

The Internet (*Cyberspace*) of Inhuman things vs Humanity

Week Three & Four

"Successful Business leaders, especially in finance, education, healthcare, and supply chain, will have to deal with the challenges of *phase change*. If they don't, their businesses will evaporate into virtual space." -Sandy Weill, philanthropist & former CEO, Citigroup

WIREIMAGE.COM

Dear AI Overlords, Don't F*ck This Up.

| Inside OpenAI
BY STEVEN LEVY

OCT 2023 COME IN PEACE



+ MORE AI

Spot the Fakes.
BY CHRISTOPHER BEAM

Talk to Whales.
BY CAMILLE BROMLEY

Write a Bestseller.
BY VAUHINI VARA

Rule the World.
BY VIRGINIA HEFFERNAN

CREATE. CONNECT. CONDÉ NAST

What is Machine Learning?

Machine Learning is a branch of artificial intelligence (AI) and computer science that focuses on using data and algorithms to **imitate** how humans learn, gradually improving its accuracy.

***Machine Learning &
Deep Learning***

Predictive Modeling

Data Engineering

**Data Science With
Cloud Computing**

**Business Applications
of Data Science**

Data Mining

Robotics

Time Series Analysis

Text Mining

Berkman Klein Center for Internet & Society
at Harvard University - exploring cyberspace,
sharing in its study,
and helping to pioneer its development

Data & Society

Data & Society is a nonprofit research institute
that studies the social implications of
data-centric technologies,
automation, and AI.

Centre for Ethics

Toronto Canada

Interdisciplinary institute

@UofT

advancing research,
teaching & public discourse on ethics

Future of Humanity Institute

Multidisciplinary research institute
at the University of Oxford.

AI Now Institute

The AI Now Institute produces diagnoses
and actionable
policy research on artificial intelligence.

MIT Media Lab

The Media Lab is an interdisciplinary
creative playground rooted squarely
in academic rigor

Georgetown Law

"Law is but the means; justice is the end."

The Alan Turing Institute

The Alan Turing Institute, UK's national institute for data science and artificial intelligence.

Other organizations that focus on the ethics of Artificial intelligence includes:

- Leverhulme Centre for the Future of Intelligence (CFI)
- Partnership on AI
- AI Now Institute
- Center for the Governance of AI
- Centre for the Study of Existential Risk
- Center for Human-Compatible AI
- OpenAI
- DeepMind Ethics & Society

Artificial Intelligence has the potential to encode and exacerbate biases by reflecting the assumptions, interests, and world views of its developers and users. In addition, machine learning - currently the most common form of AI - looks for patterns in real-world examples (“training data”), which can lead to problems in multiple ways.

In the early days (*I am talking three years ago*) of facial recognition, People with darker skin tones were matched in a database of people with skin tones of white people, and the patterns were set to scan consistently on white people; for example, the training data may omit certain types of people or ethnicities gathered within a narrow cultural context.

Why AI is trial and error: In one instance, a tool designed to sharpen the blurry images of faces was found to turn people with darker skin tones white consistently. Another application turned them dark consistently.

“ChatGPT” is a web app. **KEY** (*you can access it in your browser*) is designed specifically for chatbot applications—and optimized for dialogue.

It relies on **GPT** to produce text, like explaining code or writing poems. GPT, on the other hand, is a language model, **NOT** an app.”

GPT stands for ***Generative Pre-trained Transformers***. GPT is a family of ***neural network models*** that use transformer architecture to generate human-like text

ChatGPT is an AI-powered tool.

It was developed by OpenAI and launched in November 2022. ChatGPT uses GPT-4, a large language model with deep learning, to produce human-like text. In April 2023, it had 173 million users actively using the application daily.

ChatGPT's popularity has caused it to trend on Twitter, Facebook, Threads, etc., and even led to a brief crash of various sites. The versatility and *human-like quality* of its responses have captured the attention of the media, the tech industry, and some members of the public, especially law enforcement and regulators.

OpenAI projects that ChatGPT made \$200 million in revenue last year, forecasting \$1 billion in 2024. This doesn't include all the PLUGINS being developed by various other companies.

Here are some specific examples of how GAI/ChatGPT is being used today:

- **In customer service, GAI/ChatGPT** creates chatbots that can answer customer questions more naturally and engagingly. This can free up customer service representatives to handle more complex issues.
- **In Education, GAI/ChatGPT** is used to create personalized learning materials for students. This can help students to learn at their own pace and in a way that is most effective for them.
- **GAI/ChatGPT is used in entertainment** to write scripts for movies and TV shows. This can help to create more engaging and creative content.



WHAT YOU CAN DO WITH CHAT GPT

BERNARD MARR

Top 14 OpenAI ChatGPT Plugin Types You Can Create

- 1) *Language Translation Plugins.*
- 2) *Voice Recognition Plugins.*
- 3) *Chatbot Plugins & Chat Analytics.*
- 4) *Content Generation Plugins.*
- 5) *Sentiment Analysis.*
- 6) *Entity Recognition. (Used in HR)*
- 7) *Knowledge Base Integration.*
- 8) *Multi-channel Integration.*
- 9) *Conversational Flow Management*
- 10) *Personalization*
- 11) *Natural Language Generation*
- 12) *E-commerce*
- 13) *Customer Service*
- 14) *How Plugins are Used*
- *You Mine It, You Grow It, You Manufacture It*

There are four versions of ChatGPT:

- *The default GPT-3.5 version*
- *The paid default GPT-4 version*
- *The browsing-enabled GPT-4 version*
- *The plugin-powered GPT-4 version*

There are also several free alternatives to ChatGPT, including *Bard(Gemini), Bing, Claude, Chatsonic, Perplexity AI, YouChat, Jasper, and Character AI*. To name just a few.

What is the best alternative to ChatGPT?

The best alternative to ChatGPT completely depends on one's specific needs. However, GPT-4, IBM Watson Assistant, and DialoGPT are at the top of the list.

Overall, the combination of OpenAPI and a ChatGPT plugin could provide a powerful tool for developers and users alike, making it easier to understand and use APIs

- ***For similar capabilities:***

- **Google Gemini (formerly Bard):** built on a powerful language model and might be a good option if you're looking for something that can handle complex tasks and generate creative text formats, like ChatGPT.
- **Claude AI:** This is another strong contender, particularly good for its factual accuracy and safety features.
- **Perplexity AI:** Like (Gemini), Perplexity integrates real-time web searches, making its responses more up-to-date.

For specific uses:

- **Microsoft Copilot:** This is an excellent option for programmers, as it can generate code and suggest completions while you work.
- **Character AI:** If you're looking for a more entertaining experience, *Character AI* lets you chat with interesting fictional or historical characters.

For budget-friendly options:

- **OpenAI Playground:** This is a good choice if you want to experiment with a similar language model for free, though it has a paid tier with more features.
- **Writesonic:** This offers a free plan with some limitations, but is cost-effective for creative writing tasks.

GPT-5 might arrive this summer as a “materially better” update to ChatGPT

OpenAI is expected to release **GPT-5**, an improved version of the AI language model that powers **ChatGPT**, sometime in mid-2024—and likely during the summer. Two anonymous sources familiar with the company have revealed that some enterprise customers have recently received demos of **GPT-5** and related enhancements to **ChatGPT**.

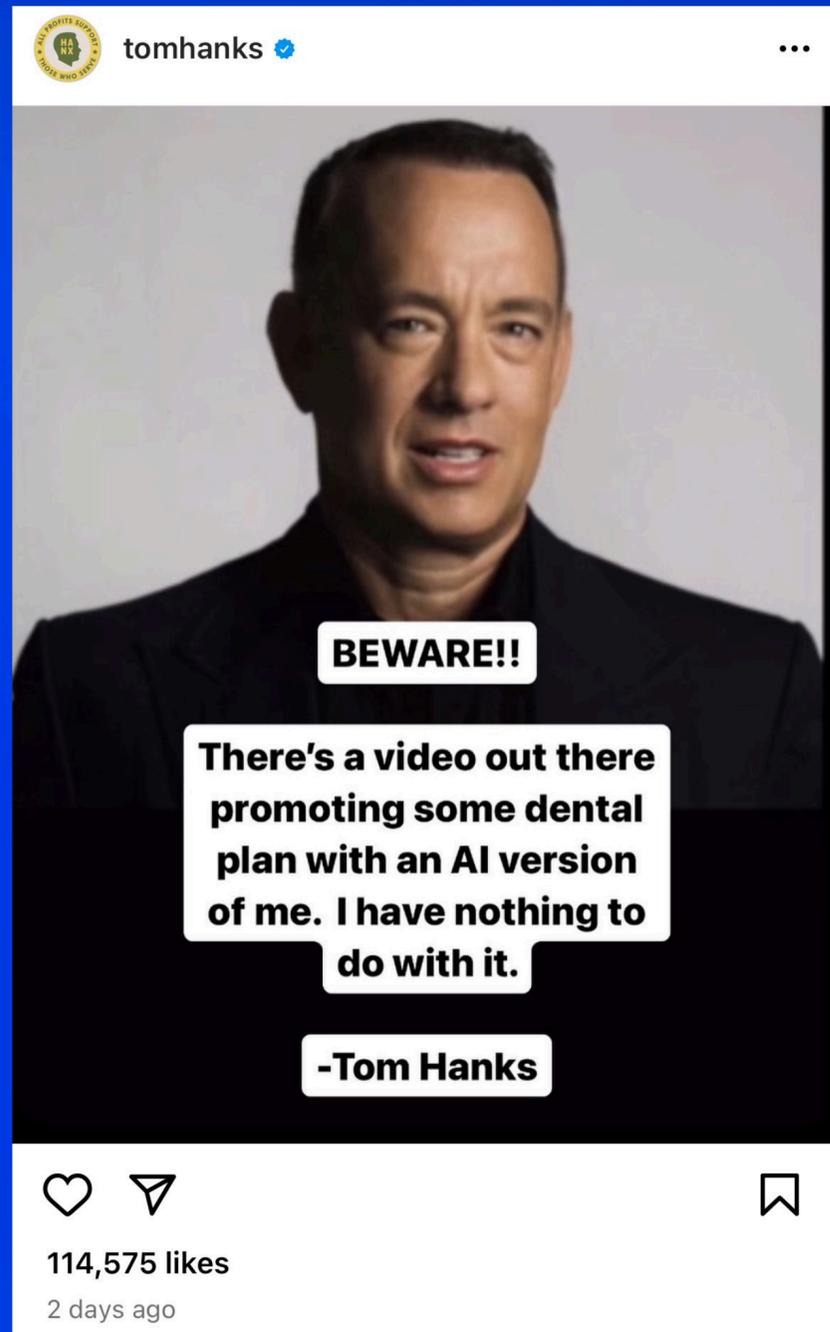
A timeline of ChatGPT-related events

Since ChatGPT's launch last year, a lot has happened in OpenAI, governments, the legal system, and with competitors. Here's a brief timeline of events related to ChatGPT's release and its impact. There have been far too many events to list all of them comprehensively here, so we're presenting an incomplete, broad overview of historically important moments.

I will email you the entire timeline.

Tom Hanks Warns of Dental Ad Using A.I. Version of Him

Tom Hanks and the CBS anchor Gayle King both said their likenesses were used in unauthorized advertisements, as worries have grown over the unregulated use of artificial intelligence.

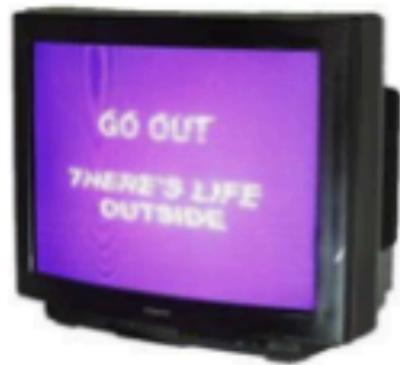


More significant to AI-generated works is the issue of **copyrightability**, or what copyright protects.

Copyright only protects original works by human creators. Under U.S. copyright law, copyright protection does not extend to **non-human creators**, which is an obvious problem for AI-generated works.

Thanks to the recent advances in **Artificial Intelligence** (specifically **Deep Learning**), generative models such as Dall-E2 (***DALL·E 2 is an AI system that can create realistic images and art from a description in natural language***) and **ChatGPT** (powered by OpenAI) can now produce unique images and texts with an uncanny human-like quality, often indistinguishable from what a human would have written.

The media, VCs, and institutional investors are abuzz with stories on **Generative Artificial Intelligence**. (GAI) Already utilized in customer service and content-writing areas, generative text-based AIs like ChatGPT seem poised to transform the web – and our society.



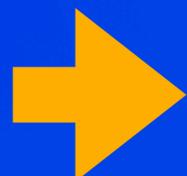
Old Media



New Media



Synthetic Media



https://phenaki.video/?mc_cid=9fee7eeb9d&mc_eid=f568038908#

Generative artificial intelligence (GAI) and ChatGPT are exciting for several reasons.

- **They can generate human-quality text.** This can potentially revolutionize many industries, such as customer service, education, and entertainment. For example, GAI can be used to create chatbots that can answer customer questions more naturally and engagingly. It can also be used to create personalized learning materials for students or to write scripts for movies and TV shows.
- **GAI systems are trained on massive text and code datasets, allowing them to learn and adapt to new information without human intervention.** This means they can become more accurate and useful over time.
- **GAI tools have the potential to democratize creativity. Anyone can use them** to create creative content, such as poems, code, scripts, and musical pieces. This means that people who are not trained artists or programmers can still express their creativity and share it with the world.

And yet, what about deepfakes and synthetic media?

ChatGPT's potential to perpetuate bias has sparked growing concern as it is trained on internet texts that may include biased or deceptive information. This raises worries about its use in journalism, education, employment, law enforcement, and healthcare. ChatGPT clearly has the potential to be used to generate deepfake texts that appear human-written yet are machine-generated. As technology advances, it will become harder to tell the difference between human-written and machine-generated texts. The implications for politics, journalism and social media are seriously worrisome, as it will most certainly be used to spread misinformation and disinformation, influencing public opinion and elections.

GPT is AI-powered. *However*, ChatGPT has some disadvantages, including Limited knowledge in these early stages:

- **Emotional intelligence:** ChatGPT is designed to have human-like conversations but lacks human emotional intelligence.
- **Takes things too literally:** ChatGPT sometimes misses the mark with its responses by failing to consider the original question.
- **Privacy issues:** ChatGPT has been criticized for user privacy and data security.
- **Inaccuracies and ambiguities:** ChatGPT sometimes produces texts that sound plausible or convincing but are incorrect.
- **Biased answers:** ChatGPT has the potential to generate biased responses.
- **Ethical issues:** ChatGPT has been prohibited in certain educational institutions.
- **Academic integrity:** Academic integrity is the primary concern for using ChatGPT in higher education.

Here are some ways to tell if AI-generated content is fake:

- **Lack of personal touch:** AI-generated content doesn't have the same raw emotions, real-life anecdotes, or personality that most readers recognize.
- **Lack of credible sources** If content cites credible sources, especially online content, it might be AI-generated.
- **Style** Some types of AI-generated content, like legal briefs, are often written dryly. Other types, like newspaper editorials, are written more colorfully and nuancedly.
- **Human instinct** Reading between the lines and using human instinct may help. For example, AI can have trouble with words.

Other ways to tell if AI-generated content is fake include:

Hands

AI can sometimes generate hands with too many fingers or fingers that blend together in weird ways.

Images

It can be hard to distinguish between an AI-generated image and a real photograph. The *Turing Test* is one method of evaluating how well AI can think like a human.

The Turing Test is arguably one of the most well-known methods of evaluating how well artificial intelligence (AI) can think like a human. Turing believed that a computer could be described as intelligent if it can mimic human responses under specific conditions.

Turing Test:

During the Turing Test, the human interrogator asks several questions to both players. Based on the answers, the interrogator attempts to determine which player is a computer and which player is a human respondent.

Player A
Computer



Player B
**Human
Responder**



Player C
Interrogator



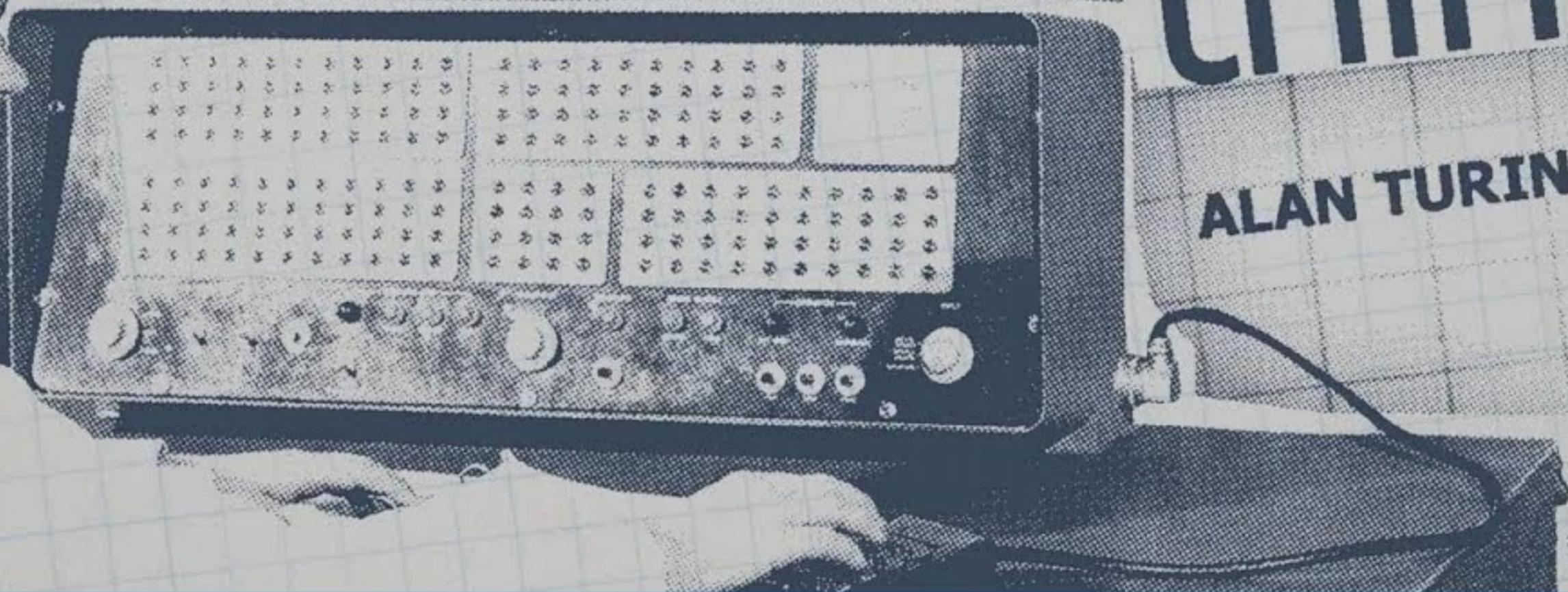
- Question to Respondents
- Answers to Question



NEWS



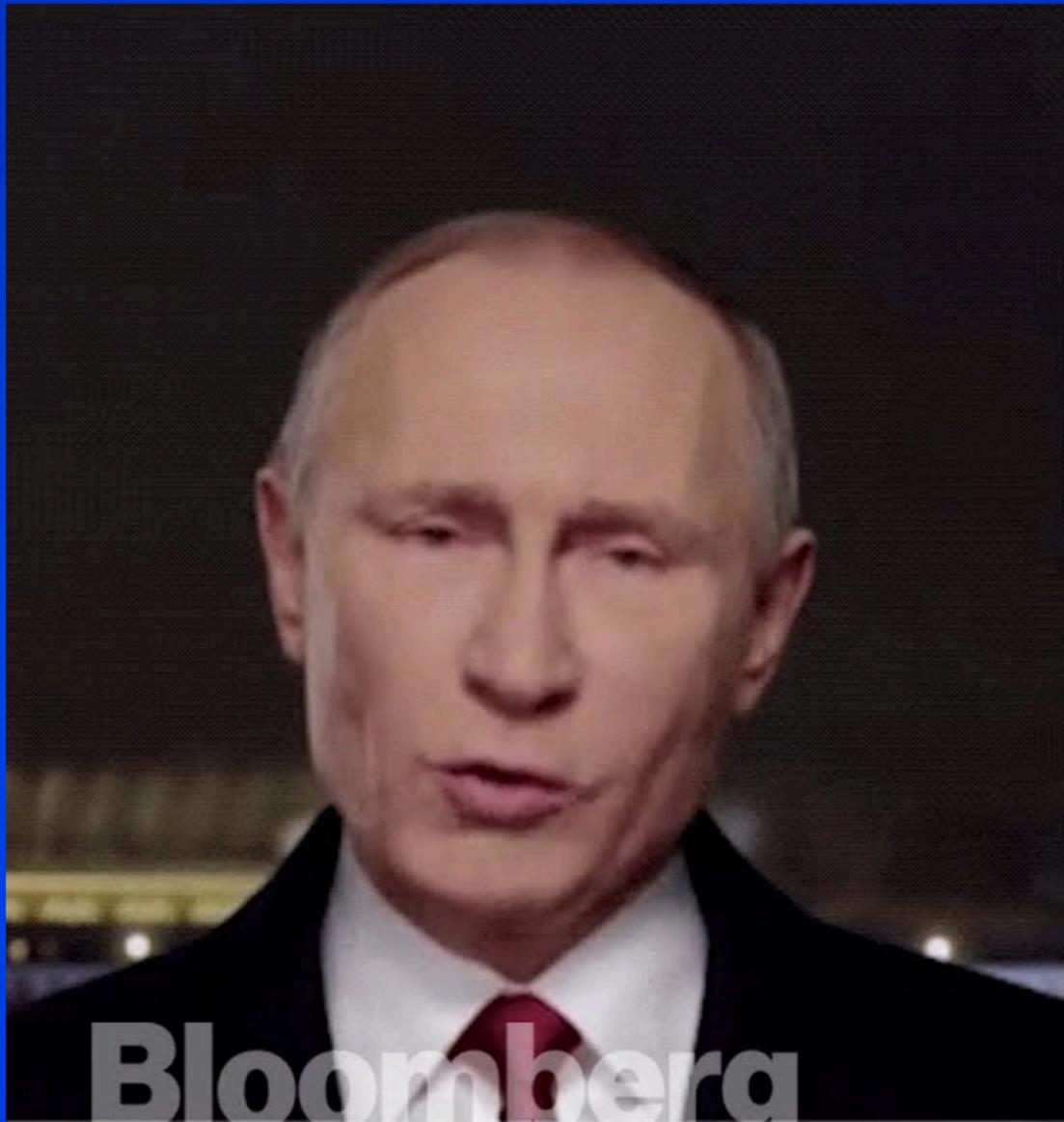
I propose to consider the question,
"can machines think?"



ALAN TURING, 1950







Reference



Our Result





The hands in this one look very realistic (even down to the barely visible hair on the hand), but the watches in the box are only approximations of watches



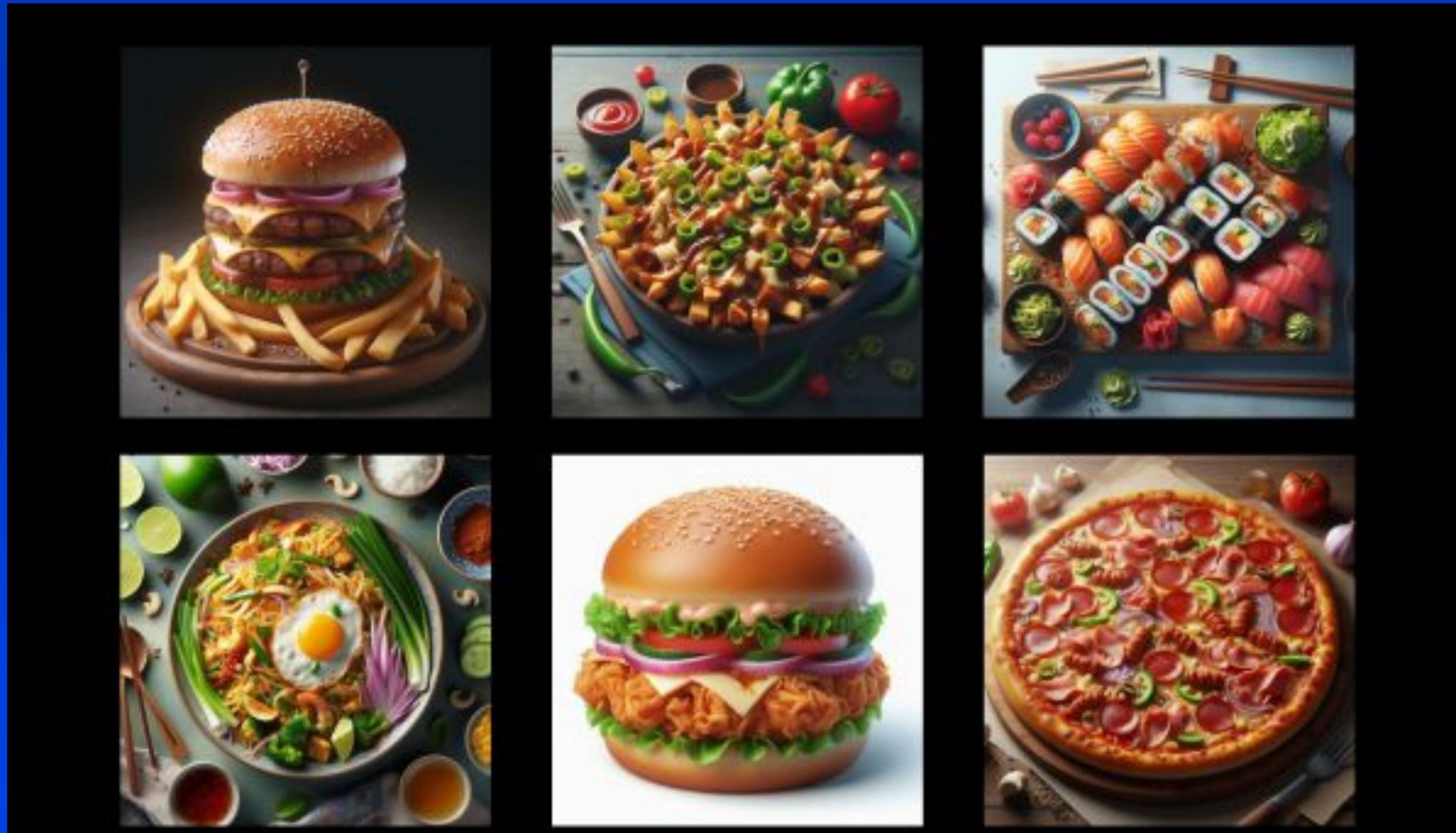
There is an uncanny sheen to the man's suitcase, suit, and hair. Even the buildings lack details when you lean in to take a closer look.



AI currently struggles with words. I asked it to generate a photograph of Montreal from 1931. First off, the tall skyscraper right of center is clearly out of place, and while the overall architecture certainly suggests Old Montreal, anyone who has been to the city will know that this is not Montreal. Spot the letters and see if you can read the words.



When asked to generate fake paintings, AI is pretty good. The above shows DALL-E aping Rembrandt. But the painting is too clean. Look at that background. It should show the wear and tear of time. Also, notice the man's collar. Those white tips are very common in AI art.



And the next time you feel like ordering a meal through a food delivery app, remember that some kitchens are using AI-generated images of their offerings. Notice the unnatural colors and the lack of details.

Turing Test today

Despite being developed more than 70 years ago, the Turing Test is still used today to assess artificial intelligence.

In 1966, a German programmer created ELIZA -- a program now considered a chatbot -- which replicated the behavior of a psychologist and was considered the first to have passed the Turing Test potentially.

More recently, in 2014, a computer program named Eugene Goodman, which simulates a teenage Ukrainian boy, was found to pass the Turing test, as did Google's LaMDA in 2022.



Model: GPT-4

AD

TV news script in 3 sentences or less explaining why AI-generated text can be difficult to detect.



Experts warn that AI-generated text can seamlessly replicate grammar, context, and even emotional nuances, making it harder to identify red flags or distinguish real news from fabricated stories. The need for improved detection tools and critical thinking skills is becoming crucial to combat the growing threat of AI-generated misinformation.