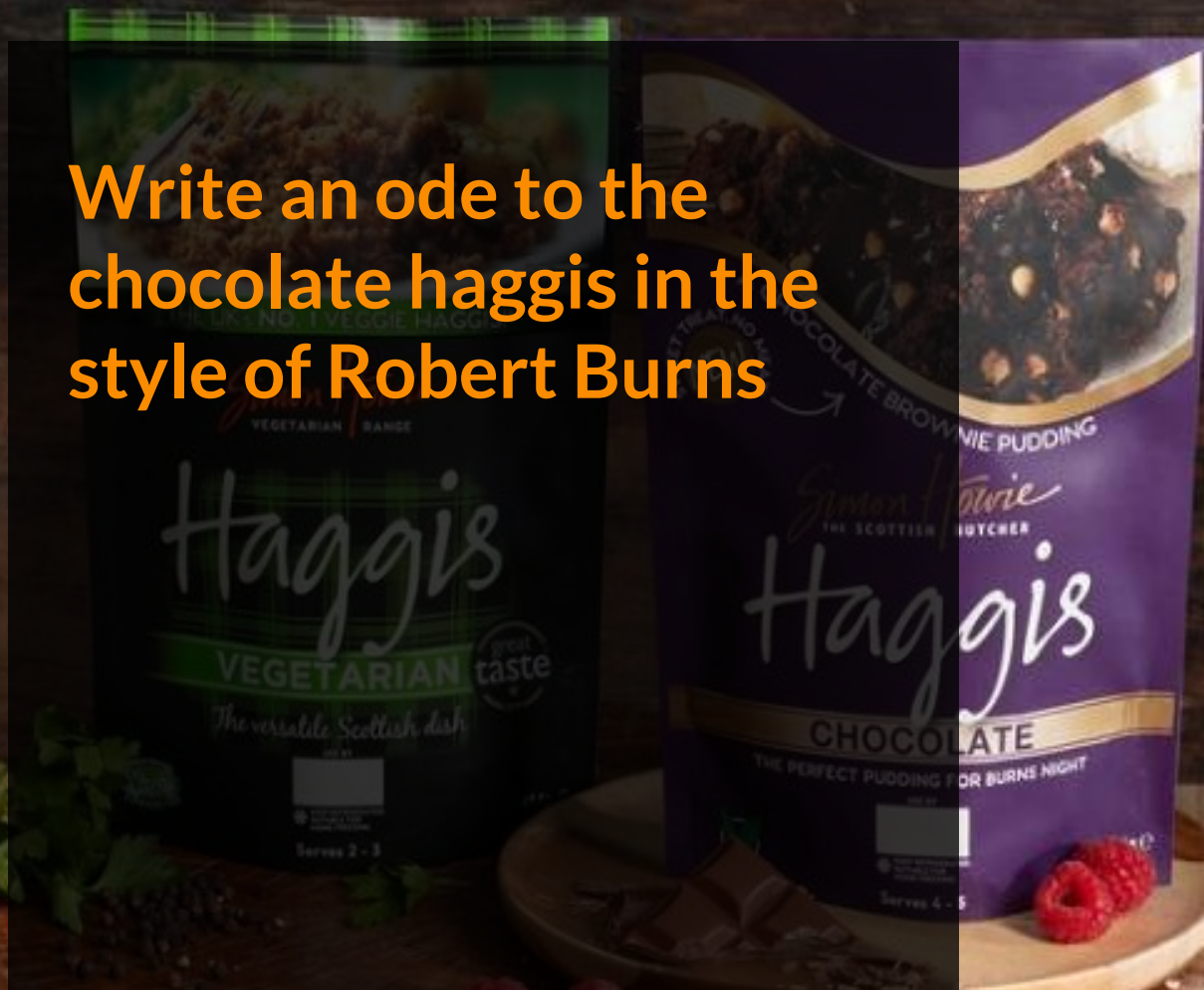

What on earth to make of ChatGPT?

Simon Thompson

BCS Kent Branch, 7 November 2024



Write an ode to the chocolate haggis in the style of Robert Burns



Write an ode to the chocolate haggis in the style of Robert Burns

Upon the brawny table laid,
A feast of oddity displayed.
A haggis made of cocoa sweet,
A daring twist, a tempting treat.



Write an ode to the chocolate haggis in the style of Robert Burns

Oh, chocolate haggis, rich and rare,
In Burns' land, a culinary dare.
No neeps or tatties by its side,
But cocoa beans in sweetness bide.



Write an ode to the chocolate haggis in the style of Robert Burns

With every bite, a rich surprise,
A Highland melody in cocoa guise.
No bagpipes play, no tartan's worn,
Yet chocolate haggis, a feast reborn.



Write an ode to the chocolate haggis in the style of Robert Burns

So here's to thee, this daring dish,
A modern twist, a chocolaty wish.
In Burns' spirit, let's all partake,
Of chocolate haggis, for Scotland's sake!



**Plausible
pastiche,
but no
sense of
right and
wrong.**



**Plausible
pastiche,
but no
sense of
right and
wrong.**

Artificial life

Talos

Protector of Crete





Silver Swan

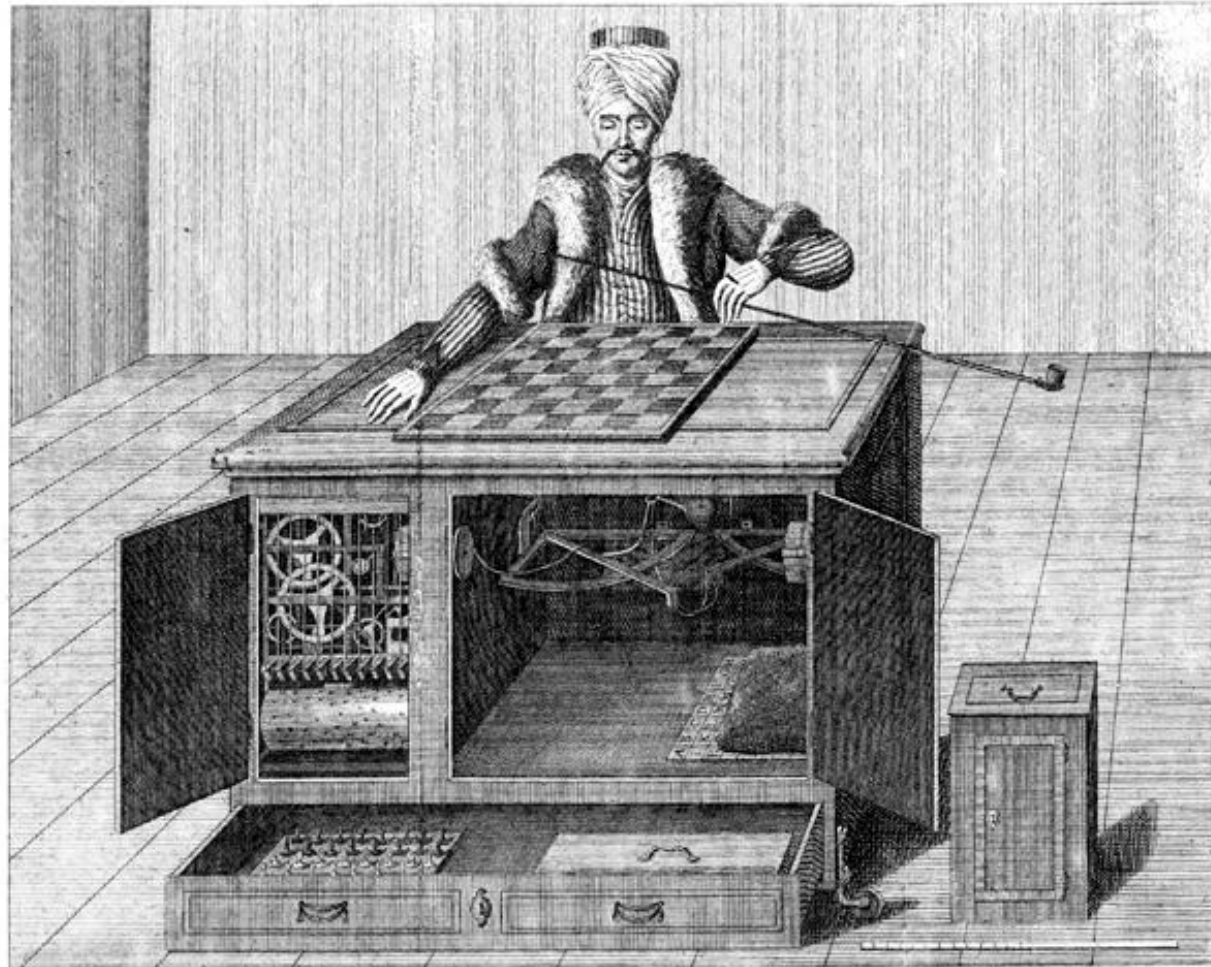
John Joseph Merlin
and James Cox, 1773

Bowes Museum,
Barnard Castle

Schachtürke

Wolfgang von Kempelen,
1770

<https://www.mturk.com>

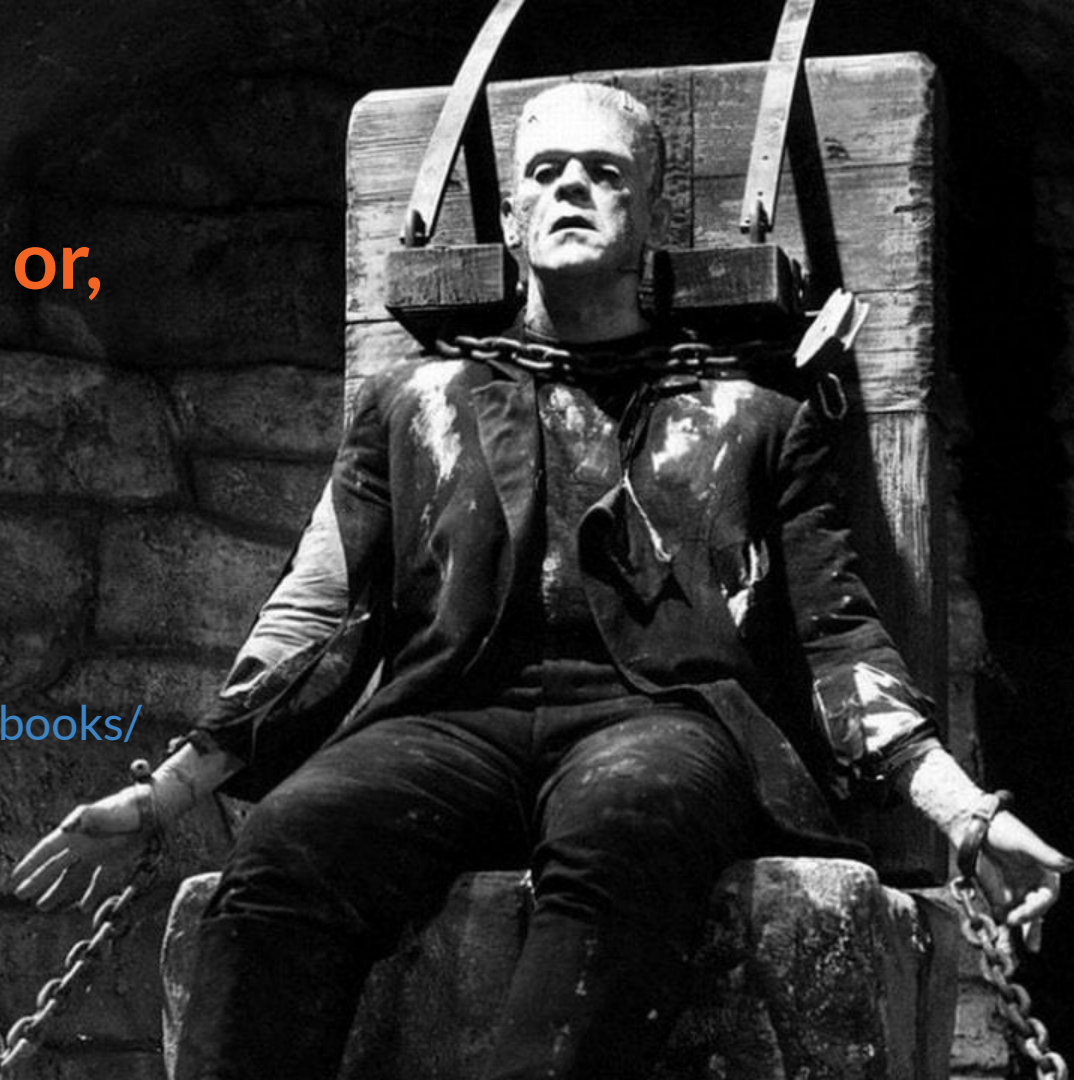


W. de Kempelen del. Che a Mechel, exaud. Basilea. P. G. Pintz, sc.
Der Schachspieler, wie er vor dem Spiel gezeiget wird von ihm. Le Joueur d'Échecs, tel qu'on le montre avant le jeu, par devant.

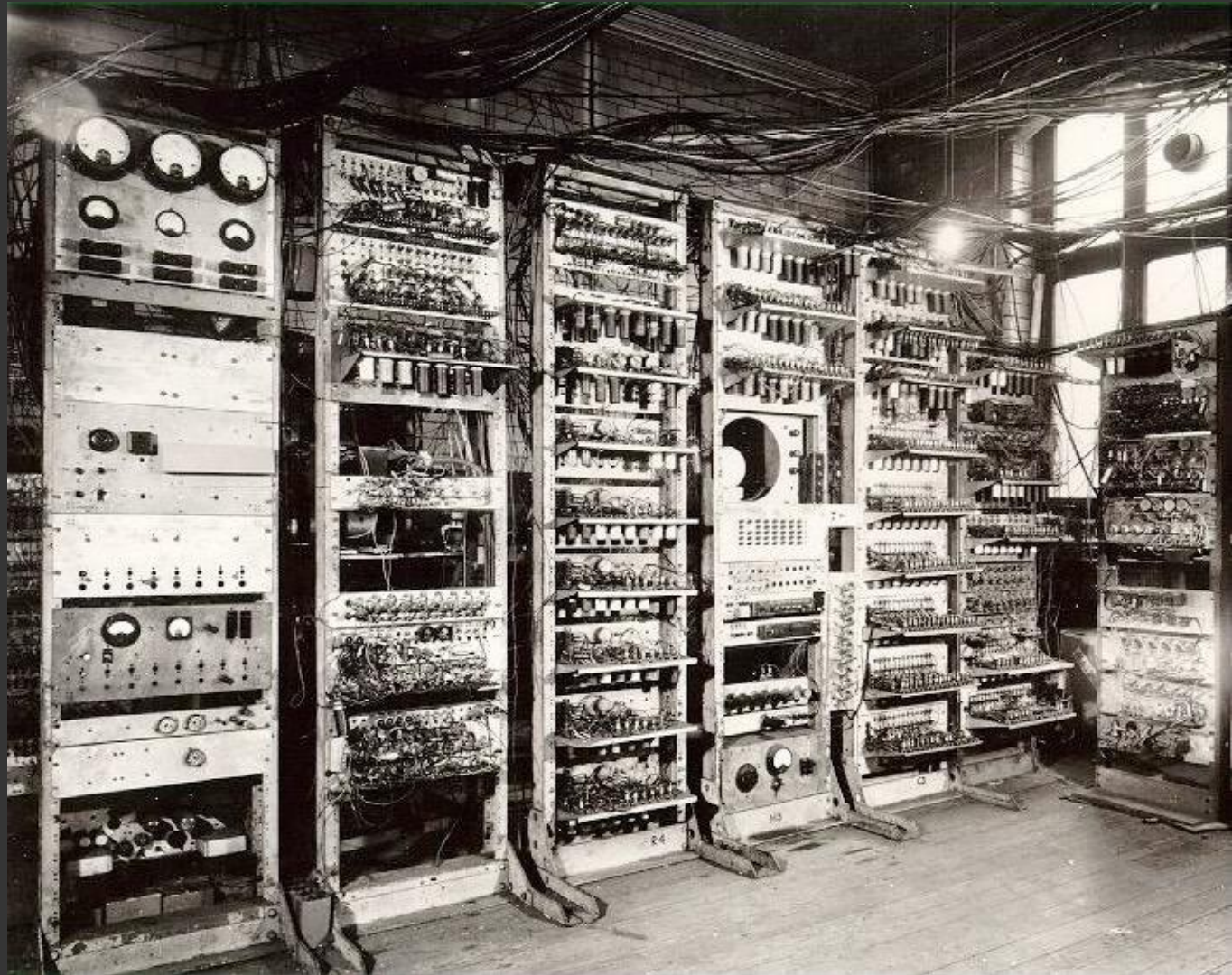
Frankenstein; or, The Modern Prometheus

Mary Shelley, 1818

[https://gutenberg.org/ebooks/
42324](https://gutenberg.org/ebooks/42324)

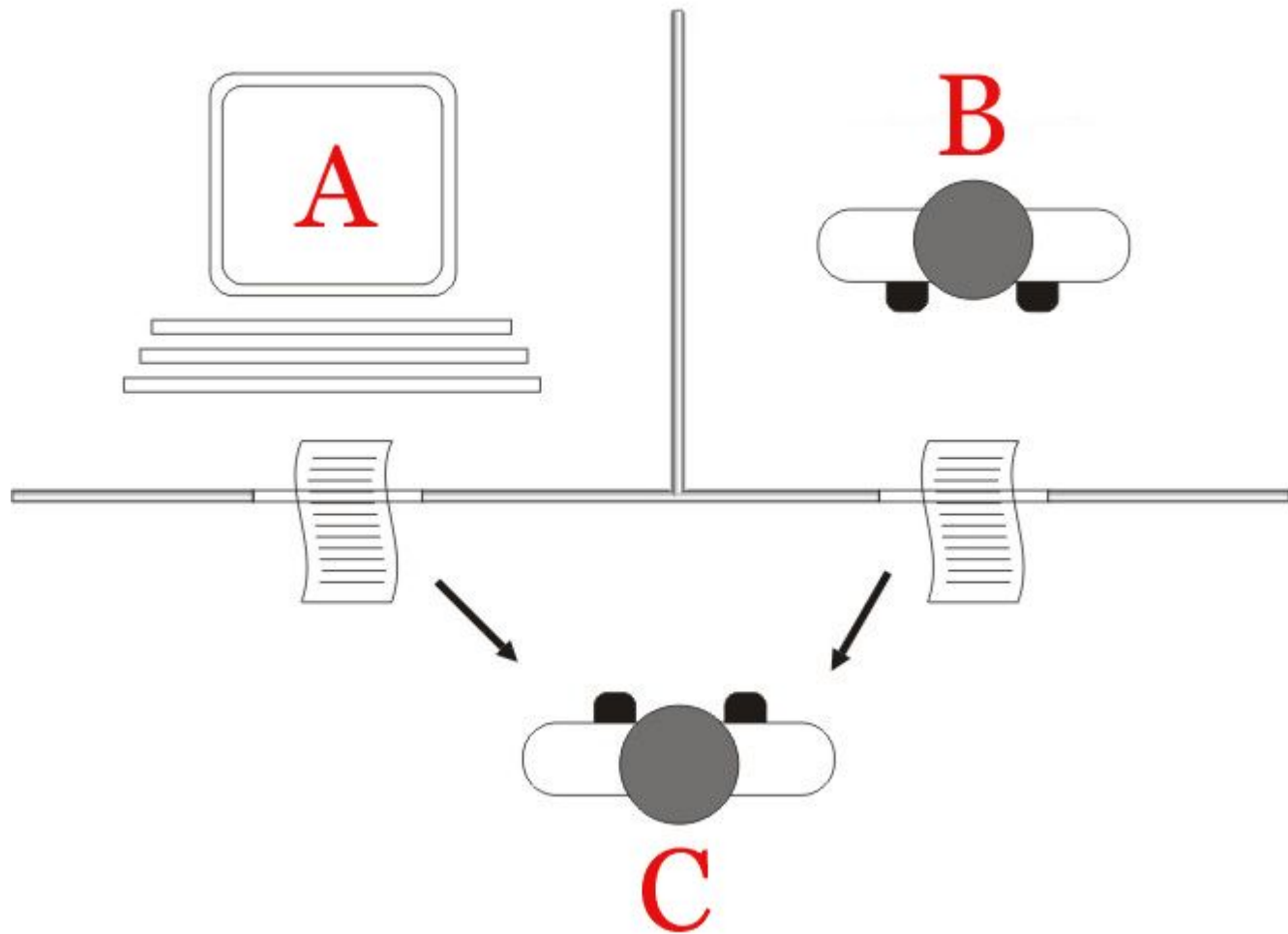


Machine
intelligence



*“Can
machines
think?”*





Artificial Intelligence is ...

... behaving in an
“intelligent” way.

... reproducing
human thought
processes.

Cognitive science

... embodying
human behaviour.

Robotics / AGI

—

Behaving intelligently ...



Calculating

Almost since their inception, computers have been much better than us at calculation.

Reasoning

Computer algebra

Formalising “real”
maths finally
happening in 2020s

Symbolic *versus*
sub-symbolic

*54·43. $\vdash :: \alpha, \beta \in 1 . \supset : \alpha \wedge \beta = \Lambda . \equiv . \alpha \vee \beta \in 2$

Dem.

$\vdash . *54·26 . \supset \vdash :: \alpha = \iota'x . \beta = \iota'y . \supset : \alpha \vee \beta \in 2 . \equiv . x \neq y .$

[*51·231] $\equiv . \iota'x \wedge \iota'y = \Lambda .$

[*13·12] $\equiv . \alpha \wedge \beta = \Lambda \quad (1)$

$\vdash . (1) . *11·11·35 . \supset$

$\vdash :: (\forall x, y) . \alpha = \iota'x . \beta = \iota'y . \supset : \alpha \vee \beta \in 2 . \equiv . \alpha \wedge \beta = \Lambda \quad (2)$

$\vdash . (2) . *11·54 . *52·1 . \supset \vdash . \text{Prop}$

From this proposition it will follow, when arithmetical addition has been defined, that $1 + 1 = 2$.



The expert

Expertise vs common sense.

Hype and “AI Winters”

“The most notorious disappointments ... have appeared in the area of machine translation, where enormous sums have been spent with very little useful result”

Lighthill report, 1973.

Fifteen years later, Expert Systems, and in particular the 5th Generation Project, did not deliver: too brittle, too hard to maintain, ...

Symbolic vs Sub-symbolic

Solve problems by manipulating a symbolic representation.

Rules express intelligence.

Offers explanations: a proof in maths, or a trace of how a solution is reached.

“GOFAI”

Symbolic vs Sub-symbolic

Solve problems by manipulating a symbolic representation.

Rules express intelligence.

Offers explanations: a proof in maths, or a trace of how a solution is reached.

“GOFAI”

Statistical methods: inferring parameters in complex models.

Intelligence grounded in training.

Mechanisms loosely based on neural behaviour.

Short on explanation: but so are we at this level.

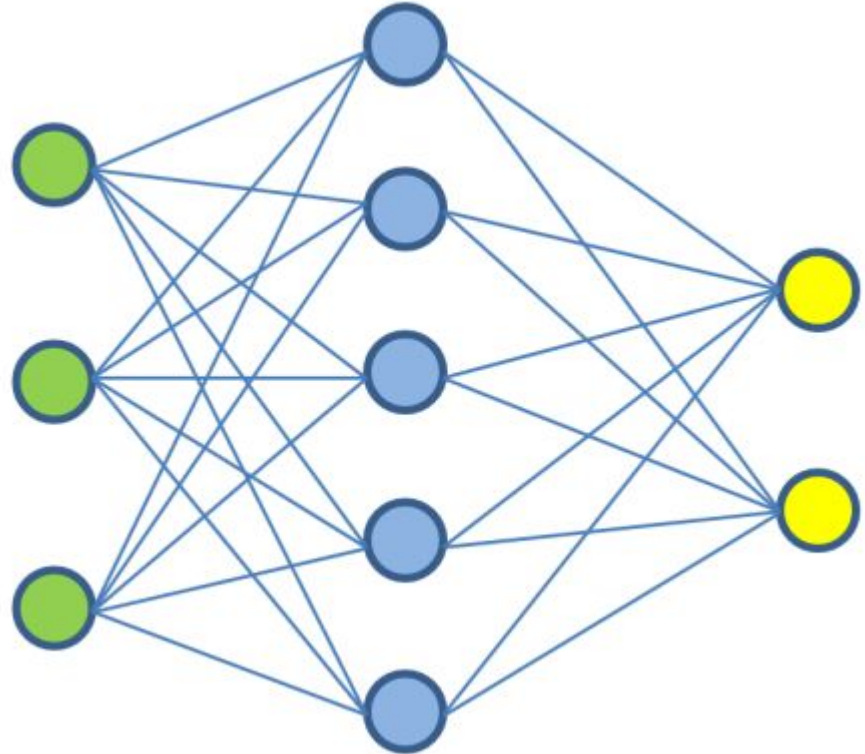
Neural networks

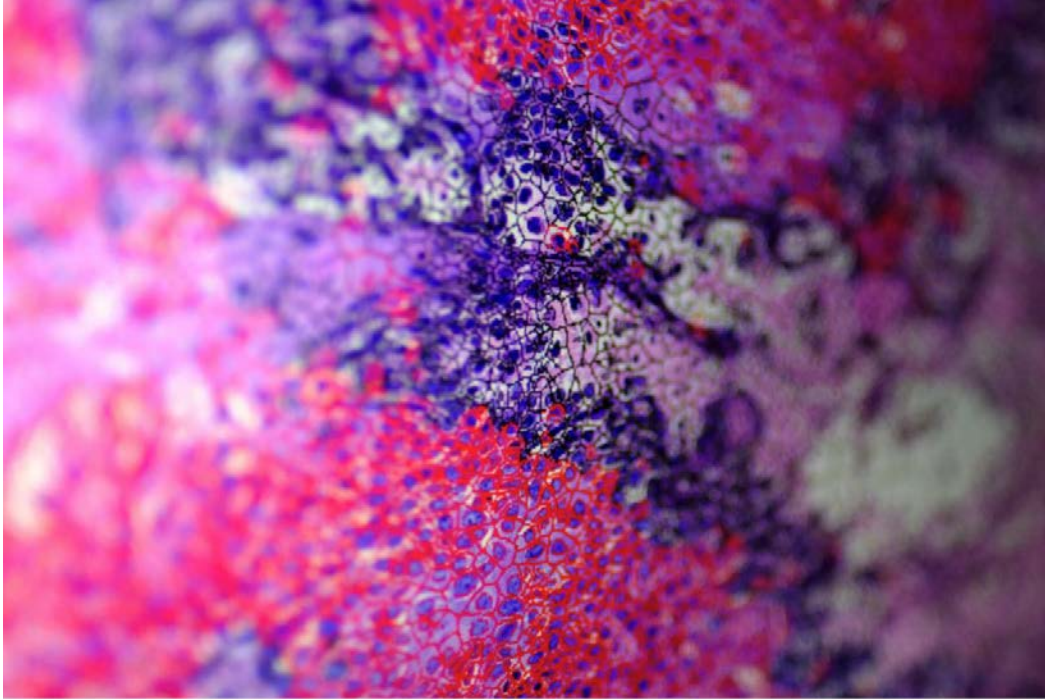
Information flow left to right.

Linear combinations + trigger values

Learn from data set

Classification (e.g. vision) or
transduction (e.g. translation)





A computer algorithm identifying individual cells in an image of a tumour. Credit: Carl Dupont/Shutterstock.com

Vision

Pattern recognition

Neural networks

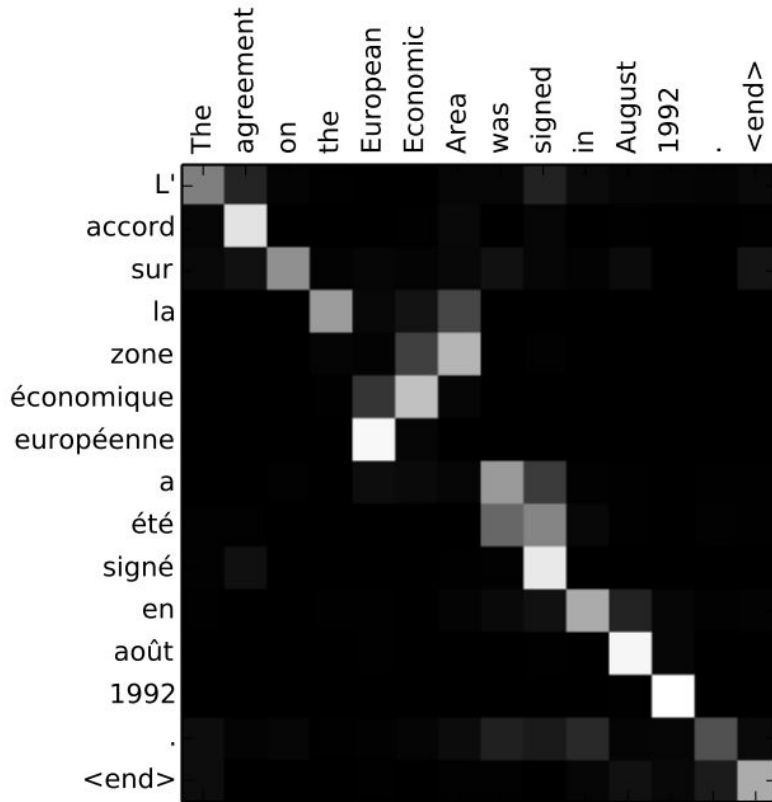
Machine intelligence



<https://deepmind.google/technologies/alphago/>



Playing
games



Transformers

Solving the problem of relevance or *attention*: using the wider context.

More efficient than previously ...
... so can train on larger data set.

Generative Pre-trained Transformers
such as *Chat-GPT*.

Language

**Speech synthesis
and recognition,
handwriting
recognition.**

Routine technology

Translation

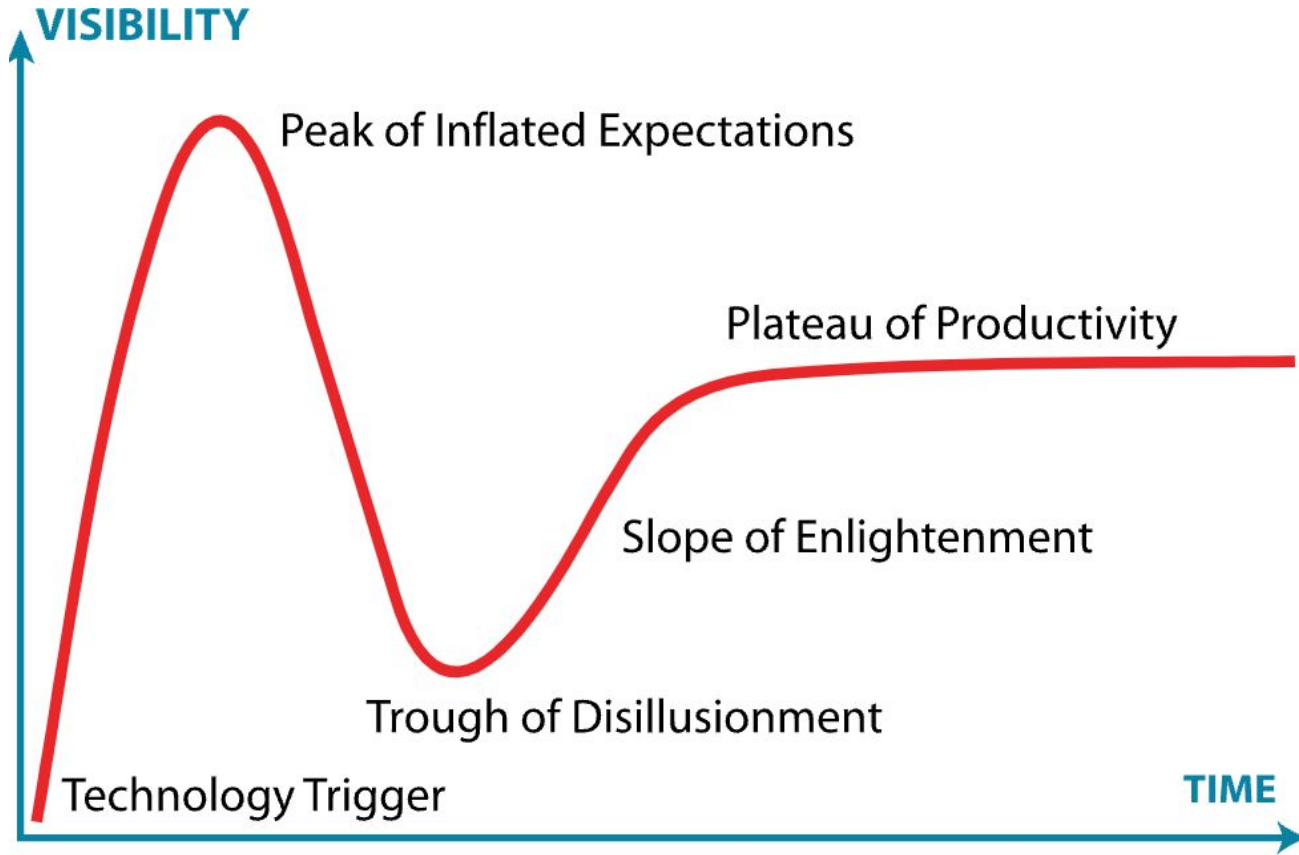
Getting there ...

**Generation via
LLMs**

ChatGPT et al

—

What on earth to make of ChatGPT and AI?







Here to stay

The screenshot shows the Wikipedia article for ChatGPT. The page title is "ChatGPT" and it is in English. The article text states that ChatGPT (Chat Generative Pre-trained Transformer) is a chatbot developed by OpenAI and launched on November 30, 2022. It is based on a large language model and allows users to refine and steer a conversation. The article also mentions that by January 2023, it had become the fastest-growing consumer software application in history, with over 100 million users. A sidebar on the left contains a table of contents with sections like "Training", "Features and limitations", "Service", "Reception", and "Use and implications".

The top screenshot shows a Google search for "chatgpt" on the google.co.uk domain. The search results show approximately 1,220,000,000 results in 0.23 seconds. Below the search results, there is a product page for ChatGPT. The page features the OpenAI logo and a description of the chatbot. It states that ChatGPT is a chatbot developed by OpenAI and launched on November 30, 2022. The page also lists the developer as OpenAI and Microsoft Corporation, the initial release date as 30 November 2022, and the license as Proprietary. A table of specifications is provided below the text.

Developer(s)	OpenAI
Initial release	November 30, 2022; 14 months ago
Stable release	February 13, 2024; 6 days ago ^[1]
Written in	Python
Engine	GPT-3.5 (free and paid) GPT-4 (paid only)
Platform	Cloud computing platforms
Type	Chatbot Large language model Generative pre-trained

A GPT-4 Capability Forecasting Challenge

This is a game that tests your ability to predict ("forecast") how well GPT-4 will perform at various types of questions. (In case you've been living under a rock these last few months, GPT-4 is a state-of-the-art "AI" language model that can solve all kinds of tasks.)

Many people speak **very** confidently about what capabilities large language models do and do not have (and sometimes even could or could never have). I get the impression that most people who make such claims don't even know what current models can do. So: put yourself to the test.

How likely do you think GPT-4 is to answer the question below correctly? Enter a number between 0 and 1, where 0 means you think the model has a 0% chance of getting the question right, and 1 means a 100% chance of getting it right. But don't be over-confident!

Question:

What is the capital of France?

Answer:

Paris

How likely is GPT-4 to solve this question correctly?

Are you right? Ask GPT-4!

What can GPT-4 actually do?

<https://nicholas.carlini.com/writing/llm-forecast>

<https://www.ft.com/content/72a27280-9aab-4ba1-86fb-79428006432f>



Hallucinations

Marketing

- Fake references
- Walking on water
- Simply incorrect
- ...

—

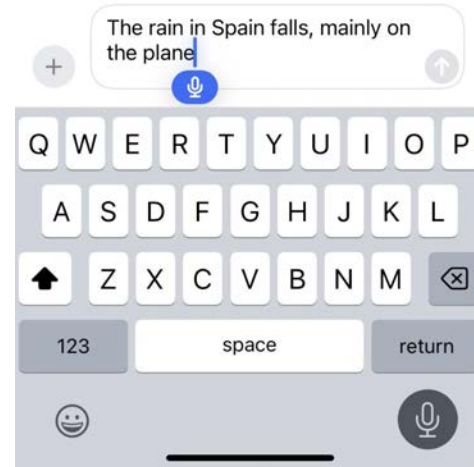
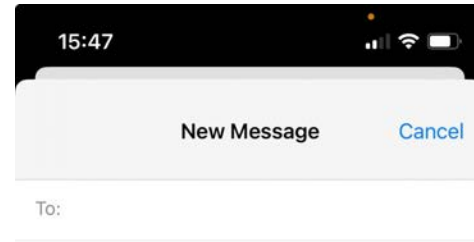
When does it make sense to use AI?

**Where is
the human
in the loop?**

Speech recognition

Cost/benefit is good.

Simple to check and correct.



Text Images Documents Websites

Detect language English Hungarian Polish

The rain in Spain falls mainly on the plain.



44 / 5,000

English Welsh Hungarian

Spanyolországban elsősorban a síkságon esik az eső.



Send feedback



History



Saved



Contribute

Text Images Documents Websites

Detect language English Hungarian Polish

The rain in Spain falls mainly on the plain.



44 / 5,000

English Welsh Hungarian

Spanyolországban elsősorban a síkságon esik az eső.



Send feedback

Text Images Documents Websites

Detect language English Hungarian Polish

Spanyolországban elsősorban a síkságon esik az eső.



51 / 5,000

English Welsh Hungarian

In Spain, it mainly rains on the plains.



Send feedback

Coding and software engineering

```
Users > simonthompson > Desktop > Test.hs > Test
1  -- write a function to reverse a list.
2
3  module Test where
4
5  reverseList :: [a] -> [a]
6  reverseList [] = []
7  reverseList (x:xs) = reverseList xs ++ [x]
8
9  myNub :: Eq a => [a] -> [a]
10
11 myNub [] = []
12 myNub (x:xs) = if member x xs then myNub xs else x : myNub xs
    where
        member _ [] = False
        member y (z:zs) = y == z || member y zs
```

Ln 12, Col 31 Spaces: 4 UTF-8 LF Haskell Spell

```
Users > simonthompson > Desktop > Test.hs > Test
1  -- write a function to reverse a list.
2
3  module Test where
4
5  reverseList :: [a] -> [a]
6  reverseList [] = []
7  reverseList (x:xs) = reverseList xs ++ [x]
8
9  myNub :: Eq a => [a] -> [a]
10
11 myNub [] = []
12 myNub (x:xs) = if member x xs then myNub xs else x : myNub xs
    where
      member _ [] = False
      member y (z:zs) = y == z || member y zs
```

Ln 12, Col 31 Spaces: 4 UTF-8 LF Haskell Spell

CHAT: GL... + ↺ ...

- o If `x` is not even, then it adds `x` once to the result of the recursive call with the rest of the list (`f xs`).

In summary, this function duplicates the even numbers in the list and keeps the odd numbers as they are.

💡 What is an example input and output for function `f`?

Ask Copilot or ty ...

```
Users > simonthompson > Desktop > Test.hs > Test > f
2
3  module Test where
4
5  reverseList :: [a] -> [a]
6  reverseList [] = []
7  reverseList (x:xs) = reverseList xs ++ [x]
8
9  myNub :: Eq a => [a] -> [a]
10
11 myNub [] = []
12 myNub (x:xs) = if member x xs then myNub xs else x : myNub xs
13   where
14     member _ [] = False
15     member y (z:zs) = y == z || member y zs
16
17 f :: Integral a => [a] -> [a]
18 f [] = []
19 f (x:xs) = if x `rem` 2 == 0 then x : x : f xs else x : f xs
```

Ln 17, Col 1 (67 selected) Spaces: 2 UTF-8 LF Haskell ✓ Spell

CHAT: GL... + ↺ ...

o If `x` is not even, then it adds `x` once to the result of the recursive call with the rest of the list (`f xs`).

In summary, this function duplicates the even numbers in the list and keeps the odd numbers as they are.

💡 What is an example input and output for function `f`?

Ask Copilot or ty ...

Test.hs x

Users > simonthompson > Desktop > Test.hs > Test > f

```
2
3  module Test where
4
5  reverseList :: [a] -> [a]
6  reverseList [] = []
7  reverseList (x:xs) = reverseList xs ++ [x]
8
9  myNub :: Eq a => [a] -> [a]
10
11 myNub [] = []
12 myNub (x:xs) = if member x xs then myNub xs else x : myNub xs
13   where
14     member _ [] = False
15     member y (z:zs) = y == z || member y zs
16
17 f :: Integral a => [a] -> [a]
18 f [] = []
19 f (x:xs) = if x `rem` 2 == 0 then x : x : f xs else x : f xs
```

Ln 17, Col 1 (67 selected) Spaces: 2 UTF-8 LF Haskell ✓ Spell

CHAT: GL... + ↺ ...

- If `x` is not even, then it adds `x` once to the result of the recursive call with the rest of the list (`f xs`).

In summary, this function duplicates the even numbers in the list and keeps the odd numbers as they are.

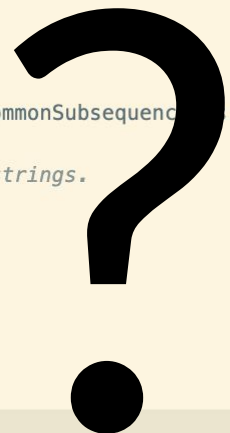
🔗 What is an example input and output for function `f`?

Ask Copilot or ty ...

Test.hs

Users > simonthompson > Desktop > Test.hs > Test > longestContiguousSubsequence

```
35
36
37  -- write a function to find the longest common subsequence of two strings.
38  longestCommonSubsequence :: String -> String -> String
39
40  longestCommonSubsequence [] _ = []
41  longestCommonSubsequence _ [] = []
42  longestCommonSubsequence (x:xs) (y:ys)
43    = if x == y then x : longestCommonSubsequence xs ys else longestCommonSubsequence xs ys
44
45  --write a function to find the longest contiguous subsequence of two strings.
46  longestContiguousSubsequence :: String -> String -> String
47
48  longestContiguousSubsequence [] _ = []
49
50  longestContiguousSubsequence _ [] = []
51
52  longestContiguousSubsequence (x:xs) (y:ys)
53    = if x == y then x : longestContiguousSubsequence xs ys else []
```



>< 0 0 0

Ln 53, Col 7 Spaces: 2 UTF-8 LF Haskell ✓ Spell

Benchmarking coding LLMs

My benchmark for large language models

by **Nicholas Carlini** 2024-02-19

I've just released a [new benchmark for large language models on my GitHub](#). It's a collection of nearly 100 tests I've extracted from my actual conversation history with various LLMs. Among the tests included in the benchmark are tests that ask a model to

- convert a python function to an equivalent-but-faster c function;
- explain the functionality of minified javascript;
- identify the encoding format (in this case, uuencoded) of some data;
- write a parser from a BNF-like grammar;
- convert some english sentences to SQL queries; and,
- write some bash oneliners.

There are two defining features of this benchmark that make it interesting. Most importantly, I've implemented a simple dataflow domain specific language to make it easy for me (or anyone else!) to add new tests that realistically evaluate model capabilities. This DSL allows for specifying both how the question should be asked and also how the answer should be evaluated. Most questions are evaluated by actually running the code the model writes^[a] but the framework supports a bunch of other evaluation methods as well. And then, directly as a result of this, I've written nearly 100 tests for different situations I've actually encountered when working with LLMs as assistants.^[b]

<https://nicholas.carlini.com/writing/2024/my-benchmark-for-large-language-models.html>

SWE-bench

SWE-bench: Can Language Models Resolve Real-World GitHub Issues?

Carlos E. Jimenez et al <https://arxiv.org/abs/2310.06770>

“an evaluation framework consisting of 2,294 problems drawn from real GitHub issues and corresponding PRs ... Given a codebase and ... an issue ... a language model is tasked with editing the codebase to address the issue.”

“state-of-the-art proprietary models and our fine-tuned model SWE-Llama can resolve only the simplest issues. The best-performing model, Claude 2, is able to solve a mere 1.96% of the issues. ”

AI in education

Problems for assessment

Talk don't mark?

Exams?

Group work?

Opportunities for learning

Faster feedback

Tailored learning

Chatbot companion?

Use or not?

Summarising

Speech recognition

Suggestions

...

Complete artifacts: projects, essays, ...

Non-textual or implicit: which code libraries
to use here?

Tacit knowledge

...

AI in the world

Everyone's talking about it ...

... but ...

“Employees turning to GenAI on their own initiative ...”

Too early to see how precisely it is going to make a fundamental difference.



Building the Generative AI advantage:
investing in skills & governance

The content is described as "Partner Content" on this site, however Deloitte and the Financial Times operate separately and independently, and are not in any form of legal partnership.

Employees are turning to Generative AI (GenAI) on their own initiative to make their jobs better. But many are doing so with tools that haven't been approved by their employers



How will AI help?

John Roberts, CEO of AO.com

“Routing of delivery vans”

“Sentiment analysis of reviews”

BBC Today programme 4-11-2024

The Henry Mance Interview Technology sector

[+ Add to myFT](#)

A chatbot that imitates the dead. Is it a good idea?

Jason Rohrer's AI tool can simulate conversations with dead people. Users should decide how far to push it

Opinion **Inside Business**

AI hype has echoes of the telecoms boom and bust

Tech transformation may take years longer than suggested by record share prices and funding targets

JUNE YOON

[+ Add to myFT](#)



Cyber Security

[+ Add to myFT](#)

Is artificial intelligence the solution to cyber security threats?

Advocates of generative AI say the technology will boost efficiency and even outpace criminals

Undercover Economist Undercover Economist

[+ Add to myFT](#)

What the birth of the spreadsheet can teach us about generative AI

Lessons from an earlier software tool that changed the world

FT Magazine Food & Drink

[+ Add to myFT](#)

Looking for an AI-safe job? Try writing about wine

Computational statistics have been wonderful for the wine industry — but the prospect of an 'artificial nose' is an AI boogeyman

A
de

Jason
users

Underc

W
sp
ab

Lessons from an earlier software tool that changed the world

AI hype has echoes of the telecoms boom and bust



Looking at the way spreadsheets are used today certainly suggests a warning. They are endlessly misused by people who are not accountants and are not using the careful error-checking protocols built into accountancy for centuries. Famous economists using Excel simply failed to select the right cells for analysis. An investment bank used the wrong formula in a risk calculation, accidentally doubling the level of allowable risk-taking. Biologists have been typing the names of genes, only to have Excel autocorrect those names into dates.

When a tool is ubiquitous, and convenient, we kludge our way through without really understanding what the tool is doing or why. And that, as a parallel for generative AI, is alarmingly on the nose.

Computational statistics have been wonderful for the wine industry — but the prospect of an 'artificial nose' is an AI boogeyman

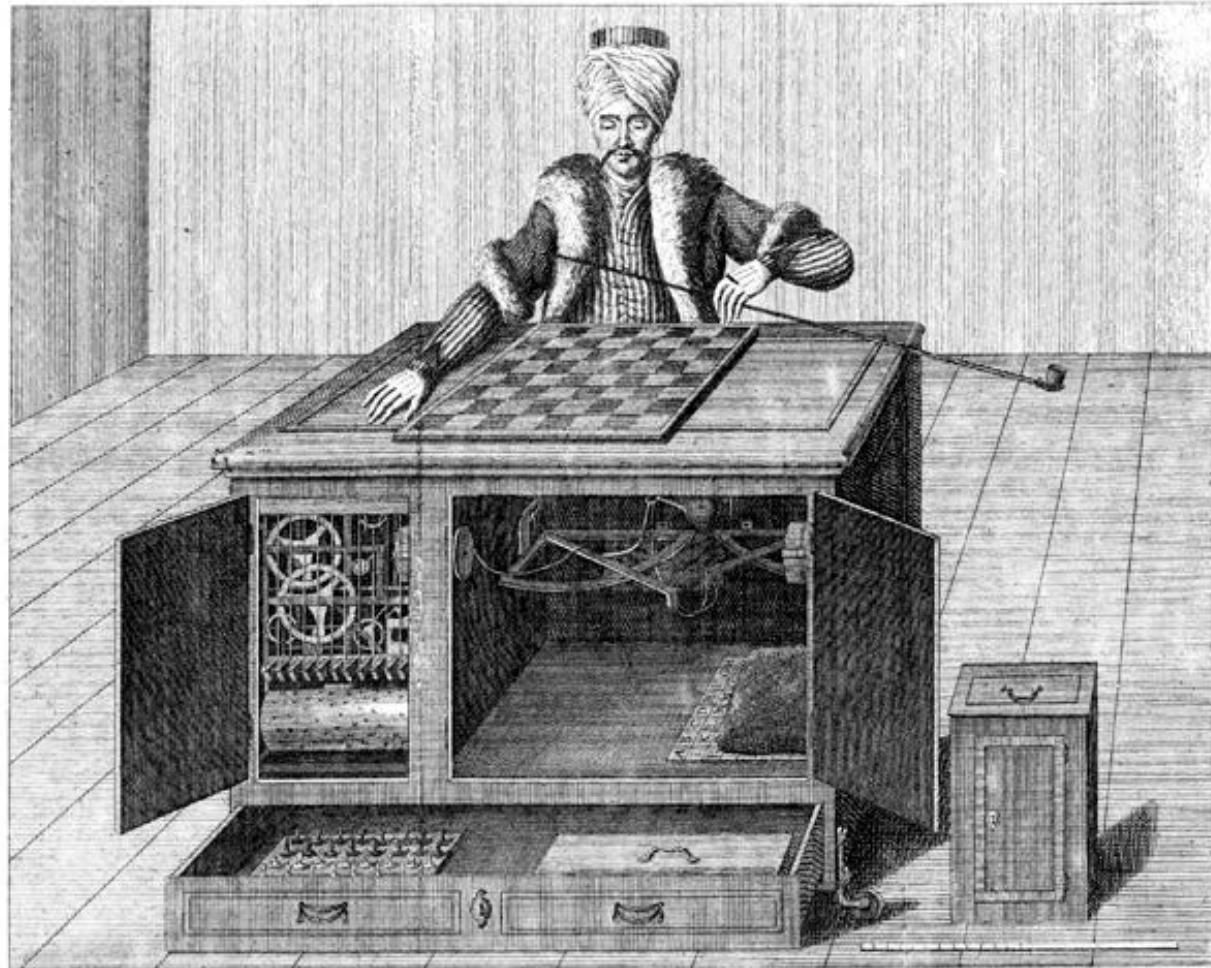
ce

b?

Humans in the loop

Positive for users ...

... less so for those employed to train the LLMs away from objectionable, harmful or incorrect content.



W. de Kempelen del. Die Schachspieler, wie er vor dem Spiel gezeigt wird von ihm. Le joueur d'échecs, tel qu'on le montre avant le jeu, par devant. P. G. Pintz, sc.

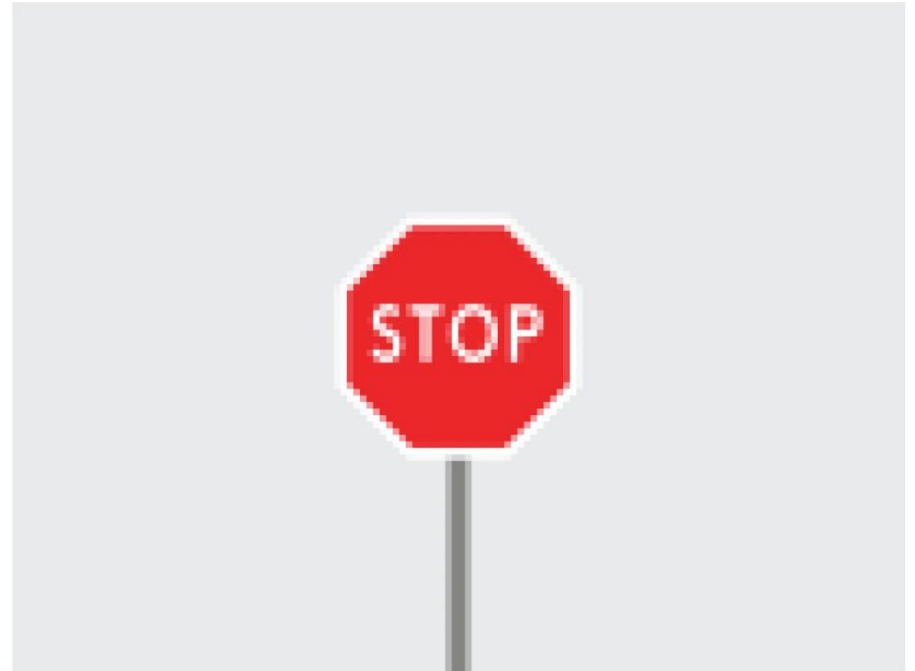
Who controls the platforms?

Stolen Intellectual Property

Deepfakes: privacy and dignity

Vulnerability: AI is just another attack surface.

<https://www.wired.com/story/a-i-has-a-hallucination-problem-t-hats-proving-tough-to-fix/>



The AI Safety Institute is a directorate of the UK Department for Science, Innovation, and Technology.

Rigorous AI research to enable advanced AI governance

[Join us](#)

[View our work](#)

Governments have a key role to play in ensuring advanced AI is safe and beneficial.

The AI Safety Institute is the first state-backed organisation dedicated to advancing this goal.

We are conducting research and building infrastructure to test the safety of advanced AI and to measure its impacts on people and society. We are also working with the wider

The AI Safety Institute is a directorate of the UK Department for Science, Innovation, and Technology.

Rigorous AI research to enable advanced AI governance

[Join us](#)

[View our work](#)

AI Safety Institute will turn into an arms-length government body.

2025 legislation will focus exclusively on ChatGPT-style “frontier” models.

CMA Digital Markets Unit

Governments have a key role to play in ensuring advanced AI is safe and beneficial.

The AI Safety Institute is the first state-backed organisation dedicated to advancing this goal.

We are conducting research and building infrastructure to test the safety of advanced AI and to measure its impacts on people and society. We are also working with the wider



Apocalypse or enshittification?

