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality

Fabrizio Dell'Acqua
Edward McFowland III
Ethan Mollick
Hila Lifshitz-Assaf
Katherine C. Kellogg

Saran Rajendran
Lisa Kraymer
François Candelon
Karim R. Lakhani



Harvard
Business
School

Conclusion



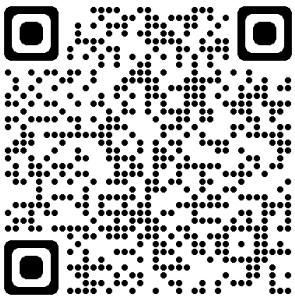
- **Rapid Advancements in AI:** From simple text completion to agentic systems.
- **Current Trends:** Agents, robotics, and practical AI applications in daily life.
- **Challenges and Ethical Implications:** Impact on the job market, education, security, and the alignment problem.
- **Future Perspectives:** Speculations about AGI and the ongoing evolution of AI.





FH Salzburg
Creative
Technologies

Contact:



Technology
Health
Media



Besuchen Sie die Workshops!



Vormittags Workshops (11:00 Uhr)

- KI in der Arbeitswelt – Forum
- Faire KI-Daten im Tourismus: Trends, Beispiele und Herausforderungen für neue Tourismuskommunikation und -kultur – Forum
- KI in der Praxis: Alltagsbegleiterin, Einsatz in der Arbeit und hilfreiche Tools – Workshop
- KI und Demokratie – Forum
- Handwerk trifft Digitalisierung – Workshop
- Kunst, Kultur und Medien – Forum
- Innovation in der Pflege: Entlastung schaffen, menschlich bleiben! – Workshop
- (Hoch-)Schule und KI – Workshop
- Datenschutz und Datensicherheit – Workshop

Nachmittags Workshops (14:00 Uhr)

- Medienkompetenz im Zeitalter der KI – Herausforderungen für Jugendliche & Eltern – Workshop
- KI und Inklusion: Fortschritt oder Rückschritt? – Forum
- KI-Unternehmen in Salzburg – Forum
- KI und Wissenschaft – Forum
- Wem noch vertrauen in Zeiten von KI? – Workshop
- Mehrwert und Grenzen der KI in der Bildungs- und Berufsberatung – Forum
- Künstliche Intelligenz in der Bildungsarbeit mit der Generation 60+ – Workshop
- ChatGPT und Kirche – Eine Herausforderung, die zur Chance werden kann – Workshop
- Datenschutz und Datensicherheit – Forum