

# AI: is it game over for humans?

Monday 4 March 2024 (09:00 – 10:15)

29th Annual International Private Client  
Tax Conference - Navigating the private  
client world in turbulent times

# Chair



**Dr. Niklas Schmidt**  
Wolf Theiss  
Vienna, Austria  
[niklas.schmidt@wolftheiss.com](mailto:niklas.schmidt@wolftheiss.com)

**Wolf Theiss**

# Speakers



**Gary Ashford**  
Harbottle & Lewis LLP  
London, UK  
[gary.ashford@harbottle.com](mailto:gary.ashford@harbottle.com)

Harbottle & Lewis



**Michael Fischer**  
Fischer, Ramp, Buchmann  
Zurich, Switzerland  
[michael.fischer@frb-law.ch](mailto:michael.fischer@frb-law.ch)

**FischerRampBuchmann**  
Rechtsanwälte .avocats .attorneys at law



**Andrea Gallizioli**  
Scarioni Angelucci  
Milan, Italy  
[a.gallizioli@scarioniangelucci.it](mailto:a.gallizioli@scarioniangelucci.it)

**SCARIONI ANGELUCCI**  
STUDIO LEGALE E TRIBUTARIO

# 1.1 Artificial Intelligence (AI)

## 1956 Dartmouth Conference: The Founding Fathers of AI



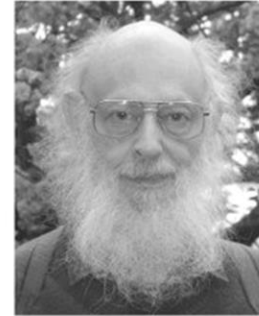
**John McCarthy**



**Marvin Minsky**



**Claude Shannon**



**Ray Solomonoff**



**Alan Newell**



**Herbert Simon**



**Arthur Samuel**



**Oliver Selfridge**

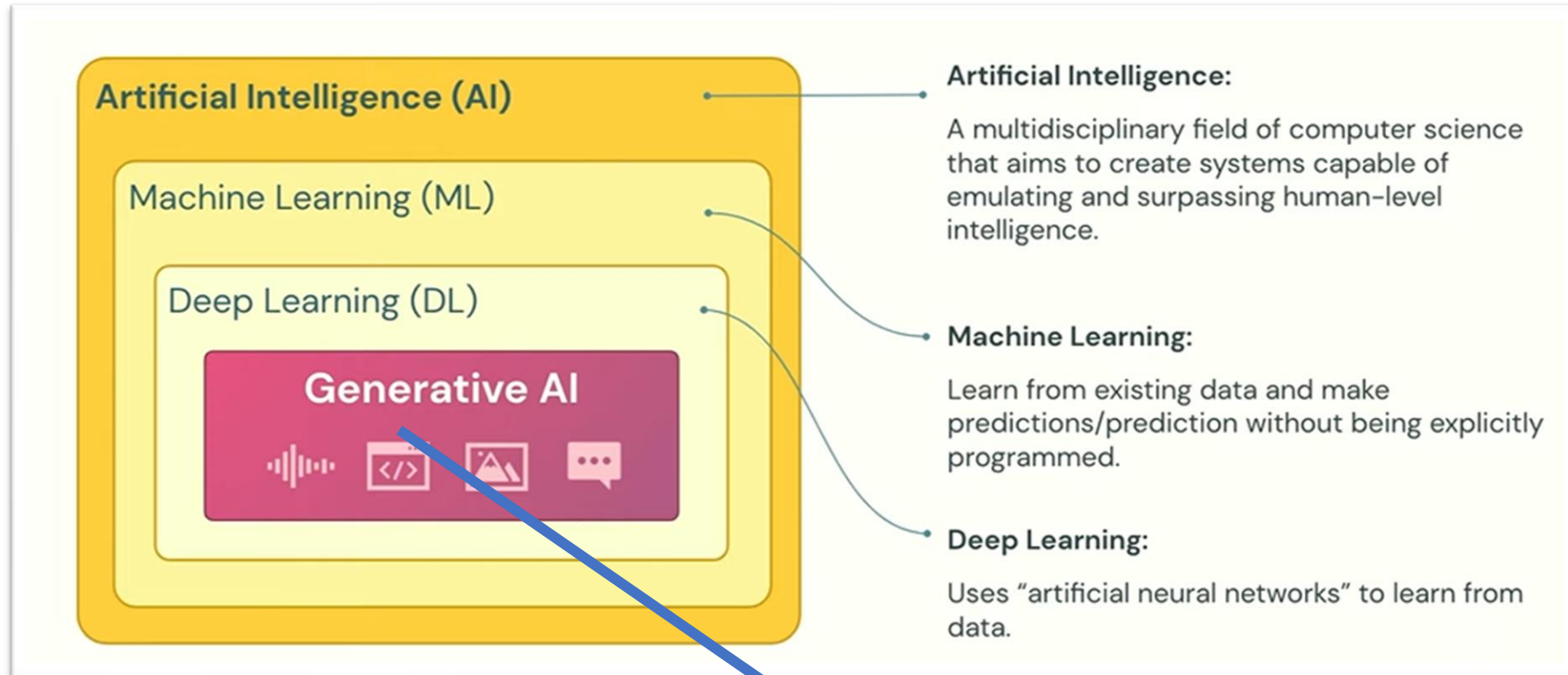


**Nathaniel Rochester**



**Trenchard More**

# 1.1 Artificial Intelligence (AI)



- Large Language Models generate texts
- Diffusion Models generate images

(both are referred to as foundation models)

# 1.1 Artificial Intelligence (AI)

In 2022, OpenAI released two important products based on its GPT (Generative Pre-trained Transformer) technology, and unleashed an arms race:

- **ChatGPT (launched in November 2022)**, a “Large Language Model” that can answer complex questions. It has processed more text than any human can read in a lifetime
- **Dall-E (launched in September 2022)**, a “Diffusion Model” that can generate digital images from natural language. It has seen millions of images.



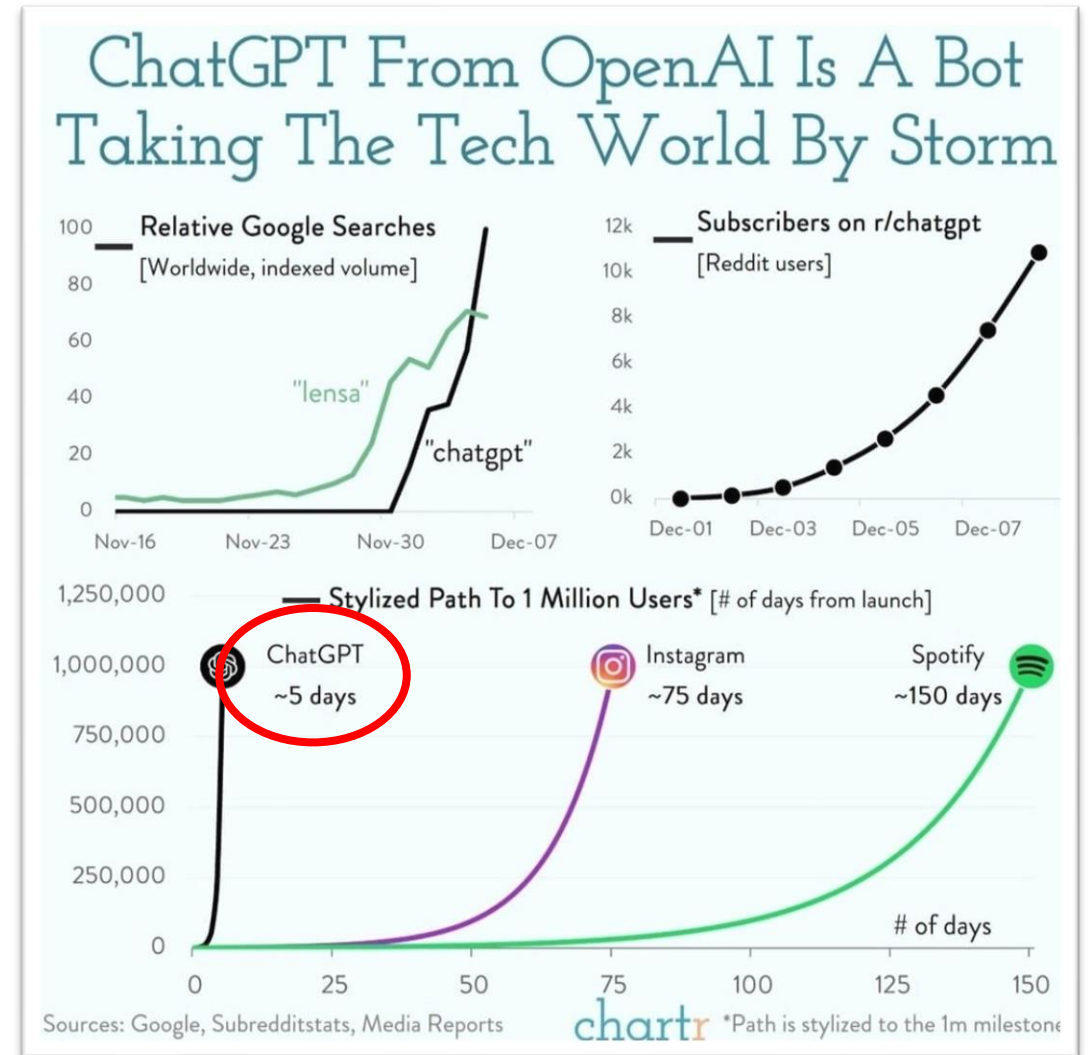
**ars TECHNICA**

YOU ARE HERE —

## AI fever turns Anguilla's “.ai” domain into a digital gold mine

Tiny island country could rake in 10% of its GDP in domain sales this year.

BENJ EDWARDS - 8/31/2023, 11:58 PM

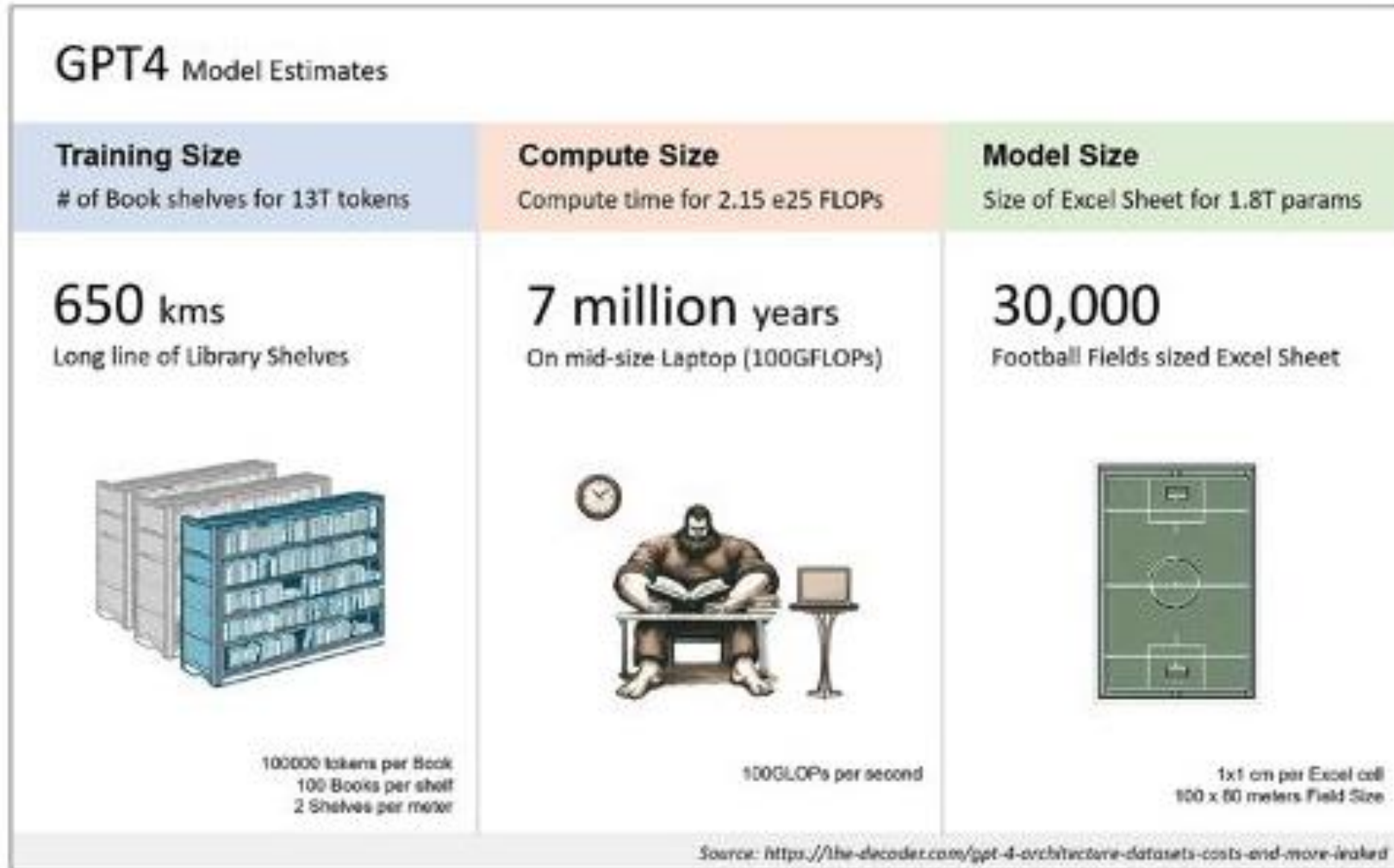


... and 200m users in half a year!

## 1.2 Large Language Models (LLMs)

- Statistical next-word predictors
- Like a parrot that is listening in on a conversation and that blindly repeats what it has heard, but:
  - this parrot has heard all conversations worldwide
  - this parrot was regularly corrected when it said stupid things

# 1.2 Large Language Models (LLMs)





# 1.2 Large Language Models (LLMs)

- **ChatGPT** by OpenAI (startup): <https://chat.openai.com> (free/commercial; based on GPT-3.5/4; app available)
- **Copilot** by Microsoft (lucky investor): <https://copilot.microsoft.com> (free/commercial; based on GPT-4; app available)
- **Gemini** by Google (the OG inventor who is now catching up): <https://gemini.google.com> (free/commercial; uses Gemini Pro 1.0/Ultra 1.0; Gemini 1.5 Pro has a context window of more than 1m tokens and is twice as fast as GPT-4; biases)
- **Claude 2** by Anthropic (ex-OpenAI engineers repeating history without noticing it): <https://claude.ai> (currently restricted to UK/US; large context window of 200k tokens; “Constitutional AI” for “helpful, harmless, honest” responses)
- **LLaMA 2** by Meta (Big Tech company with big ambitions): <https://llama.meta.com> (open source)
- **Mixtral 8X7B** by Mistral (French (!) startup with researchers from Google’s Deepmind and from Meta): <https://mistral.ai> (open source; easily beats GPT-3.5, but is still below GPT-4’s capabilities)
- **Cohere** by Cohere (Canadian startup, including an author of “Attention is All You Need”): <https://cohere.com>
- **Pi** by Inflection AI (engineers from many different AI shops): <https://pi.ai> (free)
- **Grok** by xAI (one of Elon Musk’s many companies): <https://x.ai> (works within Twitter/X on a paid plan, has a witty, rebellious personality and is not woke)



# 1.2 Large Language Models (LLMs)

Mark Zuckerberg plans on acquiring **350,000 Nvidia H100 GPUs** to help Meta build a next-generation AI that possesses human-like intelligence.

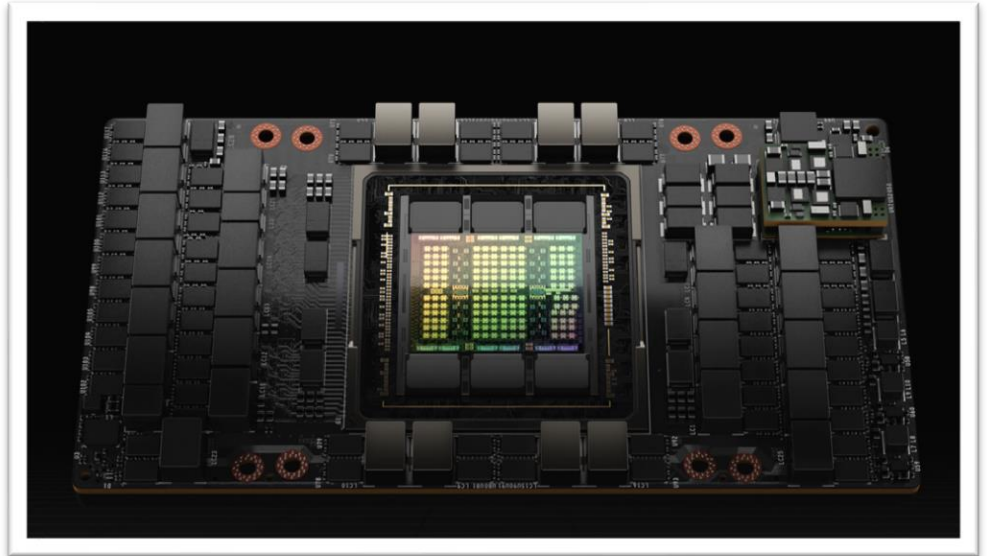
Zuckerberg mentioned the figure today as he announced his company's long-term effort to develop an artificial general intelligence (AGI), or an AI that can learn and be used to perform a variety of tasks.

Meta's CEO envisions the AGI powering a wave of cutting-edge services and devices, such as more powerful digital assistants and augmented reality glasses. "Building the best AI assistants, AIs for creators, AIs for businesses and more, that needs advances in every area of AI," he said in a video on Instagram.

But to get there, Meta is going to need Nvidia's H100, an enterprise GPU that's adept at training large language models. "We're building an absolutely massive amount of infrastructure to support this," Zuckerberg said. "By the end of this year, we're going to have around 350,000 Nvidia H100s. Or around **600,000 H100 equivalents** of compute if you include other GPUs."

The 350,000 number is staggering, and it'll also cost Meta a small fortune to acquire. **Each H100 can cost around \$30,000**, meaning **Zuckerberg's company needs to pay an estimated \$10.5 billion** just to buy the computing power, not to mention paying all the electricity costs.

The statement also provides a glimpse into how far the leading tech companies will go to develop new AI models. Other tech giants, including Microsoft, Google, and Amazon, also likely bought between 50,000 to 150,000 Nvidia H100s last year, according to research firm Omdia.



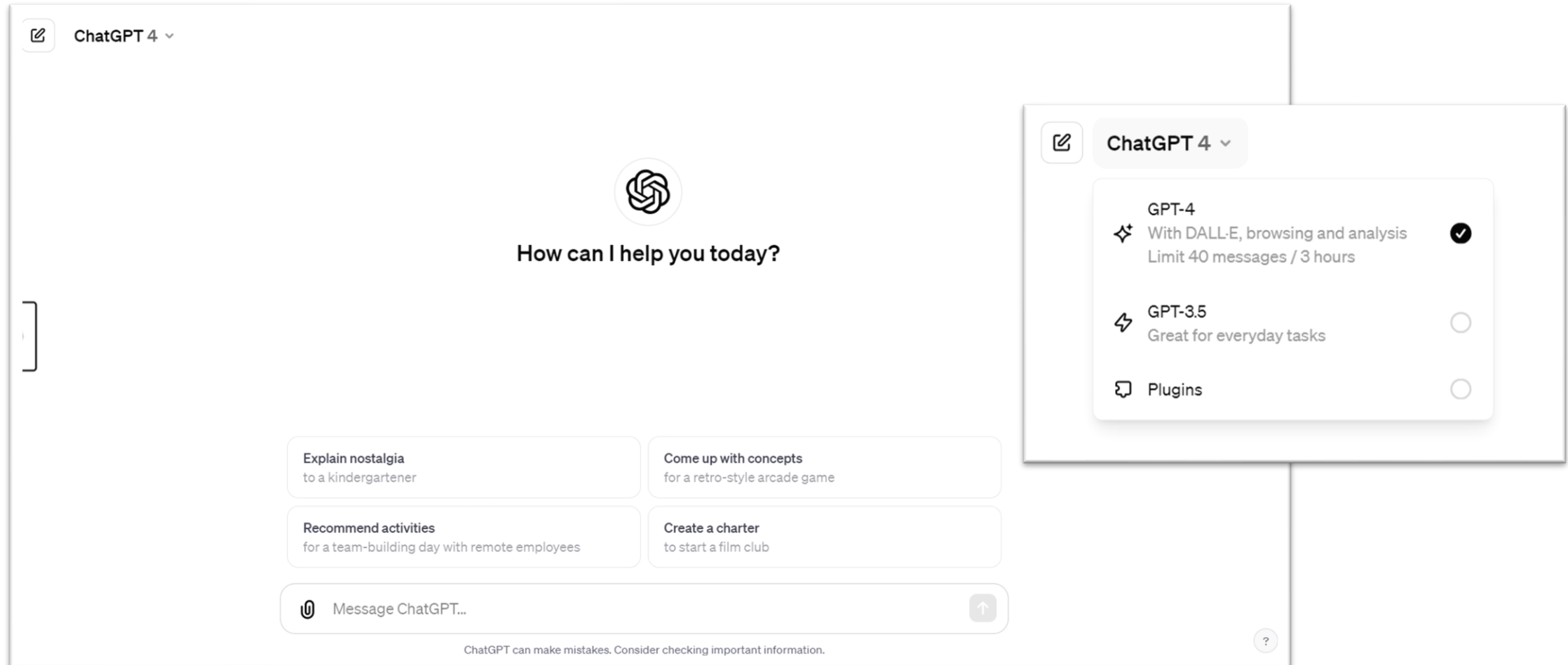
## NVIDIA H100 Tensor Core GPU

Unprecedented performance, scalability, and security for every data center.

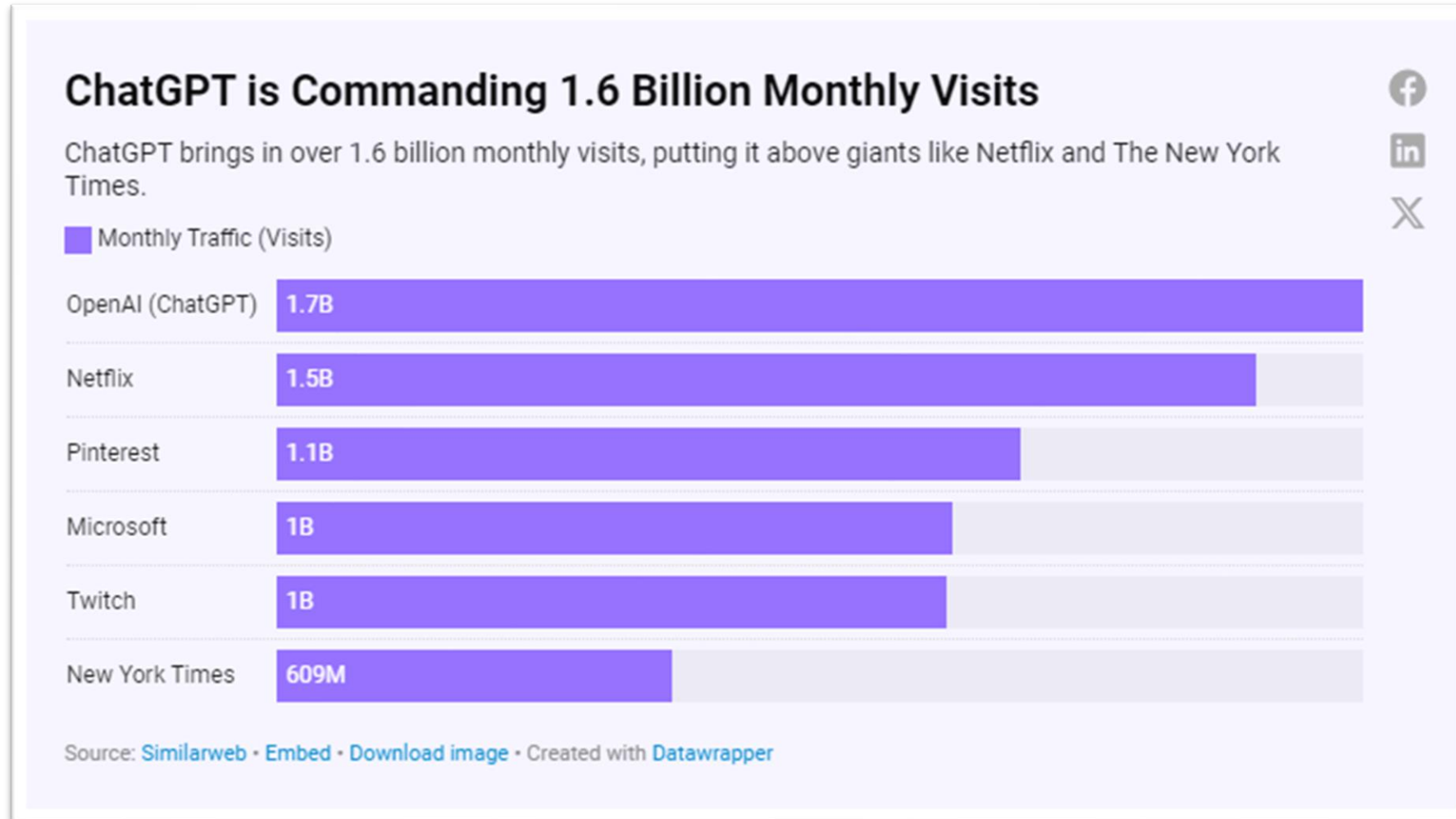
# 1.3 ChatGPT

Understand:

conversation starter, prompt, chat completion, file upload, chat history, conversational context, context window, clear chat



# 1.3 ChatGPT



## 1.3 ChatGPT

FINANCE · ENTREPRENEURS

**If you're thinking about writing your own pitch deck, think again. GPT-4 outshines humans in securing funding and impressing business owners, survey shows**

BY ORIANNA ROSA ROYLE

June 9, 2023 at 6:22 PM GMT+2

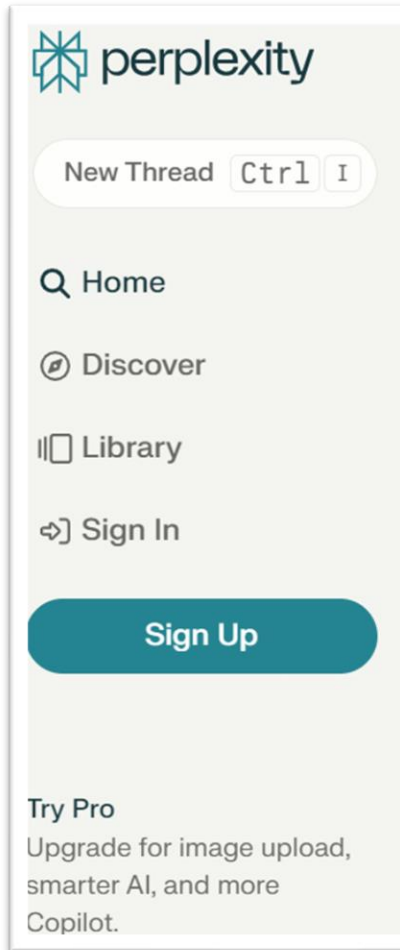


mostly syntactically + semantically correct

## 1.4 Potential use cases

- Help you to write drafts of anything
- Help you to quickly summarize texts
- Help you to quickly extract information from texts
- Help you to make your writing better
- Help you to unblock yourself
- Help you to come up with creative ideas
- Help you to check your arguments and logic
- Help you to find dangerous provisions in contracts
- Help you to check for missing items in your drafts
- Help you to get certain information better than Google

# 1.5 Search engine replacement



perplexity

New Thread **Ctrl** **I**

Home

Discover

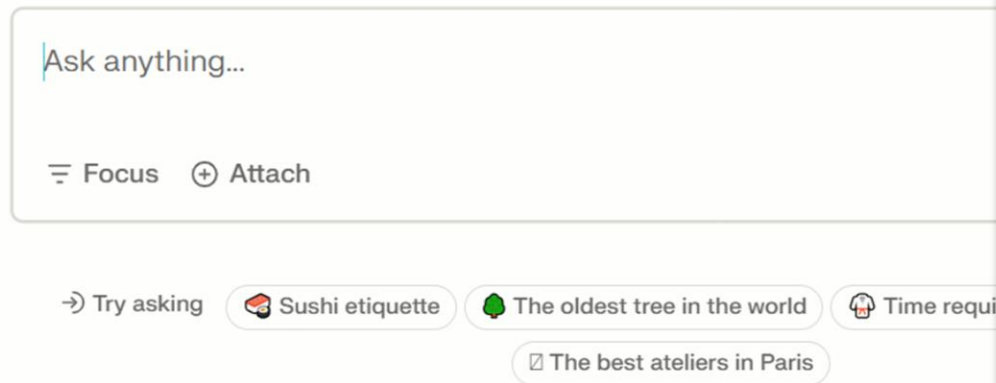
Library

Sign In

**Sign Up**

Try Pro  
Upgrade for image upload,  
smarter AI, and more  
Copilot.

## Where knowledge begins

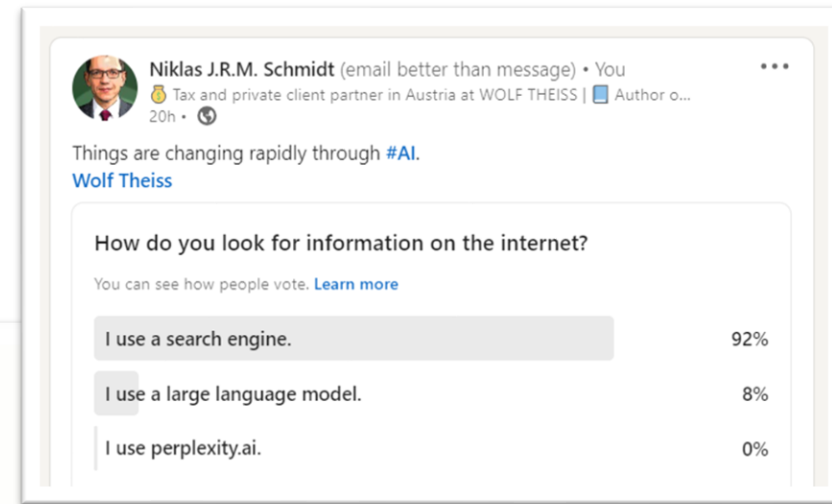


Ask anything...

Focus Attach

Try asking

- Sushi etiquette
- The oldest tree in the world
- The best ateliers in Paris
- Time requi



Niklas J.R.M. Schmidt (email better than message) • You  
Tax and private client partner in Austria at WOLF THEISS | Author o...  
20h

Things are changing rapidly through #AI.  
Wolf Theiss

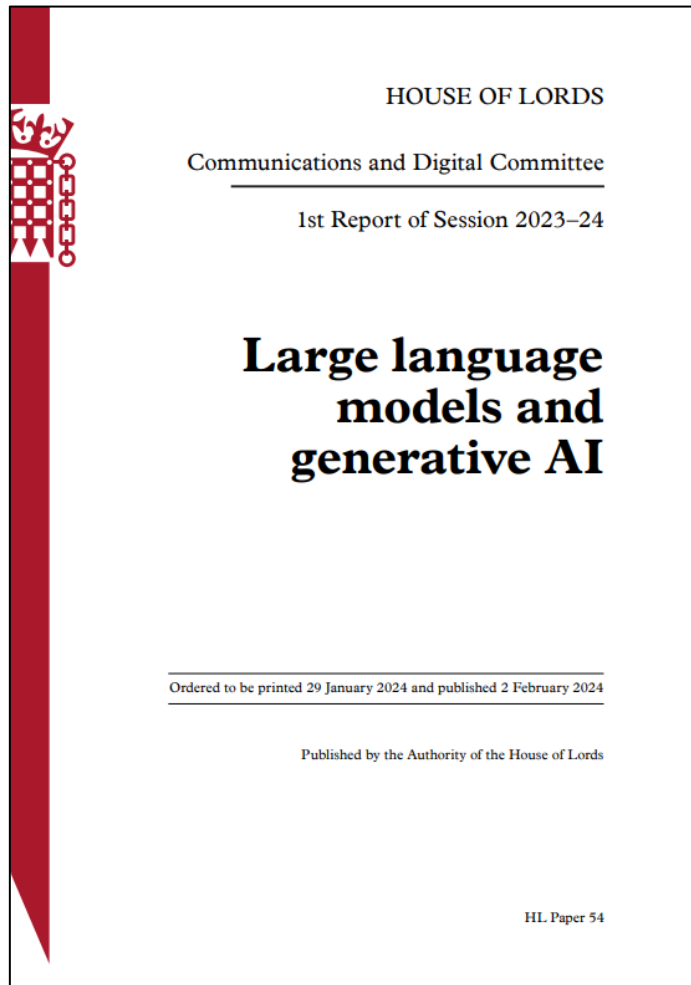
How do you look for information on the internet?

You can see how people vote. [Learn more](#)

I use a search engine.	92%
I use a large language model.	8%
I use perplexity.ai.	0%



## 2.1 House of Lords report



- 93 page House of Lords report, published on 2 February 2024
- Inquiry examined trends over next three years (!), with recommendations addressed to UK government
- Committee is “optimistic about this new technology”
- “deeper concerns about the Government’s commitment to fair play around copyright”



# 2.1 House of Lords report

## *Communications and Digital Committee*

The Communications and Digital Committee is appointed by the House of Lords in each session “to consider the media, digital and the creative industries and highlight areas of concern to Parliament and the public”.

## *Membership*

The Members of the Communications and Digital Committee are:

Baroness Featherstone

Lord Foster of Bath

Baroness Fraser of Craigmaddie

Lord Griffiths of Burry Port

Lord Hall of Birkenhead

Baroness Harding of Winscombe

Baroness Healy of Primrose Hill

Lord Kamall

The Lord Bishop of Leeds

Lord Lipsey

Baroness Stowell of Beeston (Chair)

Baroness Wheatcroft

Lord Young of Norwood Green

Baroness Lynne Featherstone, 1951  
Lord Foster of Bath, 1947  
Baroness Fraser of Craigmaddie, 1968  
Lord Griffiths of Burry Port, 1942  
Lord Hall of Birkenhead, 1951  
Baroness Harding of Winscombe, 1967  
Baroness Healy of Primrose Hill, 1955

Lord Kamall, 1967  
Lord Bishop of Leeds, 1957  
Lord Lipsey, 1948  
Baroness Stowell of Beeston, 1967 (Chair)  
Baroness Wheatcroft, 1951  
Lord Young of Norwood Green, 1942



[www.teatravellerssocietea.com](http://www.teatravellerssocietea.com)



Average age: 71

## 2.1 House of Lords report



*...is A.I. is too stupid, this A.I. is too biased, and this A.I. ..."*

*The New Yorker*

The Goldilocks problem:

“in the right hands, LLMs may drive major boosts in productivity + ground-breaking scientific insights”

vs.

“in the wrong hands LLMs may make malicious activities easier and may lay groundwork for qualitatively new risks”

## 2.1 House of Lords report

### North Korean hackers use AI for more sophisticated scams

Cyber criminals backed by Pyongyang turn to artificial intelligence as they try to secure funds



Money raised by North Korea's criminal cyber operations helps to fund the country's ballistic missile and nuclear programmes, according to a UN panel of experts monitoring implementation of international sanctions © KCNA via KNS/AFP/Getty Images

# 2.1 House of Lords report

## Getting balance right between innovation and risk

“Red Flag Act” of 1865



Seat belts – invented in 1885, mandatory only since 1983

Seat belts help protect you from death and serious injury in car accidents. From January 31st, it's the law to wear a seat belt. The wearing of seat belts is compulsory.

**THE SEAT BELT LAW**  
By law drivers and front seat passengers in most vehicles must wear their seat belts. As a result, the number of car deaths registered after 1983 has fallen significantly.

When you wear a seat belt, you must wear an approved child restraint designed for their weight. Child restraint seats must also meet the same safety standards as approved motor vehicles to use in a motor car.

When you wear a seat belt, you must wear an approved child restraint designed for their weight. Child restraint seats must also meet the same safety standards as approved motor vehicles to use in a motor car.

**Clunk-Click becomes law from January 31st.**  
**Until then, it's perfectly legal to smash your face in.**

When the belt will be more effective. It's the safest piece of all for children to travel in the car. It's designed by an approved motor car.

**THE SEAT BELT LAW**  
A law people will be happy to see. It's compulsory. But you will receive a penalty if you don't wear one.

Children who will not have to wear a seat belt. It's compulsory. It's a rule which will be enforced. It's a rule which will be enforced. It's a rule which will be enforced.

Children who will not have to wear a seat belt. It's compulsory. It's a rule which will be enforced. It's a rule which will be enforced. It's a rule which will be enforced.

Children who will not have to wear a seat belt. It's compulsory. It's a rule which will be enforced. It's a rule which will be enforced. It's a rule which will be enforced.

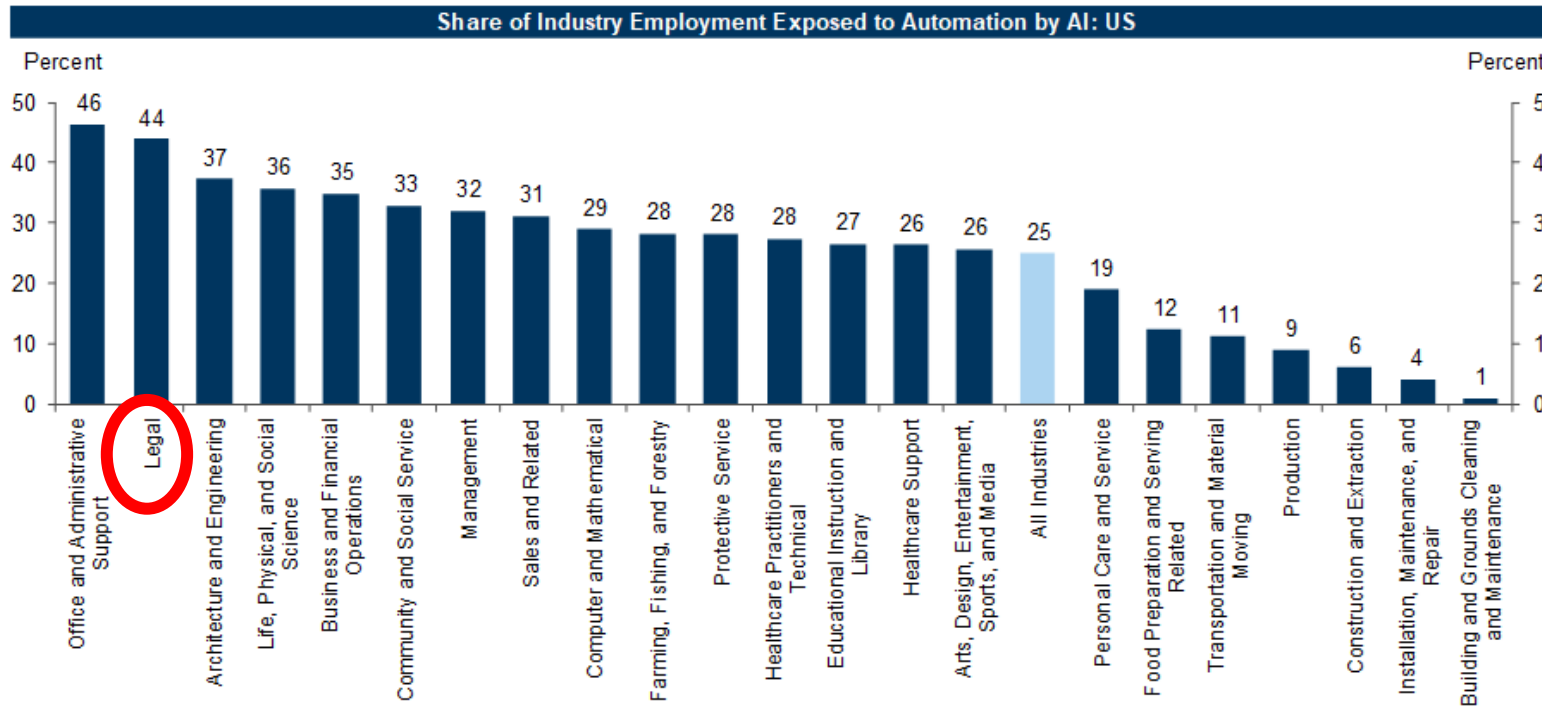
**Clunk-Click. The law of survival.**

# 2.1 House of Lords report

## Trends identified in the report

- LLMs will continue to hallucinate, exhibit bias, regurgitate private data, struggle with multi-step tasks – improvements highly likely within three years
- Level of market competition remains uncertain – multi-billion pound race to dominate the market is ongoing
- LLMs will have impact comparable to the invention of the internet
- Level of market competition remains uncertain – multi-billion pound race to dominate the market is ongoing
- “Risk of regulatory capture” – big tech firms reportedly funding salaries of US Congress staff working on AI policy
- “Labour market impacts remain uncertain”

# 2.1 House of Lords report



Goldman Sachs estimate: “one-fourth of current work tasks could be automated by AI in the US, with particularly high exposures in administrative (46%) and legal (44%) professions.

Goldman Sachs, *The Potentially Large Effects of Artificial Intelligence on Economic Growth*, 26 March 2023

## 2.1 House of Lords report

- Government cannot match big tech spending
- Current security standards unlikely to withstand attacks from sophisticated threat actors
- Cost of disinformation is plunging, example:
  - Estimated cost for online campaign targeting the US 2016 election exceeded USD 10,000,000
  - Today, same could be done for approx. USD 1,000

## 2.1 House of Lords report

Recommendation regarding approach to regulation:

*“Extensive primary legislation aimed solely at LLMs is not currently appropriate: the technology is too new, the uncertainties too high and the risk of inadvertently stifling innovation too great. [...] Broader legislation on AI governance may emerge in future.”*

Report, p. 54



## 2.2 EU's proposed AI Act



## 2.2 EU's proposed AI Act

### Timeline



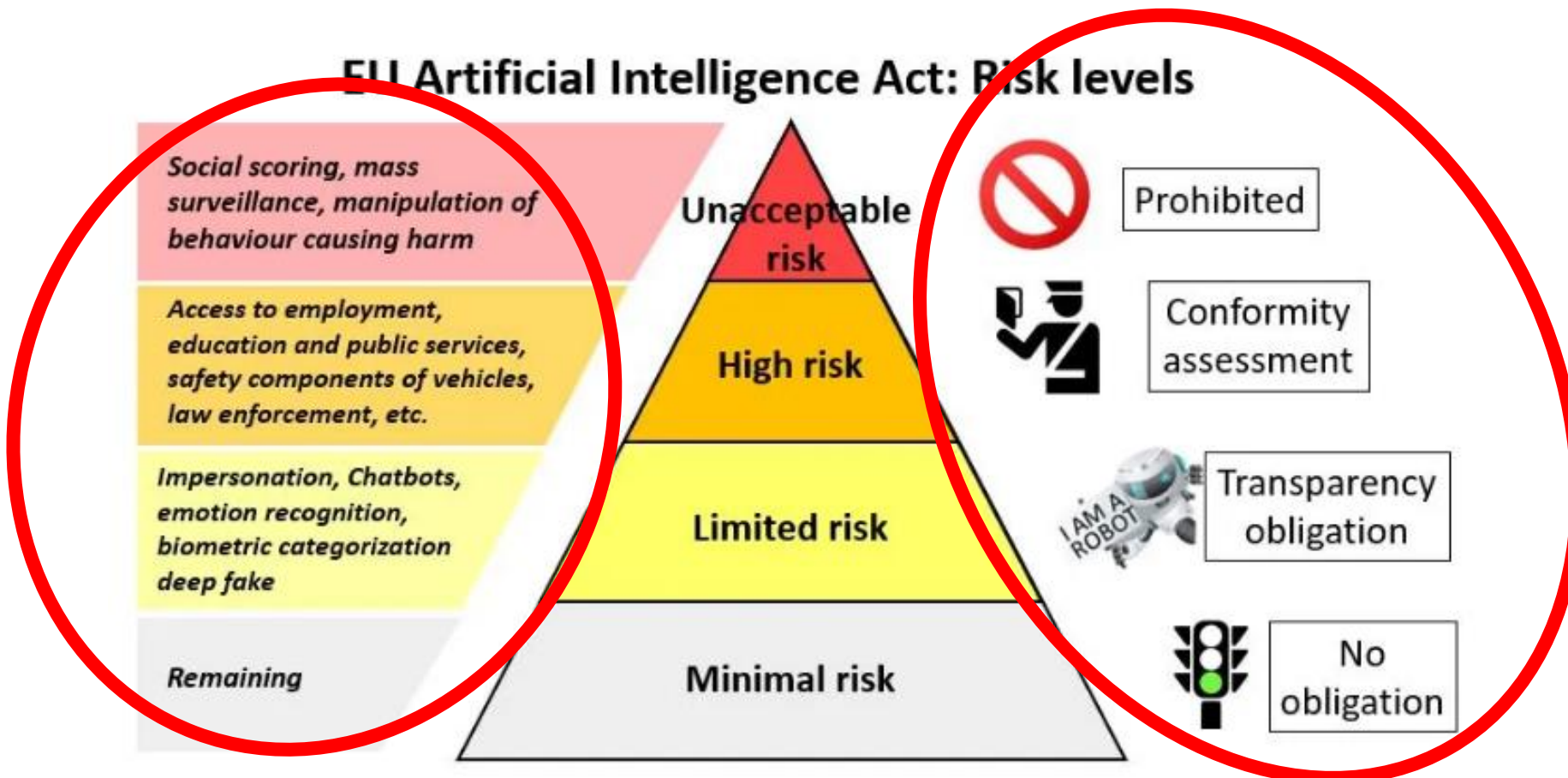
- Text of provisional agreement published on 2 February 2024
  - 93 pages of introduction + 179 pages of regulation
- Agreed text will need to be formally adopted by Parliament and Council to become law
  - Entry into force in 2025/2026 (?)

## 2.2 EU's proposed AI Act

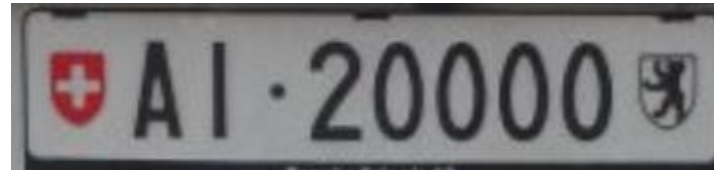
Very (very) broadly: risk-based approach, illustrated by “risk pyramid”

Penalties for non-compliance:

up to the higher of  
(a) EUR 35m or  
(b) 7% of global turnover



## 2.3 Switzerland



- Swiss government has commissioned an overview of possible regulatory approaches by the end of 2024
- Involvement of various governmental agencies, “Plateforme Tripartite” (coordinative body directed at inclusion of wide range of interests)
- Aim is to identify regulatory approaches compatible with the EU’s forthcoming AI Act, the Council of Europe’s AI Convention, and fundamental rights
- Target date for enactment is 2026 (?)

# 3.1 AI's appetite for data and associated risks

- *“AI scenarios are mainly driven and determined by the availability and evaluation of data. In other words, AI goes hand in hand with what may be referred to as an enormous appetite for data”*  
(Boris P. Paal)
- The adoption of AI in the legal profession brings both opportunities and challenges, among which it is crucial to carefully consider and address the potential risks associated with the use of data (personal and non-personal), including:
  - Compliance with GDPR
  - Compliance with professional secrecy rules
  - Exposure of clients to tax risks
  - Fiduciary relationship with clients

## 3.2 Compliance with GDPR

- The GDPR does not include specific rules applicable in the context of AI, and might be applied only by way of interpretation.
- There is a general consensus that the regulation is inadequate, and risks creating uncertainties and conducts contrary to the purpose of the same law:
  - “Personal data shall be processed [...] in a **transparent manner** in relation to the data subject [...] collected for **specified, explicit and legitimate purposes**” (Art. 5 of GDPR): in an AI scenario, is it possible to state and substantiate the specific purposes for any given data analysis in advance “*in a transparent, intelligible, and easily accessible form*”?
  - “Personal data shall be adequate, **relevant and limited to what is necessary** in relation to the purposes for which they are processed (‘data minimisation’)” (Art. 5 of GDPR): AI is based and requires the collection of large amounts of data.
  - Under GDPR, data processing is only lawful to the extent that “*specific and informed*” consent is given. In addition, the consent can be withdrawn: is this viable in a self-learning and autonomous AI systems?
- The EU’s AI Act should try to address this matter.

## 3.2 Compliance with GDPR

- In December 2022, the Italian GDPR authority launched an investigation into OpenAI and the compliance of its operations with the GDPR.
- The authority identified significant violations in ChatGPT's handling of personal data, specifically:
  - Transparency breaches: Italian users were not informed about how their data was processed
  - Unlawful data processing: OpenAI lacked a valid legal basis for collecting and using personal data to train ChatGPT
  - Data inaccuracy: ChatGPT frequently processed inaccurate personal data
- The authority issued a provisional order in March 2023 limiting ChatGPT's operations in Italy.
- The order was suspended in April 2023, conditional upon the implementation of the following measures:
  - Provide clear and comprehensive information to users about how their personal data is processed
  - Obtain valid consent for the collection and processing of personal data
  - Ensure the accuracy of data processing

## 3.3 Compliance with professional secrecy rules

- Professional secrecy is enshrined both in the law for the generality of legal professions and regulated by internal regulations.
- The scope is much more extended compared to GDPR (*“the rigorous observance of professional secrecy and the utmost confidentiality regarding **facts and circumstances learned in any way in the activity** of representation and assistance in court, as well as in the performance of legal advice and out-of-court assistance and in any case for professional reasons”*).
- The breach might represent a criminal offence (*“Whoever, having knowledge, because of his role or office, or of his profession or art, of a secret discloses it, without just cause, or employs it for his own or others' profit, shall be punished, if the fact may cause harm, with imprisonment of up to one year [...]”*).
- Utilizing third-party AI services with client data raises concerns about the unintentional disclosure of confidential information and might (easily) lead to the violation of professional secrecy.



## 3.4 Exposure of clients to tax risks

- Using AI for tax risk analysis is one of the cornerstones of the Italian tax reform (under implementation) , enabling the Italian Revenue Agency to do the following:
  - processing large volumes of data from various sources
  - identifying potential tax evaders and fraudsters, optimising the control activities of the Revenue Agency
  - developing predictive models to prevent illicit behaviour, allowing for targeted verification interventions
- AI might have a role also in the context of ruling procedures.
- There is a general concern about risks associated with the Revenue Agency using AI, including classification errors (AI could mistakenly classify a taxpayer as high-risk), lack of transparency and general privacy concerns.
- At the same time, there are measures trying to ensure:
  - compliance with the GDPR
  - a definition, by ministerial decree, of specific protection profiles in line with the GDPR
  - the consultation with the Italian GDPR authority to identify the necessary measures to safeguard citizens' rights

## 3.5 Fiduciary relationship with clients

- The average private client is obsessed with confidentiality and management of personal data (*e.g.*, request for “ad hoc” NDAs is increasing) and reluctant to provide information (*e.g.*, for purposes of KYC and AML procedures).
- In general, clients are probably more relaxed, and thoughtless behaviour is not uncommon (*e.g.*, uploading pictures and information on social networks, exchanging information on WhatsApp public groups).
- Attention to the confidentiality and management of the data of a client is at the root of our profession and often goes well beyond what is required by regulation and professional bodies.

## 3.6 What can be done to address the problem?

- Strong internal policies concerning (i) the selection of AI tools to be used and (ii) the mode of usage of those tools
- Communication with (*e.g.*, in engagement letters) and raising awareness among clients
- GDPR knowledge and the possible evolutions
- Knowledge of AI technology and ability to evaluate how AI systems work

## 3.7 Evaluation of AI tools

- First layer (AI tools, *e.g.*, Harvey)
  - Is data being stored or not?
  - Which data: personal data and other content?
  - For how long is it being stored?
  - In which country is it being stored (EU, non-EU, multiple countries)?
  - What happens in the case of data leaks?
- Second layer (underlying AI technology, *e.g.*, GPT-4): same analysis as before

## 3.8 Is data being stored?



### 3. Texts and Translations - DeepL Translator (free version)

When using our translation service, please only enter texts that you wish to transfer to our servers. The transmission of these texts is necessary in order for us to provide the translation and offer you our service. We process your texts, the documents you upload and their translations for a limited period of time to train and improve our neural networks and translation algorithms. This also applies to corrections you make to our translation suggestions. The corrections are forwarded to our servers to check them for accuracy and, if necessary, to update the translated text according to your changes. If you use the glossary feature and enter certain pairs of terms there, this data is only stored locally and is not forwarded to our servers. Therefore, you cannot use your glossary entries in another browser or on another device.

Please note that according to its Terms and Conditions, you may not use DeepL Translator for the translation of texts containing personal data of any kind. The translation of personal data is only possible as part of a DeepL Pro subscription (please see section 5 [↗](#)).

### 5. Texts and Translations - DeepL Pro

When using our DeepL Pro subscription to translate texts, the texts or documents you submit will not be permanently stored and will only be kept temporarily to the extent necessary for the production and transmission of the translation. After complete performance of the contractually agreed services all submitted texts or documents and their translations will be deleted. When using DeepL Pro, your texts will not be used to improve the quality of our services. For further information on the processing of your data within the DeepL Pro subscription, please refer to section 6 of this privacy policy and our DeepL Pro Terms and Conditions [↗](#).

Please note that using DeepL Pro for the translation of texts containing personal data of any kind is only permitted if there is a justification for this under data protection law. Therefore, our T&C provide for the conclusion of a data processing agreement (see section 8.1.3 of DeepL Pro Terms and Conditions [↗](#)). To enter into such an agreement, please contact [sales\(at\)deepl.com](mailto:sales(at)deepl.com).

# 3.8 Is data being stored?



## 3. Content

**3.1 Customer Content.** You and End Users may provide input to the Services (“Input”), and receive output from the Services based on the Input (“Output”). We call Input and Output together “Customer Content.” As between you and OpenAI, and to the extent permitted by applicable law, you (a) retain all ownership rights in Input and (b) own all Output. We hereby assign to you all our right, title, and interest, if any, in and to Output.

**3.2 Our Obligations for Customer Content.** We will process and store Customer Content in accordance with our Enterprise privacy commitments. We will only use Customer Content as necessary to provide you with the Services, comply with applicable law, and enforce OpenAI Policies. We will not use Customer Content to develop or improve the Services.

## 3.9 For how long is data being stored?

We'll retain your Personal Data for only as long as we need in order to provide our Service to you, or for other legitimate business purposes such as resolving disputes, safety and security reasons, or complying with our legal obligations. How long we retain Personal Data will depend on a number of factors, such as:

- Our purpose for processing the data (such as whether we need to retain the data to provide our Services);
- The amount, nature, and sensitivity of the data;
- The potential risk of harm from unauthorized use or disclosure of the data;
- Any legal requirements that we are subject to.



3.1.3 DeepL will only access the Content and/or the Processed Content in the event and to the extent required to carry out a diagnosis and to solve technical issues which may eventually compromise the availability of the Products. To the extent required for the purposes mentioned above and in derogation from Section 3.1.2 DeepL may, in exceptional cases, automatically store Content and/or Processed Content for a maximum period of 72 hours in case certain error patterns occur during the processing of the translation request. The Content and/or the Processed Content will be stored in an encrypted form for the duration of the debugging process and will be automatically deleted afterwards.

## 3.10 In which country is data being stored?

We will transfer your Personal Data to recipients outside of the EEA, Switzerland and the UK for the purposes described in this Privacy Policy



By using our Services, you understand and acknowledge that your Personal Data will be processed and stored in our facilities and servers in the United States and may be disclosed to our service providers and affiliates in other jurisdictions.



forwarded to our servers. Therefore, you cannot use your glossary entries in another browser or on another device. When using additional functionalities (e.g., summaries), your texts and files may be transferred to servers of external third parties. These servers are located exclusively in the European Economic Area (EEA).



# 3.11 What does AI think about risks for advisors?



**You**

As a private client advisor do I incur any risk to my clients by using Artificial intelligence tools?



**ChatGPT**

Using artificial intelligence (AI) tools as a private client advisor can potentially present certain risks to your clients. These risks include:

1. **Accuracy and Reliability:** AI tools, including chatbots like ChatGPT, may not always provide accurate or reliable information. Relying solely on AI-generated insights without proper validation or human oversight could lead to incorrect advice or decisions that may adversely affect your clients.
2. **Data Privacy and Security:** Interacting with AI tools often involves sharing client information or sensitive data. There's a risk that this information could be intercepted, accessed by unauthorized parties, or compromised in some way, potentially violating client privacy and confidentiality. Ensuring robust data encryption and implementing stringent data security measures can help mitigate these risks.
3. **Ethical Considerations:** There are ethical implications associated with using AI technology in client advisory roles. It's essential to maintain transparency with clients about the use of AI tools, obtain informed consent, and adhere to professional standards of conduct and integrity. Additionally, ensuring that AI tools are used responsibly and ethically, and that clients' best interests are prioritized, is crucial.

# 4.1 Felicity Harber v. The Commissioners for HMRC

- Background and Facts
- Penalty for Failure to Notify
- Appeal for Reasonable Excuse
- Taxpayer put forward chatbot generated case law (cases didn't exist!)
- “Hallucinations”
- Case dismissed!

## 4.2 What are hallucinations?

- AI hallucination is a phenomenon wherein a large language model (LLM) – often a generative AI chatbot or computer vision tool – perceives patterns or objects that are nonexistent or imperceptible to human observers, creating outputs that are nonsensical or altogether inaccurate.
- Generally, if a user makes a request of a generative AI tool, they desire an output that appropriately addresses the prompt (*i.e.*, a correct answer to a question). However, sometimes AI algorithms produce outputs that are not based on training data, are incorrectly decoded by the transformer or do not follow any identifiable pattern. In other words, it “hallucinates” the response.
- The term may seem paradoxical, given that hallucinations are typically associated with human or animal brains, not machines. But from a metaphorical standpoint, hallucination accurately describes these outputs, especially in the case of image and pattern recognition (where outputs can be truly surreal in appearance).
- AI hallucinations are similar to how humans sometimes see figures in the clouds or faces on the moon. In the case of AI, these misinterpretations occur due to various factors, including overfitting, training data bias/inaccuracy and high model complexity.
- Source: <https://www.ibm.com/topics/ai-hallucinations>

## 4.3 Other cases of hallucinations

- Steven A. Schwartz and Peter LoDuca of Levidow, Levidow & Oberman
  - The two lawyers submitted fake legal research generated by ChatGPT.
  - USD 5,000 fine and a scolding by a federal judge.
- Michael Cohen court application
- Google's Bard chatbot incorrectly claiming that the James Webb Space Telescope had captured the world's first images of a planet outside our solar system.
- Microsoft's chat AI, Sydney, admitting to falling in love with users and spying on Bing employees.
- Meta pulling its Galactica LLM demo in 2022, after it provided users inaccurate information, sometimes rooted in prejudice.

## 4.4 Protection of information/privilege

- AI training
  - Data scraping and international data processing
- Data retention
  - Keep all inputs, outputs and error notices
- Production orders
  - Enforcement agencies secure inputs and outputs where client specific?

## 4.5 HMRC Information Powers

- Schedule 36 FA 2008 and Schedule 23 FA 2011
  - Schedule 36 covers Information AND Documents [in possession or power], also extends to “Computer Records, electronic documents and equipment”
  - Communication data (and overrides Investigator Powers Act 2016 for the purposes of the civil powers)
  - Bulk Data Gathering Powers

## 4.6 UK Government White Paper on AI Regulation

- Not a statutory framework but framework for regulating risk
  - safety, security and robustness
  - appropriate transparency and explainability
  - fairness
  - accountability and governance
  - contestability and redress

## 4.7 Law Society Principles

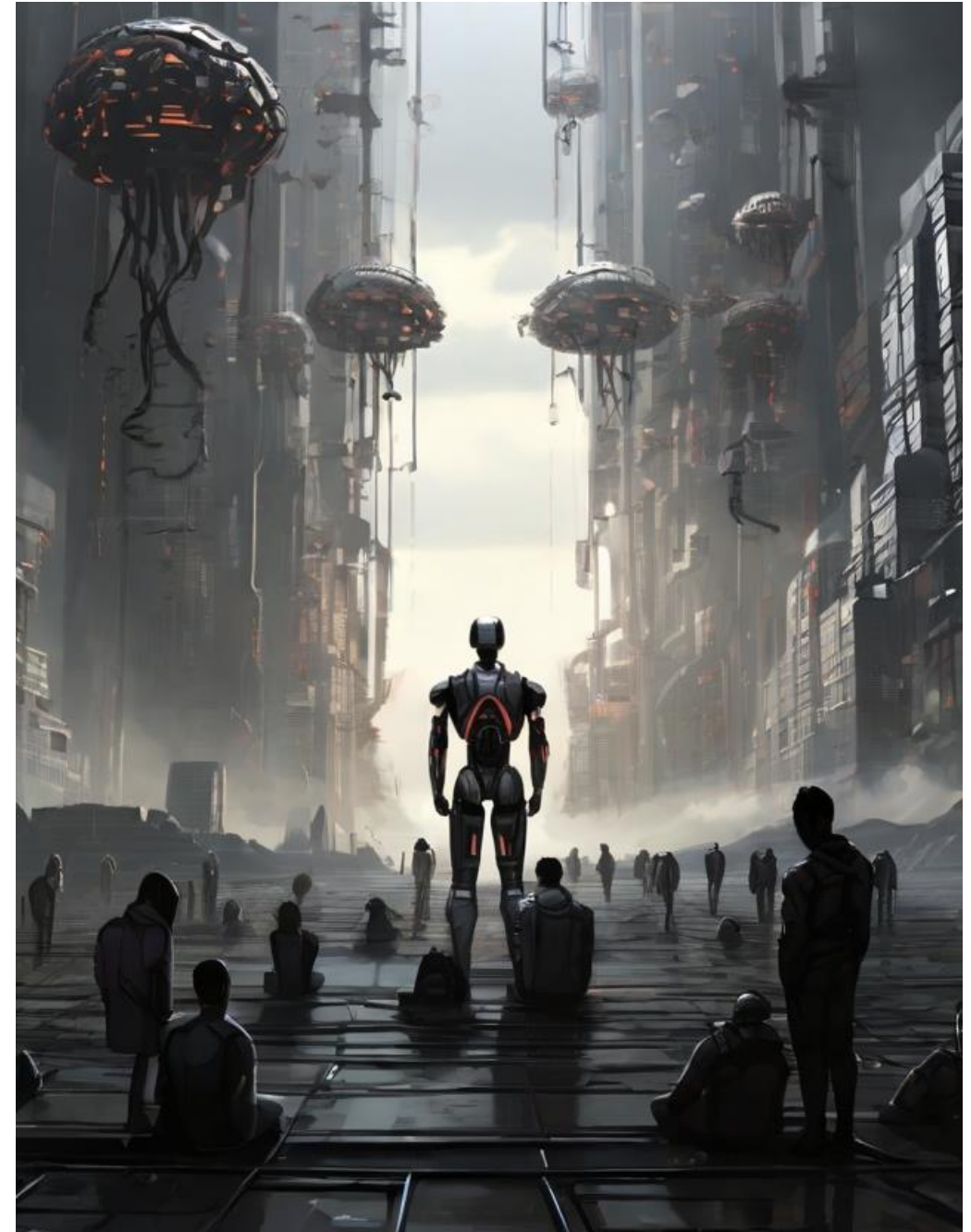
- Five main principles (reminders) identified to inform law tech design, development and deployment
  - Compliance
  - Lawfulness
  - Capability
  - Transparency
  - Accountability



## 4.8 AI in the legal and tax industry

- Opportunities
  - Chartered Institute of Taxation (CIOT) – Diploma in Tax Technology
  - Firms creating their own internal chatbots

AI: is it game over for  
humans?



Questions?



the global voice of  
the legal profession®