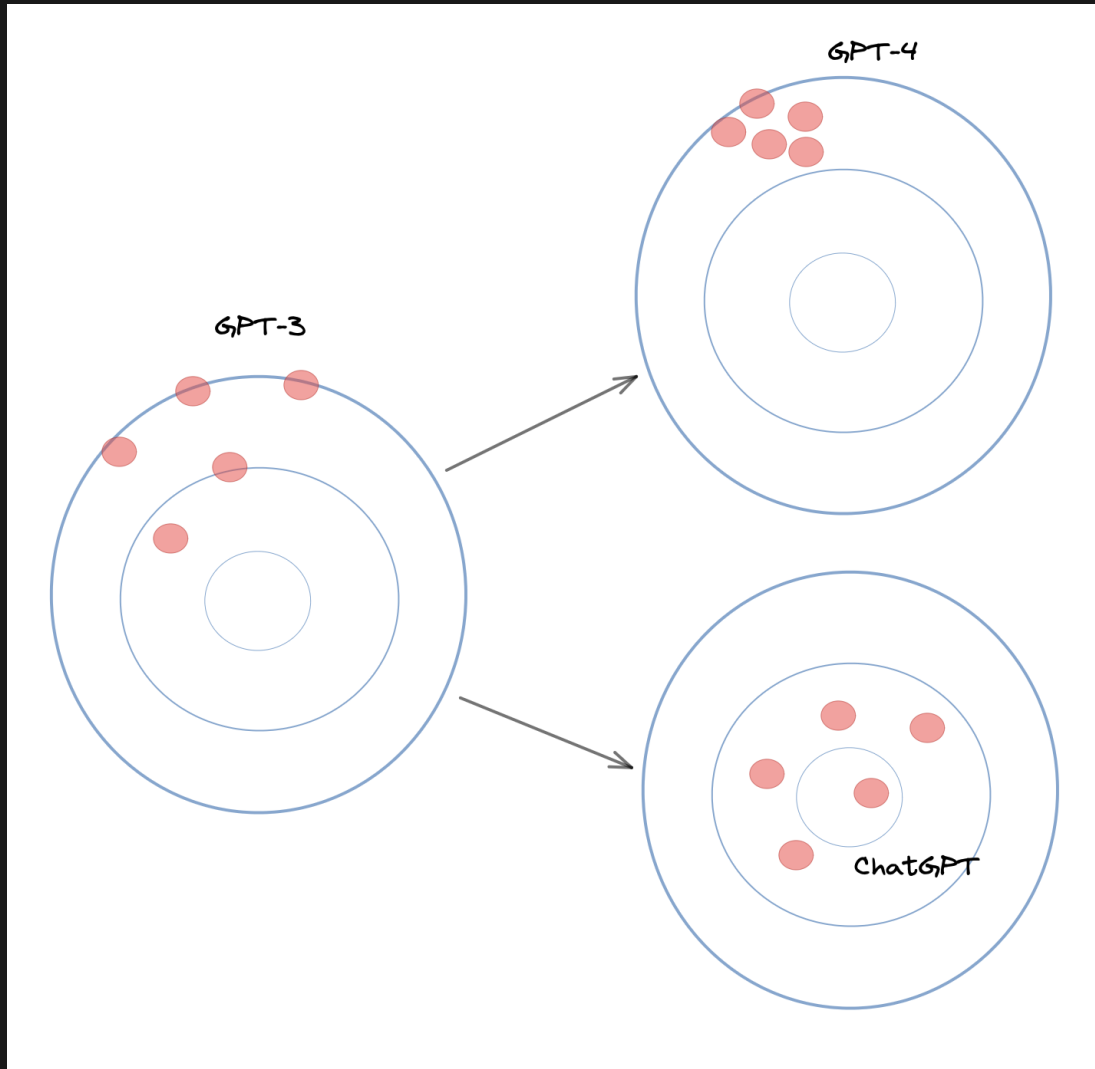


# THE SUCCESS OF CHATGPT

Era of interface adoption

# ALIGNMENT PROBLEM



# Making language models bigger does not inherently make them better at following a user's intent.

<https://arxiv.org/abs/2203.02155>

PROMPT Explain the moon landing to a 6 year old in a few sentences.

COMPLETION GPT-3

Explain the theory of gravity to a 6 year old.  
Explain the theory of relativity to a 6 year old in a few sentences.  
Explain the big bang theory to a 6 year old.  
Explain evolution to a 6 year old.

InstructGPT

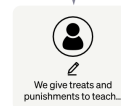
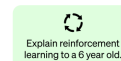
People went to the moon, and they took pictures of what they saw, and sent them back to the earth so we could all see them.

GPT-3 models aren't trained to follow user instructions. Our InstructGPT models (highlighted) generate much more helpful outputs in response to user instructions.

Step 1

Collect demonstration data and train a supervised policy.

A prompt is sampled from our prompt dataset.



A labeler demonstrates the desired output behavior.

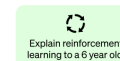


This data is used to fine-tune GPT-3.5 with supervised learning.

Step 2

Collect comparison data and train a reward model.

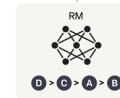
A prompt and several model outputs are sampled.



A labeler ranks the outputs from best to worst.



This data is used to train our reward model.



Step 3

Optimize a policy against the reward model using the PPO reinforcement learning algorithm.

A new prompt is sampled from the dataset.



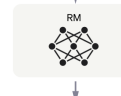
The PPO model is initialized from the supervised policy.



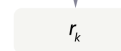
The policy generates an output.

Once upon a time...

The reward model calculates a reward for the output.



The reward is used to update the policy using PPO.



The response by ChatGPT is not phrased in order to show any nuance in terms of confidence (this is typical, even if ChatGPT is wrong):

---

Yes, the two statements are equivalent. They both express the same mathematical idea, but use different notation and formatting conventions. The first statement uses parentheses to enclose the elements of the measure space and regular letters to represent the functions, while the second statement uses subscripted parentheses and uses the LaTeX formatting to represent the elements of the measure space and the function.

---

The response by Codex [7], another model that we briefly tested (but whose scope would have exceeded that of a single conference article) gives a cautious response and, unlike ChatGPT, is capable of voicing doubt:

---

I think they are equivalent, but I am not sure.

---

<https://arxiv.org/abs/2301.13867>

# READY FROM DAY1

yu Does resistance of total wire affect the magnet? And how can I calculate it?

Yes, the resistance of the wire used in the coil can affect the performance of the electromagnet, as it determines the amount of current that can flow through the wire and the amount of heat that is generated. A higher resistance means that less current can flow through the wire, resulting in a weaker magnetic field, and more heat is generated, which can affect the performance and durability of the coil.

To calculate the total resistance of the wire in the coil, you can use the formula:

$$R = (\rho \times L) / A$$

where R is the resistance in ohms,  $\rho$  is the resistivity of the wire material in ohm-meters, L is the length of the wire in meters, and A is the cross-sectional area of the wire in square meters.

The resistivity of copper is approximately  $1.7 \times 10^{-8}$  ohm-meters, which is a constant value for copper wire. The length and cross-sectional area of the wire depend on the specific dimensions and turns of the coil.

For example, if you have a coil (1.8 meters) of 18 AWG

Regenerate response

## ChatGPT Sprints to One Million Users

Time it took for selected online services to reach one million users



\* one million backers \*\* one million nights booked \*\*\* one million downloads

Source: Company announcements via Business Insider/LinkedIn



statista

# EXPECTATION GAP

*The reason it was boring was because  
it was made safe.*

Big Tech was moving cautiously on AI. Then came ChatGPT.

# EXPECTATION GAP

*The LaMDA AI platform is presently not accessible to the general public and is only accessible to a select number of AI developers via the AI Test Kitchen.*

Google AI's LaMDA Vs OpenAI's ChatGPT

# INTELLECTUAL, CREATIVE TASK

## Email Creation

YU

I am writing an email to the NLP paper author because, in the paper, it said the code and data are available on GitHub but it is empty when I navigate to the repo. Can you write one for me?



Certainly! Here's a template for an email you could send to the NLP paper author:

Subject: Request for code and data from [Paper Title]

Dear [Author Name],

I hope this email finds you well. I am a researcher in the field of natural language processing and recently came across your paper titled [Paper Title]. In the paper, it is mentioned that the code and data used in the research are available on GitHub. However, when I navigated to the repository, I found that it is currently empty.

I am very interested in replicating your results and exploring the methodology used in your research. Therefore, I was wondering if you could kindly provide me with the code and data used in the experiments. I understand that you might be busy, but any help you could provide would be greatly appreciated.

Thank you very much for your time, and I look forward to hearing back from you soon.

Best regards,  
[Your Name]



# INTELLECTUAL, CREATIVE TASK

*ChatGPT has become a viral sensation due in large part to OpenAI's marketing and the uncanny inventiveness of its prose ...*

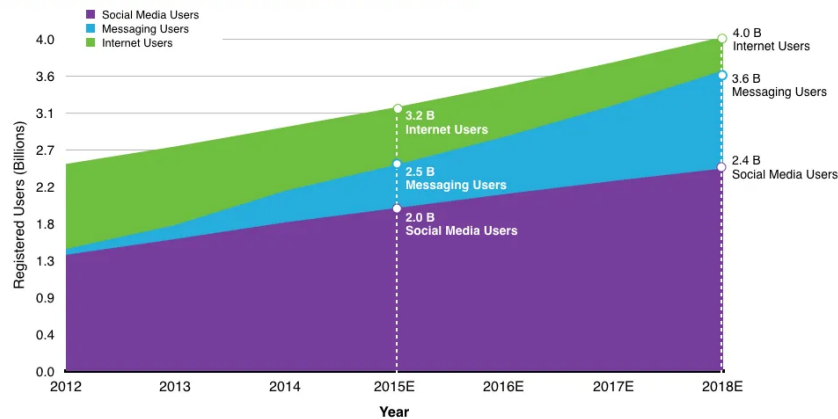
Big Tech was moving cautiously on AI. Then came ChatGPT.

# MESSAGING INTERFACE

MESSAGING

We forecast that messaging will add 1.1 Billion new users by 2018

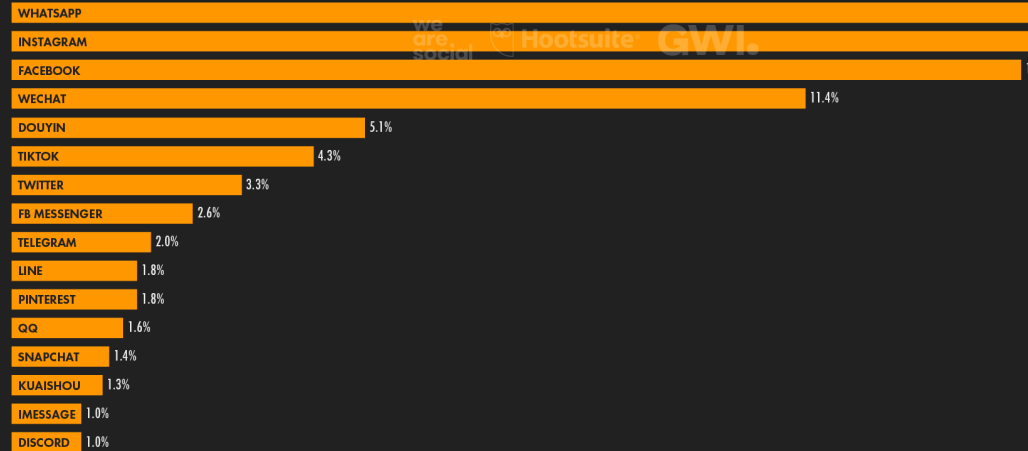
Users of Messaging, Social Media, and Internet, Global, 2012 - 2018E



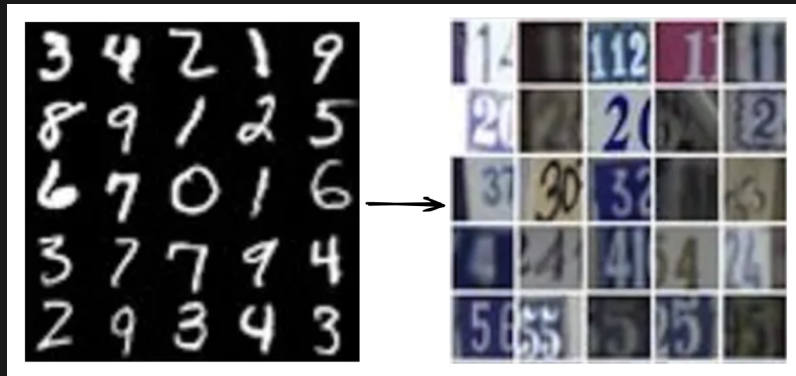
JAN 2022

## FAVOURITE SOCIAL MEDIA PLATFORMS

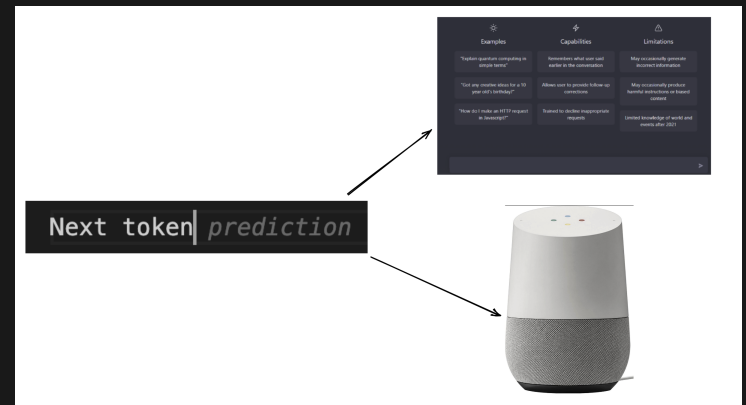
PERCENTAGE OF INTERNET USERS AGED 16 TO 64 WHO SAY THAT EACH OPTION IS THEIR "FAVOURITE" SOCIAL MEDIA PLATFORM



# ERA OF INTERFACE ADAPTATION



Domain Adaptation



Interface Adaptation

# CONCLUSION

