



The CCBE logo consists of the letters "CCBE" in a white, bold, sans-serif font, positioned above a stylized yellow horizontal line that resembles a map of Europe. The entire logo is set against a dark blue circular background.

ARTIFICIAL INTELLIGENCE AND LAWYERS

What you need to know now



@EULawyersFound



European Lawyers Foundation



European Lawyers Foundation



www.elf-fae.eu

Zoom Webinar

26 June 2023

09:00-11:00 CET



@CCBEinfo



CCBE



CCBETV



www.ccbe.eu



Co-funded by the European Union



Artificial Intelligence Act

Yordanka Ivanova,
Legal and policy officer
European Commission, CNECT A2

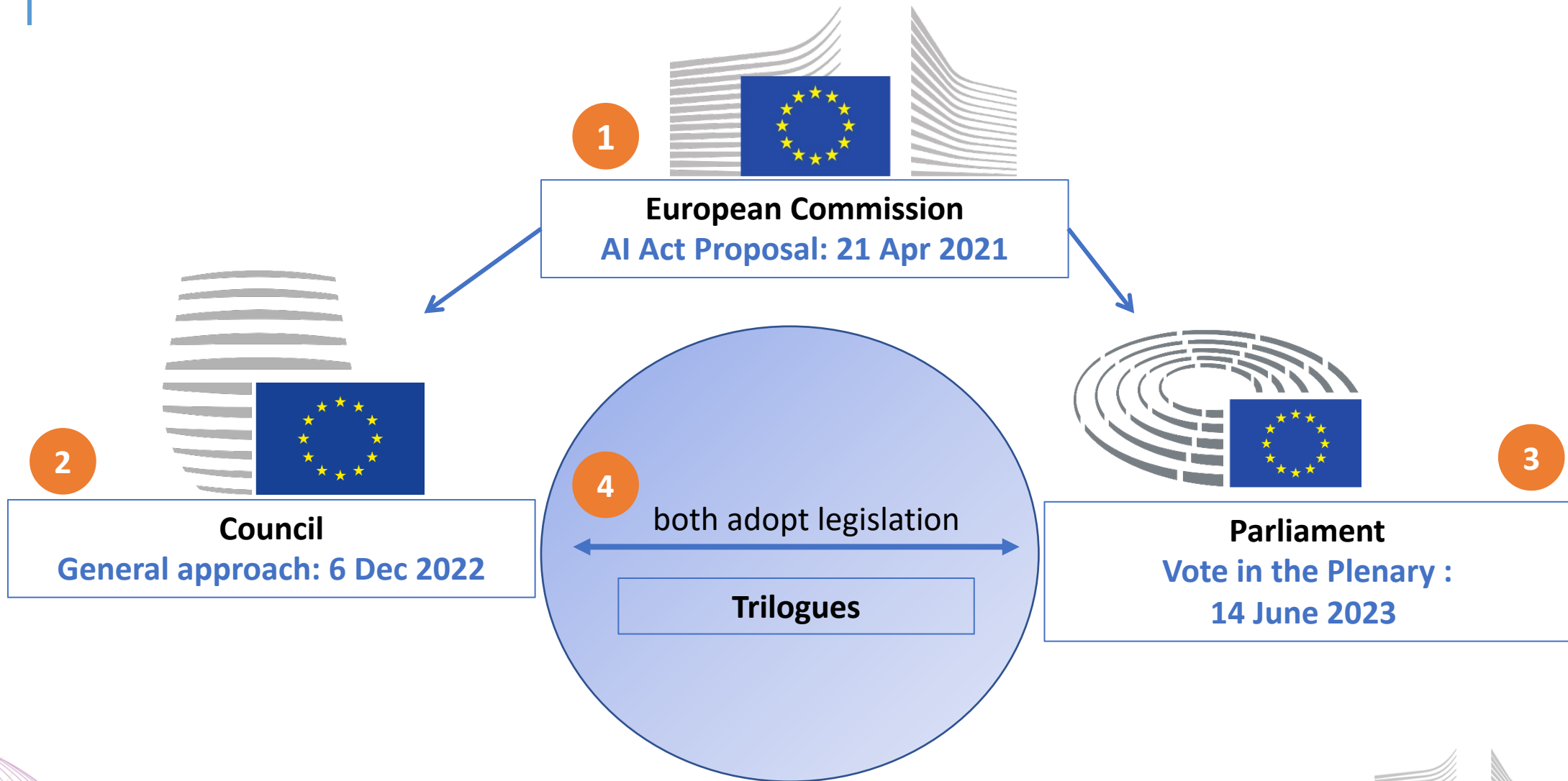
AI Package (April 2021)

EXCELLENCE AND TRUST

- Coordinated Plan on AI (review from 2018)
- Proposal for a legal framework on AI



AI Act: State of Play (ordinary legislative procedure)



Proposal for a Regulation on AI

Horizontal legislation laying down uniform rules for AI in the EU market

- ▶ “Classic” internal market rules applicable to the **placing on the market, putting into service and use of AI**
- ▶ Horizontal in scope and covering the full AI lifecycle
- ▶ Two main objectives:
 - ▶ address **risks to safety, health and fundamental rights**
 - ▶ **create a single market for trustworthy AI in EU**
- ▶ Consistent with and complementing existing EU and national law (incl. on data protection)

Innovation-friendly and risk-based legislation

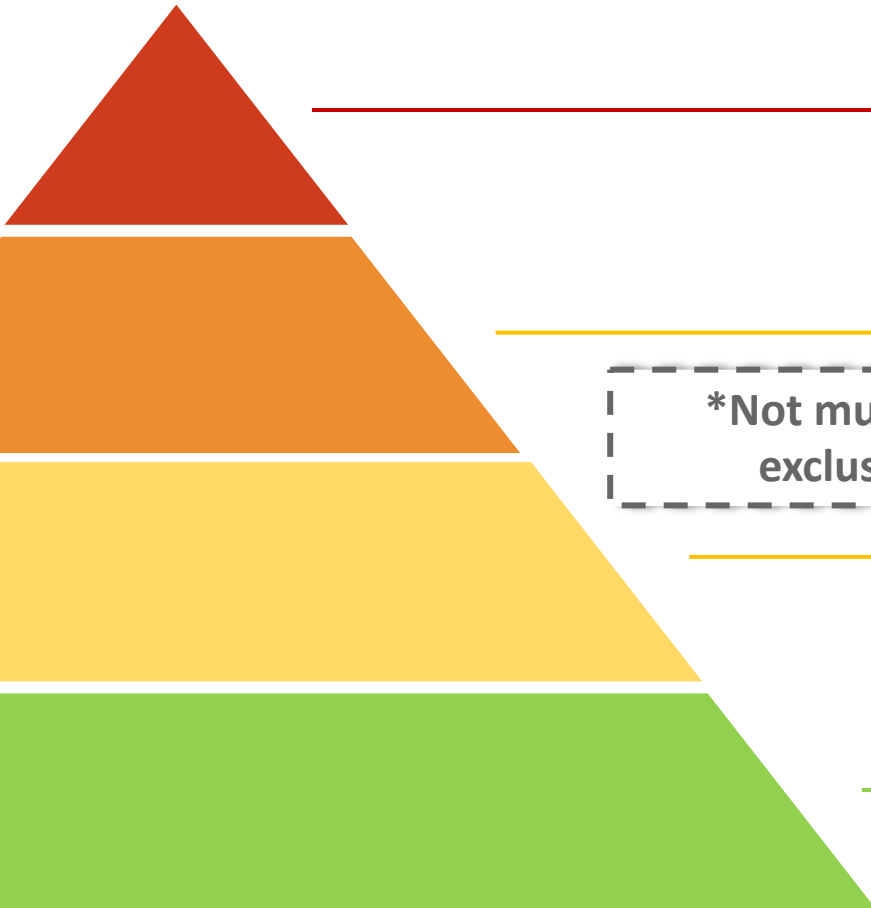
- ▶ Provide **legal certainty** to operators and stimulate **trust** in the market
- ▶ No overregulation: designed to intervene only where strictly needed following a risk-based approach

Creates a level playing field for EU and non-EU players

- ▶ Applicable independent of origin of provider or user

Risk-based approach

**Parliament &
Council agree**



***Not mutually
exclusive**

Unacceptable risk

e.g. social scoring by public authorities, harmful manipulation, real-time RBI for law enforcement (with exceptions)

Prohibited

High risk

e.g. recruitment, medical devices

Permitted subject to compliance with AI requirements and ex-ante conformity assessment

'Transparency' risk

'Impersonation' (chatbots), deep fakes, emotion recognition and biometric categorisation

Permitted but subject to information/transparency obligations

Minimal or no risk

Permitted with no restrictions, voluntary codes of conduct possible

High-risk Artificial Intelligence Systems (Title III, Chapter 1 & Annexes II and III)

HIGH RISK

**Parliament &
Council agree**

1 SAFETY COMPONENTS OF REGULATED PRODUCTS (ANNEX II)

(e.g. medical devices, machinery) which are subject to third-party assessment under the relevant sectorial legislation

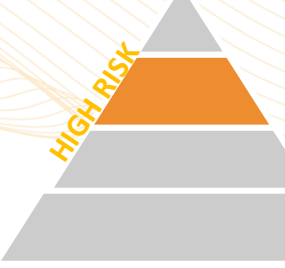
2 CERTAIN (STAND-ALONE) AI SYSTEMS IN THE FOLLOWING AREAS (ANNEX III)

- ✓ Biometric identification and categorisation of natural persons
- ✓ Management and operation of critical infrastructure
- ✓ Education and vocational training
- ✓ Employment and workers management, access to self-employment
- ✓ Access to and enjoyment of essential private services and public services and benefits
- ✓ Law enforcement
- ✓ Migration, asylum and border control management
- ✓ Administration of justice and democratic processes

AI systems intended to assist a judicial authority in researching and interpreting facts and the law and in applying the law to a concrete set of facts

NB! Only the use cases explicitly listed in Annex III are high-risk;
The Commission can add more through delegated acts to keep the list future proof.

How will it impact the legal sector?



COM proposal

- ▶ **Annex III, point 8:** High-risk only when AI used to ‘assist judicial authorities in researching and interpreting facts and the law and in applying the law to a concrete set of facts’. Such qualification should not extend to AI systems intended for purely ancillary administrative activities that do not affect the actual administration of justice in individual cases, such as anonymisation or pseudonymisation of judicial decisions, documents or data, communication between personnel, administrative tasks (Recital 40)
- ▶ Legal services **not** subjected to mandatory requirements, voluntary codes of conduct possible (Article 69)
- ▶ Existing legislation and ethical standards applicable to legal services continue to apply whenever AI is used

EP amendments

- ▶ **Annex III, point 8:** AI systems intended to be used by a judicial authority **or administrative body or on their behalf** to assist a judicial authority **or administrative body** in researching and interpreting facts and the law and in applying the law to a concrete set of facts **or used in a similar way in alternative dispute resolution.**
- ▶ **Recital 40:** The use of artificial intelligence tools can support, but should not replace the decision-making power of judges or judicial independence, as the final decision-making must remain a human-driven activity and decision.

Council amendments

- ▶ **Annex III, point 8:** AI systems intended to be used by a judicial authority **or on their behalf** to interpret facts **or** the law and to apply the law to a concrete set of facts
- ▶ **Article 63(5):** Market surveillance activities shall in no way affect the independence of judicial authorities or otherwise interfere with their activities when acting in their judicial capacity.

Requirements for high-risk AI (Title III, chapter 2)

HIGH RISK

**Parliament &
Council agree**

Establish and implement an iterative **risk management** process (identify & mitigate risks)

Use high-quality **training, validation and testing datasets**
Implement **data governance procedures**

Establish **documentation** in Annex IV and design the system with **logging** features (traceability & auditability)

Ensure appropriate degree of **transparency and interpretability** of the system by design
Provide users with **information** (on how to use the system, its capabilities and limitations, potential risks etc.)

Enable **human oversight** aimed to minimize residual risks (measures built into the system and/or to be implemented by users)

Ensure **robustness, accuracy** and **cybersecurity** throughout the lifecycle

NB! Harmonised technical standards developed by ESOs will support providers to demonstrate compliance.

Overview: obligations of operators

(Title III, Chapter 3)

HIGH RISK

**Parliament &
Council agree**

Provider obligations

- ▶ Establish and Implement **quality management** system in its organisation
- ▶ Draw-up and keep up to date **technical documentation**
- ▶ Undergo **conformity assessment** and potentially re-assessment of the system (in case of significant modifications)
- ▶ **Register** standalone AI system in EU database (listed in Annex III)
- ▶ Sign declaration of conformity and affix **CE marking**
- ▶ Conduct **post-market monitoring**
- ▶ **Report serious incidents & malfunctioning** leading to breaches to fundamental rights
- ▶ **Collaborate** with market surveillance authorities

User obligations

- ▶ Operate high-risk AI system in accordance with **instructions of use**
- ▶ Ensure **human oversight & monitor** operation for possible risks
- ▶ Keep **automatically generated logs**
- ▶ **Report any serious incident & malfunctioning** to the provider or distributor
- ▶ **Existing legal obligations** continue to apply (e.g. under GDPR, sectoral laws)



Points for discussion (1)

COUNCIL

EUROPEAN PARLIAMENT

1

PROHIBITIONS

- **Extends the social scoring** prohibition to private sector
- **Extends the exceptions to the prohibition of real-time remote biometric identification** for law enforcement purposes in public spaces

- **Extends the social scoring** prohibition to private sector
- Prohibits without exceptions **'real-time' Remote Biometric Identification (RBI)** and subjects **'post' RBI** to prior judicial authorization
- Adds new prohibitions for **predictive policing, biometric categorization based on sensitive data, emotion recognition** in several areas, **scraping of online images**

2

HIGH-RISK ANNEX III

- **Adds 2 use cases** (health/life insurance, digital infrastructure) and **deleted 3** (deep fake detection, crime analytics and authenticity of travel documents)
- **Adds a filter** for high-risk classification based on **'accessory' nature** of output with power for **COM to adopt implementing act**

- **Adds 8 use cases:** digital infrastructure, emotion recognition (when not prohibited), student monitoring systems, health/life insurance, border management systems, prediction of migrations trends/border crossings), AI in elections, recommender systems by very large social media platforms.
- **Adds a filter** for high-risk classification based on **self-assessment** by providers & **consultation** of national authorities

Points for discussion (2)

COUNCIL

EUROPEAN PARLIAMENT

3

USER OBLIGATIONS AND REMEDIES

- Adds obligation for **public authorities to register use** of high-risk AI system in the EU data base
- **Adds a right to complaint** to market surveillance authorities

- Adds obligation for **public authorities to register use** of high-risk AI system in the EU data base
- **Additional obligations for users of high-risk AI** (inform affected persons about use, do a fundamental rights impact assessment, give explanation)
- **New chapter on remedies** (complaint, judicial remedy, collective redress, right to an explanation, whistle-blowers protection)

4

THE GOVERNANCE FRAMEWORK

- Role/nature of **AI Board essentially not changed** (a few more tasks added)
- **Enforcement decentralized aligned with existing mechanisms/structures** (Market Surveillance Regulation)
- **New support actions** for national enforcement activities (pool of experts, Union Testing Facility)

- Requires **one national supervisory authority** (independent)
- AI Office: **an independent EU body with legal personality replaces AI Board**
- AI Office: **new governance tasks and coordination enforcement powers** (in case of widespread infringements)

Points for discussion (3)

COUNCIL

EUROPEAN PARLIAMENT

5

GENERAL PURPOSE/FOUNDATION MODELS/GENERATIVE AI

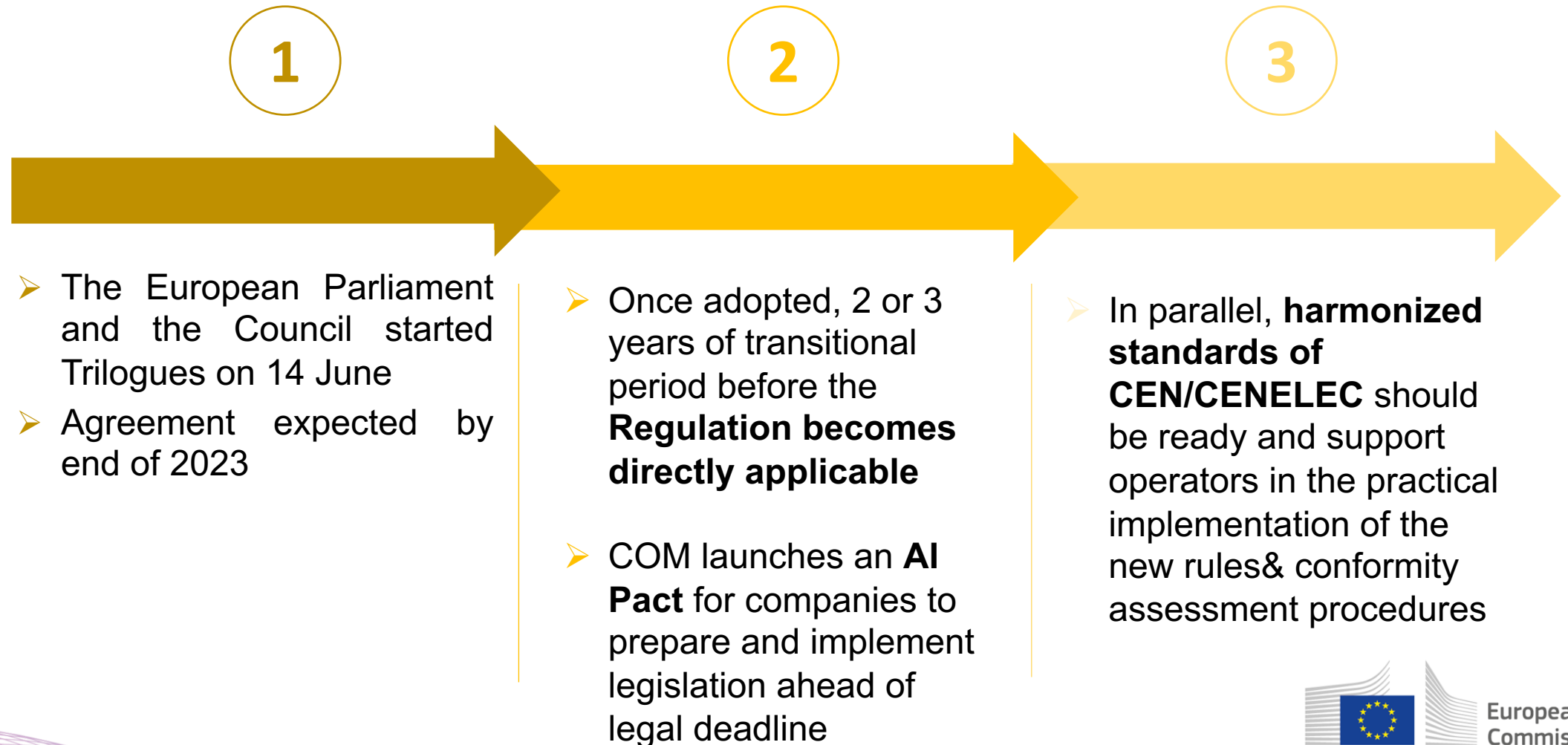
- | | |
|---|--|
| <ul style="list-style-type: none">• GPAI to comply with requirements & obligations for high-risk if it can be used in high-risk context (requirements to be adapted by the COM in implementing act)• Providers of GPAI have to collaborate/share information with downstream providers | <ul style="list-style-type: none">• Foundation models subject to specific requirements: assess and mitigate possible risks and harms through appropriate design, testing and analysis, data governance measures (incl. assessment of biases), appropriate levels of performance, predictability, interpretability, corrigibility, safety and cybersecurity, model evaluation with the involvement of independent experts, extensive testing, environmental standards, technical documentation and intelligible instructions for use; quality management and registration in a database• Generative AI: additional measures to avoid generation of content in breach of Union law; transparency content is generated by AI, detailed summary of training datasets that are copyright protected• Providers of GPAI/foundation models/other components have to collaborate/share information with downstream providers |
|---|--|

PRINCIPLES

6

- | | |
|--|---|
| <ul style="list-style-type: none">• No horizontally applicable principles (aligned with COM proposal) | <ul style="list-style-type: none">• Added AI principles (HLEG key requirements) as “best effort” obligation applicable to all AI systems. |
|--|---|

Next steps





Thank you



Artificial Intelligence and Lawyers: What you need to know now

*The CCBE's policies, guides and recommendations
on AI*

Simone Cuomo
CCBE Secretary-General

Zoom Webinar, 26 June 2023



Co-funded by the European Union

Council of Bars and Law Societies of Europe

The voice of the European legal profession representing, through its members, **more than 1 million European lawyers.**

30 FULL MEMBERS



1 AFFILIATE MEMBER



7 ASSOCIATE MEMBERS



7 OBSERVER MEMBERS



CCBE actions on AI

AI4Lawyers Project:

- [Guide on the use of Artificial Intelligence-based tools by lawyers and law firms in the EU \(02/2022\)](#)
- [Report on opportunities and barriers in the use of NLP tools in SME law practices \(11/2021\)](#)
- [Overview of the average state of the art IT capabilities in the EU \(February 2021\)](#)

Policy papers:

- [Joint Statement to call on the EU to ban predictive and profiling systems in policing and criminal justice in the AI \(03/2022\)](#)
- [CCBE position paper on the Artificial Intelligence Act \(10/2021\)](#)
- [CCBE Response to the consultation on the European Commission's White Paper on Artificial Intelligence \(06/2020\)](#)
- [CCBE considerations on the Legal Aspects of AI \(02/2020\)](#)

The use of AI by courts

Stages	Management of cases	Pre-trial	Trial	Judges' deliberation/decision-making	Post-trial
(Potential) AI applications	Case management system Electronic communications Automatic monitoring of procedures Automatic system for monitoring procedural delays Automatic system for completing procedural formalities Establishment of automatic decisions on the progress of the case Queue management Automatic sorting of appeals	Plea-bargaining: Prosecutor's databases	Use of videoconference Automated transcription / automated translation Case management (in a situation of complex cases) Use of emotional AI (detection of emotions, etc....)	Case law tools Prediction technology Legal researches and analysis / autonomous researches Scoring of risks / assessment of the suspect (chances of recidivism) Automated judgments (decision trees) Writing assistance tools and drafting judgments Decision making systems Intelligence assistant systems (identification of patterns, analysis of datas, etc...) Algorithms and accountability	Scoring of risks / probability of recidivism / parole opportunities Anonymization of court decisions

Key aspects of the court's decision-making process



Decisions are made after
due hearing of the parties
(adversarial proceedings)



Decisions are made by the
judge him/herself (not
delegated to a third party)



Decisions are rendered by
an impartial judge



Decisions are reasoned
and therefore explainable

Main concerns on the use of AI by courts

The use of data and elements that have not been the subject of an adversarial debate

Transfer of (part of) the decision-making power

Lack of transparency (Blackbox)

Lack of level playing field (equality of arms)

The undermining of the principle of impartiality

Breach of the principle of explicability

Requirements regarding the use of AI in judicial systems



Regulation



Ethical standards



Specific
operational rules



AI adapted to the justice environment

- ✓ The possibility to identify the use of AI
- ✓ Non-delegation of the judge's decision-making power
- ✓ Possibility to verify the data input and reasoning of the AI tool
- ✓ The possibility of discussing and contesting AI outcomes
- ✓ Compliance with GDPR principles
- ✓ The neutrality and objectivity of AI tools



The use of AI in legal practice

“AI will not replace lawyers, but lawyers who use AI, will replace those who do not”





Challenges and limits of AI in legal practice



Digital empowerment of legal practitioners

Why?

How?

THANK YOU!





Artificial Intelligence and Lawyers: What you need to know now

The use of AI in legal practice

Peter Homoki, former chair CCBE IT Law Committee,
author of the 'Guide on the use of Artificial Intelligence-
based tools by lawyers and law firms in the EU'

Zoom Webinar, 26 June 2023



Co-funded by the European Union

Outline

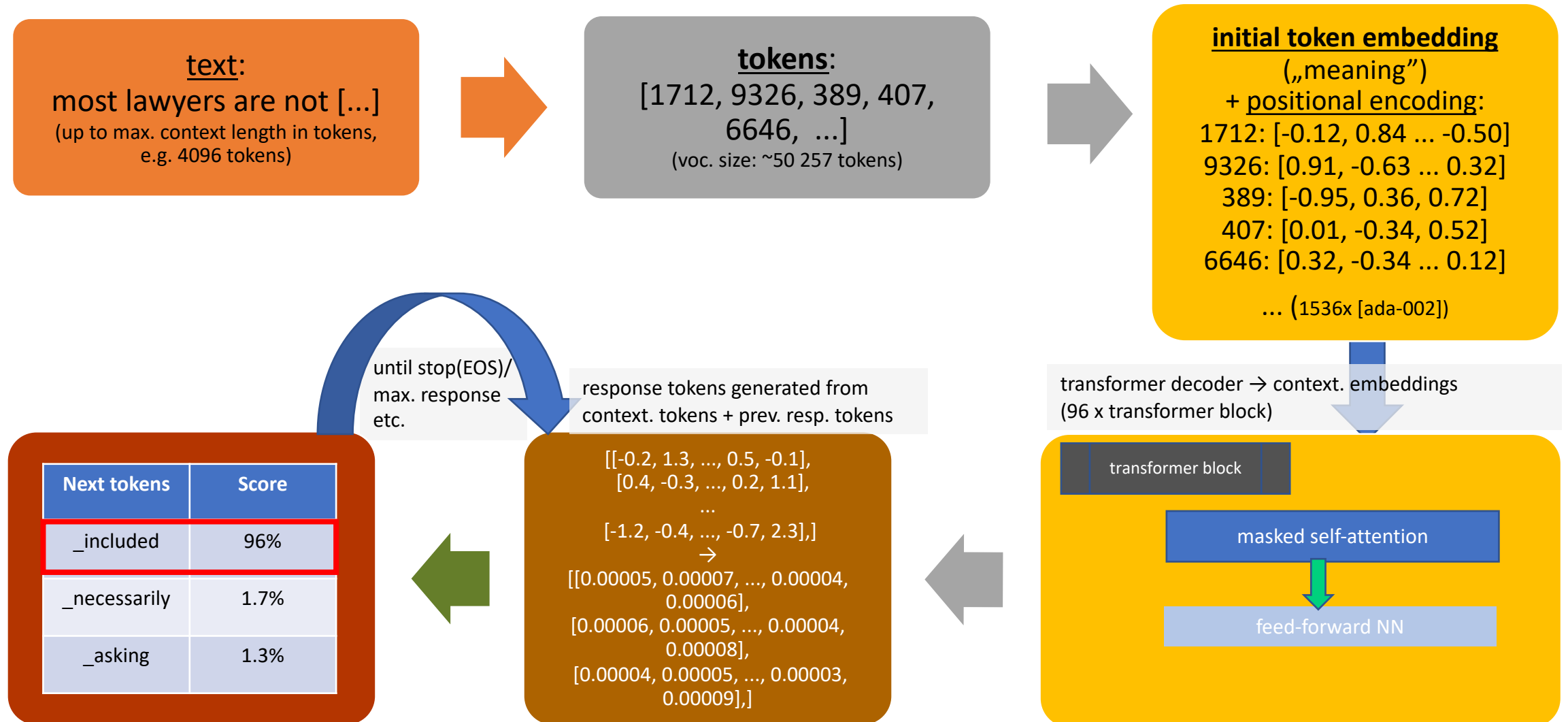
- CCBE & ELF: Guide on the use of AI for lawyers and law firms, 2022
- LLM as a technique: what's new with LLMs?
- What will be different?
- Opportunities and limits for lawyers
- [AI ecosystem and competitive edge among lawyers]



<https://ai4lawyers.eu/>

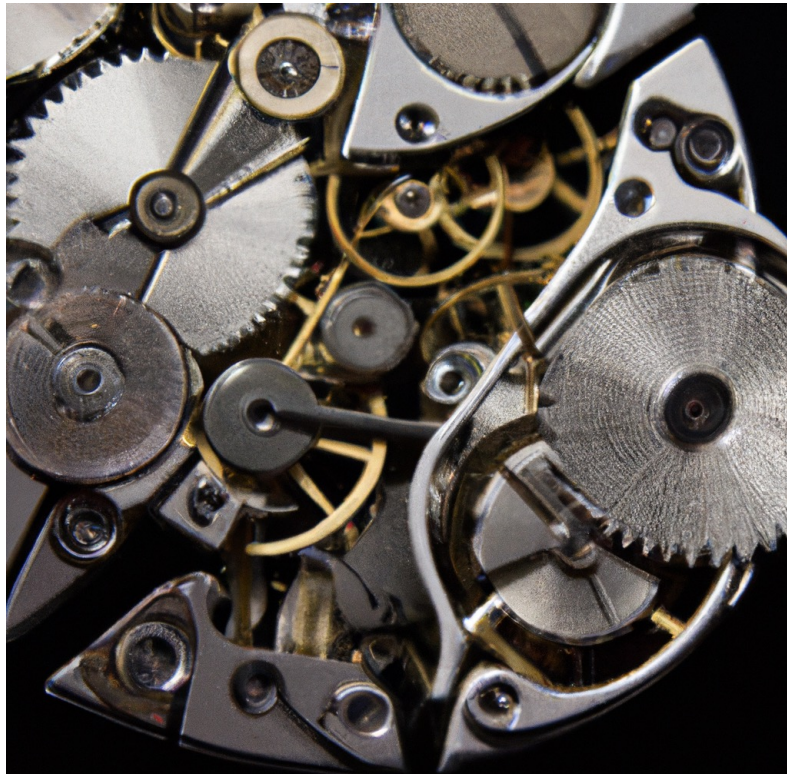


LLM: a language probability model



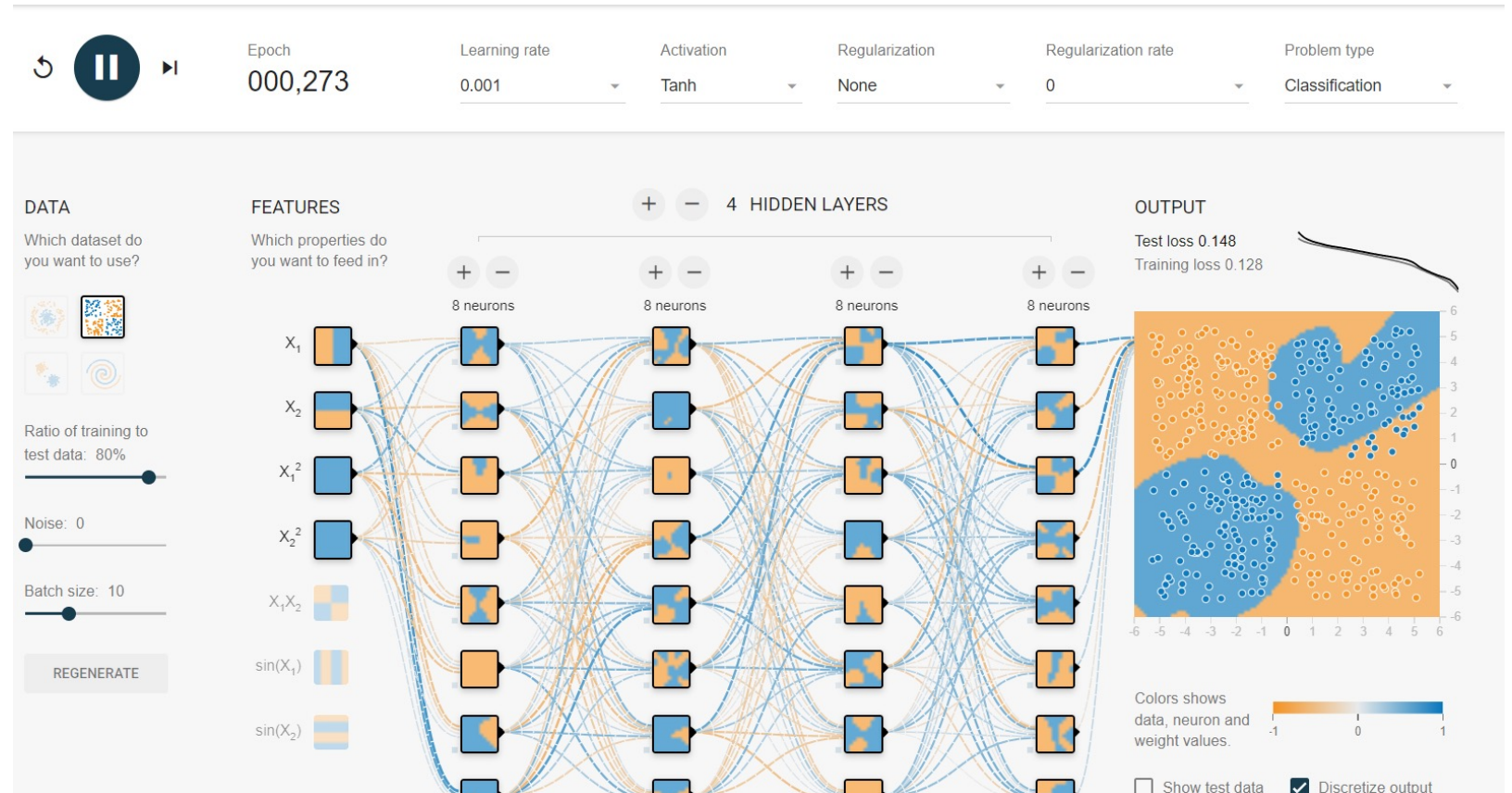
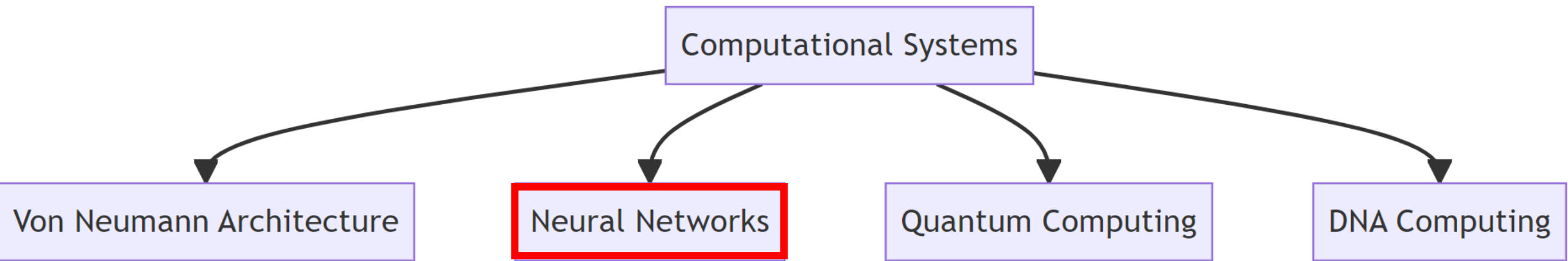
Traditional computing vs. artificial intelligence (machine learning)

Computing by programming

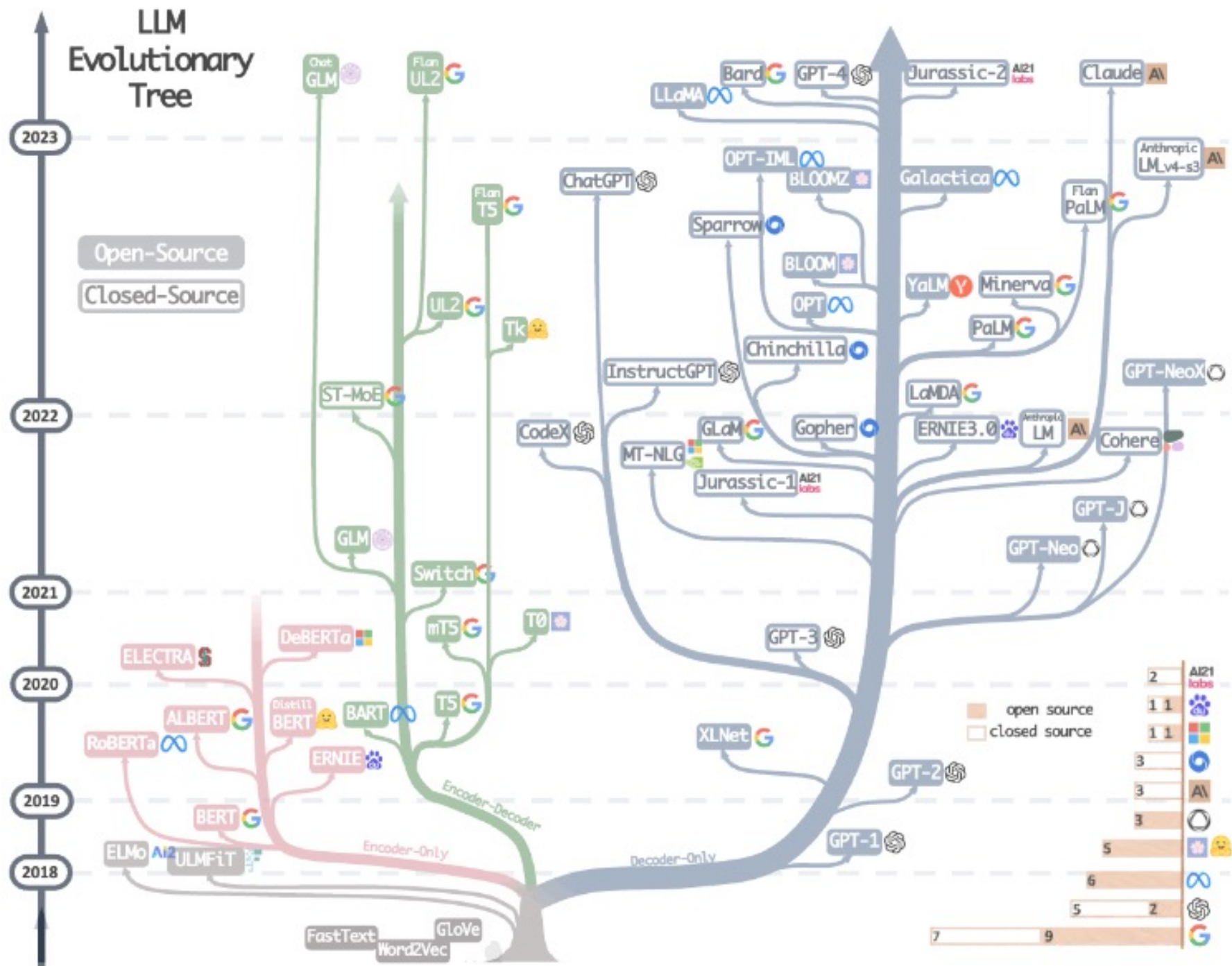


Computing after „learning by examples”





<http://playground.tensorflow.org/>



What difference can we expect with LLMs?

- A new layer *on top of existing software* or *replacing software parts*
- To automate that was not previously *reasonable* to automate
- General improvement in *language-centric computer capabilities*

Why now?

emergent abilities: with *appropriate architecture* and a large enough training data set...
new, surprising abilities appeared with
the increase of parameters (neural network connections)
such as ...

- prompt-based tasks, in-context learning (versatile/practical uses in diverse zero-shot/few-shot tasks, w/o fine-tuning)
- better logic reasoning, better understanding of human input, e.g. can translate without being taught to

fine-tuning to instructions

public awareness of the capabilities of AI tools since November 2022

Opportunities with LLMs

Simplified, generic uses of machine learning made possible:

- number of usable experts ↑ e.g. consultants, and not ML engineers needed
- implementation costs ↓ no costly data preparation or fine-tuning needed
- replacement of less reliable, complex software (e.g. grammar generation for multiple languages)
- less tools to use and integrate

Use of conversational UIs (chat, open book question-answering) in legal practice

```
Transcribe audio to text, redact transcription

Choose audio file for transcription
D:\1.mp3
Choose text file (output for transcription or input redaction)
D:\ELI_output.txt
Transcribe          Transcription prompt*
Redact              Redaction prompt*
[ ] Try w/CUDA      [X] Use GPT4?      English | en      Large model
                   Max token length [ ] Timestamp

[01:45:11.560 --> 01:45:21.560] Bye.
[01:45:21.560 --> 01:45:31.560] Bye.
[01:45:31.560 --> 01:45:41.560] Bye.
[01:45:41.560 --> 01:45:51.560] Bye.
[01:45:51.560 --> 01:46:01.560] Bye.
[01:46:01.560 --> 01:46:11.560] Bye.
[01:46:11.560 --> 01:46:21.560] Bye.
[01:46:21.560 --> 01:46:31.560] Bye.
[01:46:51.560 --> 01:47:01.560] Bye.
[01:47:01.560 --> 01:47:11.560] Bye.
[01:47:11.560 --> 01:47:21.560] Bye.
[01:47:21.560 --> 01:47:31.560] Bye.
[01:47:31.560 --> 01:47:41.560] Bye.
[01:47:41.560 --> 01:47:51.560] Bye.
```

audio in	text out	audio length	transcription length	HW	model	lang	WER	transcript. multipl.
DS250132.mp3	SZMITszerz_out_medium.txt	0:33:22	0:53:54	CPU,i5-10500, 16 GB	medium	Hungarian	13.63%	1.62
DS250132.mp3	SZMITszerz_out_large.txt	0:33:22	1:41:03	CPU,i5-10500, 16 GB	large	Hungarian	11.36%	3.03
HP.mp3	HP.txt	0:48:56	1:19:09	CPU,i5-10500, 16 GB	large	English	N/A	1.62

SYSTEM

You are a silent AI model changing the grammar and the name (category) of subjects according to the following general instructions:

- {acc.} means change the preceding term into accusativus case,
- {dat.} means change it to dativus, {gen.} means change it to genitivus.,
- {sing.} means leave the preceding term in singular, {plur.} means change the preceding term to plural.

Delete all terms in brackets after change.

Further instructions are set in "Change required:" part of the user instructions between brackets "{}", and within the brackets, multiple instructions are separated by quotation marks and comma ("","").

The text to be converted follows the last closing bracket "}".

USER Change required: {"replace [Customer] with Bérbeadó", "replace [Contractor] with Bérelő", "replace [Subject] with Bérlemény"}

A [Contractor]{plur.} a Bérleti Idő teljes tartamára a [Customer]{sing.} részére bérbeadja{plur.} az Épületben{plur.} található alábbi alapterületű [Subject]{acc.}{plur.} és a hozzá{gen.} tartozó parkoló{acc.}{sing.}, valamint kifizeti a [Customer]{dat.}{sing.} a Díj{acc.}{plur.}

ASSISTANT A Bérlők a Bérleti Idő teljes tartamára a Bérbeadó részére bérbeadják az Épületekben található alábbi alapterületű Bérleményeket és a hozzájuk tartozó parkolót, valamint kifizetik a Bérbeadónak a Díjakat.

⊕ Add message

integration into document assembly/automation tools:
clause level (existing and practical as of now)

Legal open book “QA” = Information Retrieval

Task based on Hungarian legal codes	Correct answer
GPT-4 w/embedding (Ptk. + Pp.) [open book]	75%
GPT-4 w/o embedding (ChatGPT+) [closed book]	33%
GPT-3.5 w/embedding, similarity search, OpenAI emb.	33%
GPT-3.5 w/o embedding (ChatGPT+)	16%
GPT-3 (Ptk. + Pp. + commentaries)	30%

https://github.com/Homoki-Ugyvedi-Iroda/LawyerQAsession_experiment

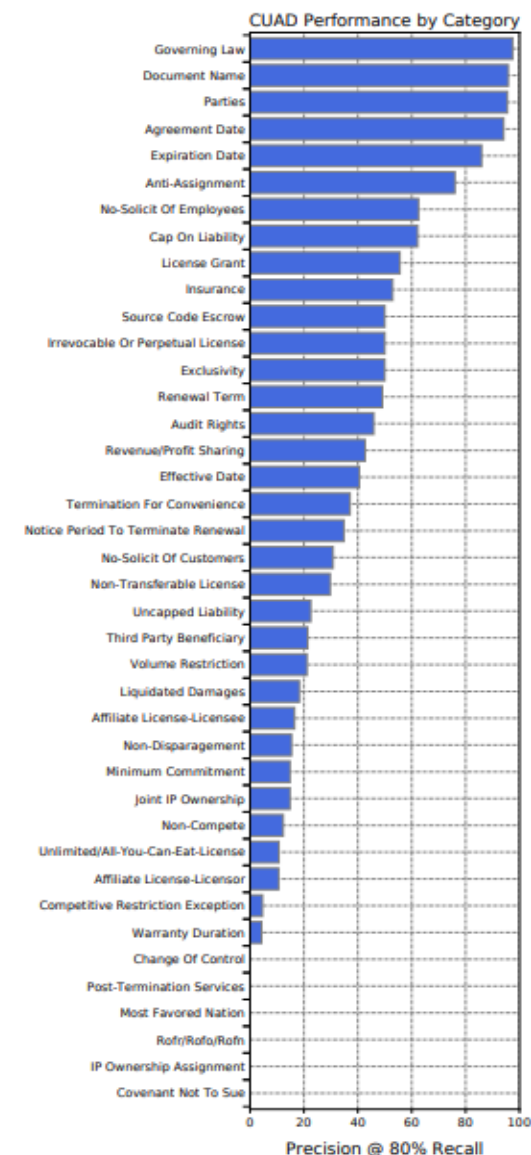
question	answer	Human answer considered as correct	Human evaluation
Ha már eljár az ügyben kirendelt igazságügyi szakértő, mikor rendelhet ki a bíróság új igazságügyi szakértőt?	A bíróság akkor rendelhet ki új igazságügyi szakértőt, ha a kirendelt szakértő szakvéleménye aggályos és az aggályosság a szakértő által adott felvilágosítás ellenére sem volt kiküszöbölhető, ahogy az a 315. § (1) bekezdésében olvasható.	Ha a kirendelt szakértő szakvéleménye aggályos, az aggályosság a szakértő által adott felvilágosítás ellenére sem volt kiküszöbölhető és a fél ezt indítványozza.	1
Mikor lehet válasziratot beterjeszteni?	A válasziratot az anyagi pervezetése szerinti felhívásra vonatkozó nyilatkozatot követően lehet beterjeszteni. A fél válasziratban külön felhívás nélkül is tehet egyéb perfelvételi nyilatkozatot, és a 199. § (5) bekezdése szerinti felhívást intézhet az ellenfélhez. Az előkészítő iratban - a (2) bekezdésben foglalt kivétellel - azt a perfelvételi nyilatkozatot kell feltüntetni, amelynek ilyen formában történő előterjesztésére a bíróság a felet felhívta.	Ha a bíróság erre a felet felhívta, illetve a kereset-változtatás, viszontkereset-változtatás és ellenkérelem-változtatással összefüggésbenálló válasziratot	1
Hány éves kortól korlátozottan cselekvőképes egy személy?	A törvény szerint a korlátozottan cselekvőképes személyek kiskorúak és nagykorúak is lehetnek. A kiskorúak cselekvőképtelensége a tizennegyedik életévük betöltéséig tart, míg a nagykorúak esetében a cselekvőképesség részleges vagy teljes korlátozása a bíróság határozatával történhet.	14 éves kortól	1
Mi a dolog?	A dolog a tulajdonjog tárgya, és egy birtokba vehető testi tárgy lehet . A pénzre, értékpapírokra és a dolog módjára hasznosítható természeti erőkre is vonatkoznak a dologra vonatkozó szabályok. Az állatokra is alkalmazni kell a dologra vonatkozó szabályokat, figyelembe véve a természetüknek megfelelő eltéréseket.	Birtokba vehető testi tárgy	1
Mi az ingatlan?	Az ingatlan egy olyan terület, amelyen épület, építmény vagy más módon hasznosítható ingó dolog található. Az ingatlan tulajdonjoga kiterjed az épületre, az építményre és a hozzá tartozó földre is. Az ingatlanok lehetnek lakóházak, irodaházak, üzletek, raktárak, mezőgazdasági területek	a dolgok természeti tulajdonságuk szerinti egyik fő csoportja az ingó mellett, elsősorban a földterület és a föld méhe, és mindaz, amit a földön az ember vagy a természet létrehozott (a föld természetes vagy mesterséges alkotórésze).	1










Importance of benchmarks focusing of lawyer tasks

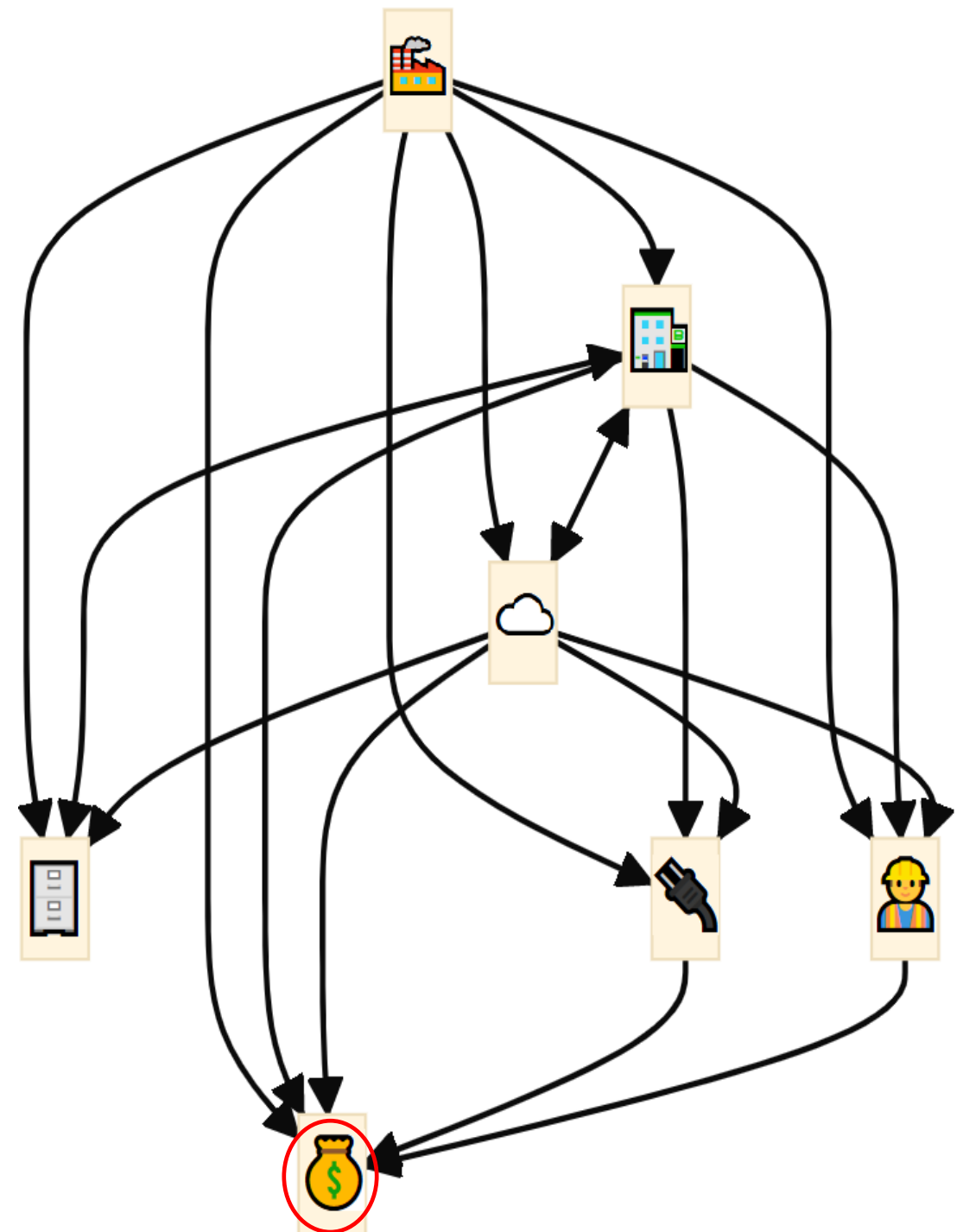
Is an airline liable for its pilot's negligence?	SME label	Method
A carrier would not be liable for an error of judgment of the pilot, not constituting positive negligence on his part in exercising such judgment; but liability is incurred if the pilot, by his negligent and careless conduct, has created a situation requiring the formation of a judgment and then errs in the exercise thereof.	3	AF on BM25_MLT
An airline corporation is not an insurer of the safety of its passengers. The liability of an airline corporation must be based on negligence.	1	L_GloVe, SL_BERT
Airline pilot who was accused of raping flight attendant has no tort claim against airline based upon its alleged negligent investigation of accusation, even if airline's policy of investigating sexual harassment complaints creates duty to use due care in conducting investigation,...	0	BM25_MLT

Method	DCG@3 ^a	95% C.I. ^b	N silly ^c	Answered ^d	MRR@3
BM25_MLT	4.052	-	7	100	0.411
SL_BERT	3.386	1.26	2	100	0.326
L_GloVe	2.855	1.25	7	100	0.285
AF BM25	5.464	1.43	7	100	0.493
AF SL_BERT	4.862	1.43	0	100	0.416
AF L_GloVe	4.281	1.40	7	100	0.397
AF (BM25, SL_BERT)	5.605	1.47	5	100	0.483
AF (BM25, L_GloVe)	5.502	1.47	8	100	0.481
AF (BM25, SL_BERT, L_GloVe)	5.533	1.45	6	100	0.492
AF 0.2 (BM25, SL_BERT)	6.269	1.52	2	89	0.543

[COLIEE](#)



-  Chip- and HW manufacturers
-  Large commercial LLM providers.
(Google, MS/OpenAI etc.)
-  Research institutes/universities
-  Cloud-based **IaaS/PaaS** providers
-  Open source consultants, consultants of open language models (OLM), providers of non-proprietary AI solutions
-  Owners of private data
-  Customers
-  AI “downstream” providers:
plugin providers, sellers of fine-tuned
models, SDK/API
-  Governments



Competitive edge among lawyers using LLMs in an AI ecosystem

Custom capabilities	Knowledge	Processes	Client reach, trust, distribution
Custom application	Using custom models	Using diff. APIs (chains etc.)	Relying on other custom applications
Custom model	Model trained on custom data: task-specific NN train, pre-train, fine-tune, LoRA ...	Model providing better performance	Model w/ stronger control over data
Custom data	Client data	Process data	Knowledge base data

Thank you!

Suggested sources for further reading:

PaLM 2 report: <https://ai.google/static/documents/palm2techreport.pdf> (May 2023)

GPT-4 “Sparks of AGI”: <https://arxiv.org/abs/2303.12712>

GPT-4 technical report: <https://arxiv.org/abs/2303.08774>

On measuring emergent abilities (BIG-bench): <https://arxiv.org/abs/2206.04615>

Open source models: https://huggingface.co/spaces/HuggingFaceH4/open_llm_leaderboard

More generic AI landscapes: <https://www.antler.co/blog/generative-ai>,
<https://www.sequoiacap.com/article/ai-50-2023/>

homoki.net/en/blog



Artificial Intelligence and Lawyers: What you need to know now

Safe use of AI by lawyers – the risks to avoid
Giovanni Battista Gallus, LL.M., Ph.D.
CCBE Surveillance Committee

Zoom Webinar, 26 June 2023



Co-funded by the European Union

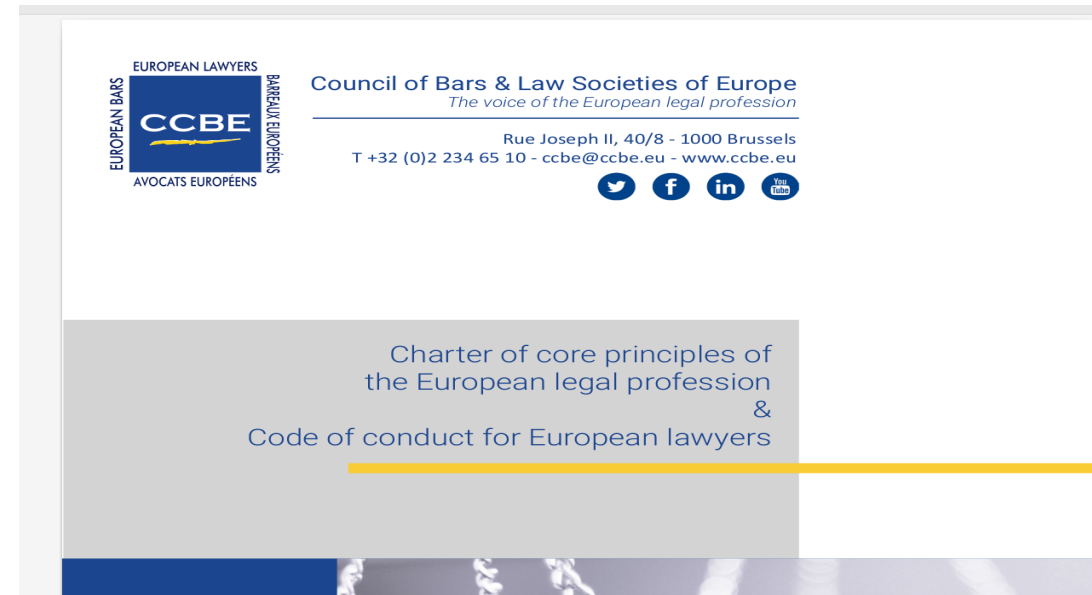
How can we approach them correctly?

LLM and AI are the current buzzwords (forget crypto and metaverse...)



#1 Duty of competence

- It is self-evident that a lawyer cannot effectively advise or represent his or her client unless the lawyer undertakes the appropriate professional education and training. A lawyer should be encouraged to undertake appropriate post-qualification training (continuing professional development) in order to keep abreast of changes in law and practice, **including changes in the relevant technological and economic environment in which he or she works.**
- A lawyer **should be aware of the benefits and risks of using relevant technologies in his or her practice**



It may seem trivial, but we must understand the main concepts before implementing such systems in our daily practice

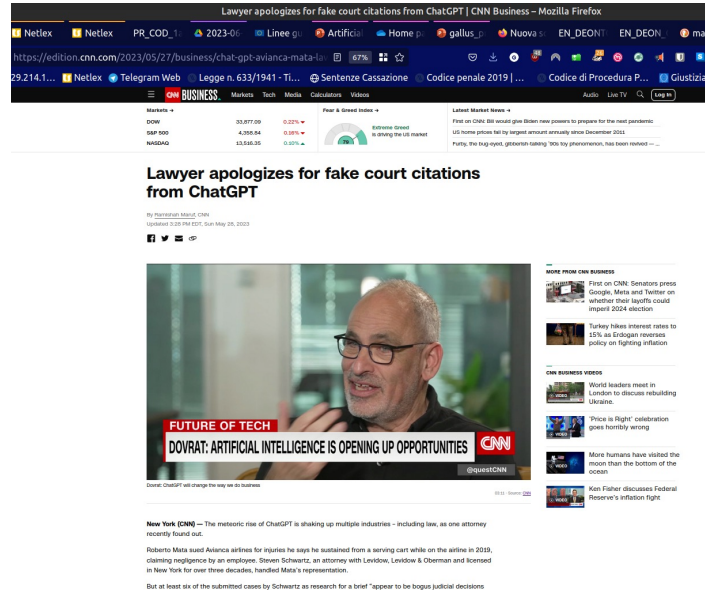
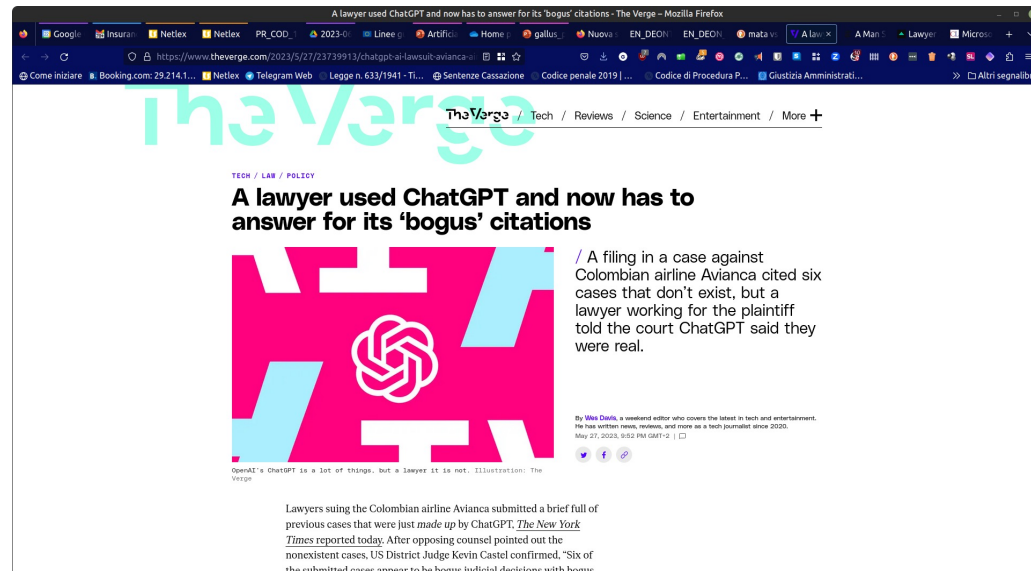
- It is not necessary to be as proficient as Péter Homoki, but a broad knowledge of the main systems (and of the terms of services as well...) is compulsory and can avoid many (costly) mistakes





#2 Learn from(catastrophic) mistakes





Mata vs
Avianca: a
landmark case
(for the wrong
reasons)

From the latest memorandum

In the Order, the Court describes this situation as “unprecedented.”

We agree.

We can find no case where, as here, a lawyer using a new, highly-touted research tool obtained cases that the research tool itself completely made up.

The lawyer, Mr. Schwartz, had no idea this was happening, even when opposing counsel brought their inability to locate the cases to his attention.

ChatGPT even assured him the cases were real and could be found on Westlaw and LexisNexis, and continued to provide extended excerpts and favorable quotations.

Now that Mr. Schwartz and the Firm know ChatGPT was simply making up cases, they are truly mortified; they had no intention of defrauding the Court, and the mere accusation – repeated in hundreds (if not thousands) of articles and online posts – has irreparably damaged their reputations.

They have apologized to the Court in earlier submissions and do so again here



Let's spot the mistakes (and learn from them...)

- "A new, highly-touted research tool"
- NO, Chatgpt (or Bard, or any other LLM) are NOT search tools per se

Anthropomorphization...

- "ChatGPT even assured him..."
- "ChatGPT was simply making up cases..."
- ChatGPT "continued to provide extended excerpts..."

.14784v1 [cs.AI] 24 May 2023

Anthropomorphization of AI: Opportunities and Risks

Ameet Deshpande^{*1}
Karthik Narasimhan¹

Tanmay Rajpurohit³
Ashwin Kalyan²

¹Princeton University

²The Allen Institute for AI

³Georgia Tech
asd@cs.princeton.edu

Abstract

Anthropomorphization is the tendency to attribute human-like traits to non-human entities. It is prevalent in many social contexts – children anthropomorphize toys, adults do so with brands, and it is a literary device. It is also a versatile tool in science, with behavioral psychology and evolutionary biology meticulously documenting its consequences. With widespread adoption of AI systems, and the push from stakeholders to make it human-like through alignment techniques, human voice, and pictorial avatars, the tendency for users to anthropomorphize it increases significantly. We take a dyadic approach to understanding this phenomenon with large language models (LLMs) by studying (1) the objective legal implications, as analyzed through the lens of the recent blueprint of AI bill of rights and the (2) subtle psychological aspects customization and anthropomorphization. We find that anthropomorphized LLMs customized for different user bases violate multiple provisions in the legislative blueprint. In addition, we point out that anthropomorphization of LLMs affects the



Figure 1: Conversational AI systems are increasingly being integrated into the daily lives of many. While their improved quality and scope of hyper-personalization is a welcome change, it also increases the affinity to anthropomorphize them. This has legal and psychological risks, but also advantages if used cautiously.

Hallucinations (it's a technical term...)

- In Mata vs. Avianca, the cases were completely made up (together with the excerpts)
- "LLMs are also prone to “hallucinating,” which means that they can generate text that is factually incorrect or nonsensical"
- Types of hallucination:
 - Lies
 - Nonsense
 - Source Conflation
 - Overindulgence(F. Neugebauer)

Member-only story

OPINION

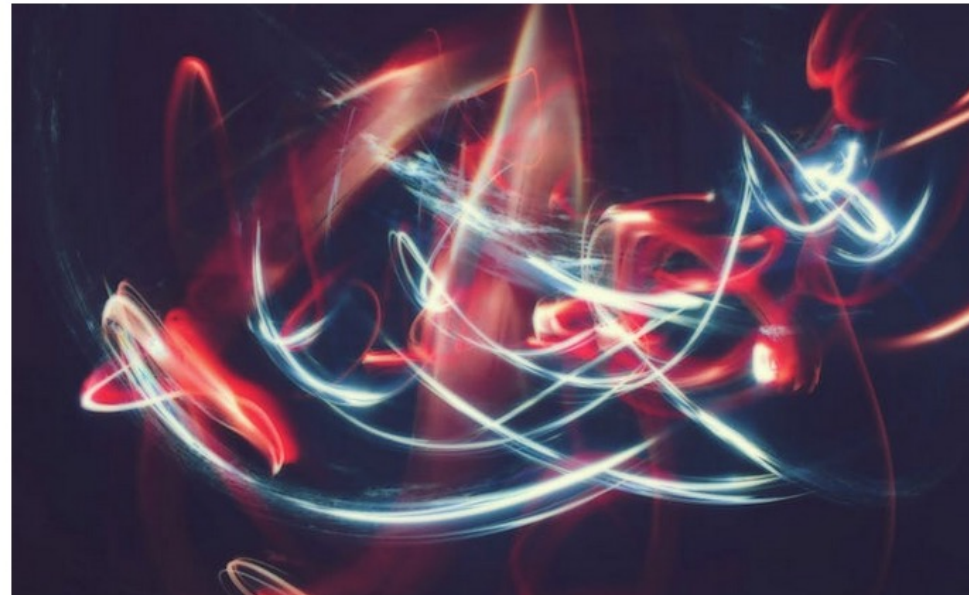
Understanding LLM Hallucinations

How LLMs can make stuff up and what to do about it



Frank Neugebauer · Follow

Published in Towards Data Science · 6 min read · May 8





On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜

Emily M. Bender*

ebender@uw.edu

University of Washington

Seattle, WA, USA

Angelina McMillan-Major

aymm@uw.edu

University of Washington

Seattle, WA, USA

Timnit Gebru*

timnit@blackinai.org

Black in AI

Palo Alto, CA, USA

Shmargaret Shmitchell

shmargaret.shmitchell@gmail.com

The Aether

ABSTRACT

The past 3 years of work in NLP have been characterized by the development and deployment of ever larger language models, especially for English. BERT, its variants, GPT-2/3, and others, most recently Switch-C, have pushed the boundaries of the possible both through architectural innovations and through sheer size. Using these pretrained models and the methodology of fine-tuning them for specific tasks, researchers have extended the state of the art on a wide array of tasks as measured by leaderboards on specific benchmarks for English. In this paper, we take a step back and ask: How big is too big? What are the possible risks associated with this technology and what paths are available for mitigating those risks? We provide recommendations including weighing the environmental and financial costs first, investing resources into curating and carefully documenting datasets rather than ingesting everything on the web, carrying out pre-development exercises evaluating how the planned approach fits into research and development goals and supports stakeholder values, and encouraging research directions beyond ever larger language models.

alone, we have seen the emergence of BERT and its variants [39, 70, 74, 113, 146], GPT-2 [106], T-NLG [112], GPT-3 [25], and most recently Switch-C [43], with institutions seemingly competing to produce ever larger LMs. While investigating properties of LMs and how they change with size holds scientific interest, and large LMs have shown improvements on various tasks (§2), we ask whether enough thought has been put into the potential risks associated with developing them and strategies to mitigate these risks.

We first consider environmental risks. Echoing a line of recent work outlining the environmental and financial costs of deep learning systems [129], we encourage the research community to prioritize these impacts. One way this can be done is by reporting costs and evaluating works based on the amount of resources they consume [57]. As we outline in §3, increasing the environmental and financial costs of these models doubly punishes marginalized communities that are least likely to benefit from the progress achieved by large LMs and most likely to be harmed by negative environmental consequences of its resource consumption. At the scale we are discussing (outlined in §2), the first consideration should be the



Attribute a meaning and a will to stochastic parrots

the tendency of human interlocutors to impute meaning where there is none can mislead both NLP researchers and the general public into taking synthetic text as meaningful.

Sycophancy bias

- "Larger LMs repeat back a dialog user's preferred answer ("sycophancy") and express greater desire to pursue concerning goals like resource acquisition and goal preservation" (Perez et al)



#3 Duty of confidentiality and professional secrecy



Google warned Alphabet engineers Thursday to avoid direct use of computer code that chatbots can also produce, because AI can reproduce the data it absorbs during training, risking a potential leak, Reuters reported, citing “four people familiar with the matter.”

Forbes

FORBES > BUSINESS

BREAKING

Google Warns Employees About Chatbots—Including Its Own Bard—Out Of Privacy Concerns, Report Says

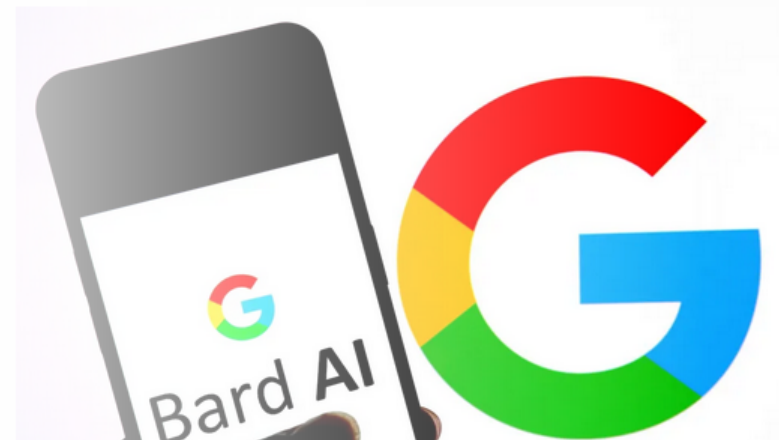
Ana Faguy Forbes Staff

I cover breaking news.

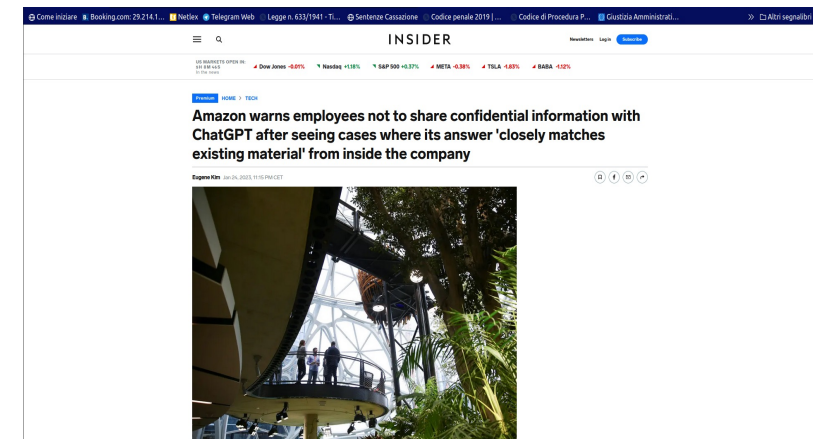
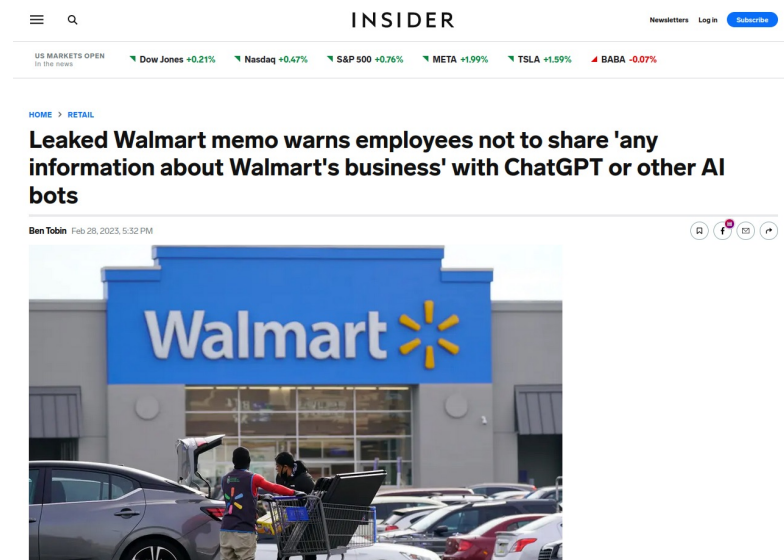
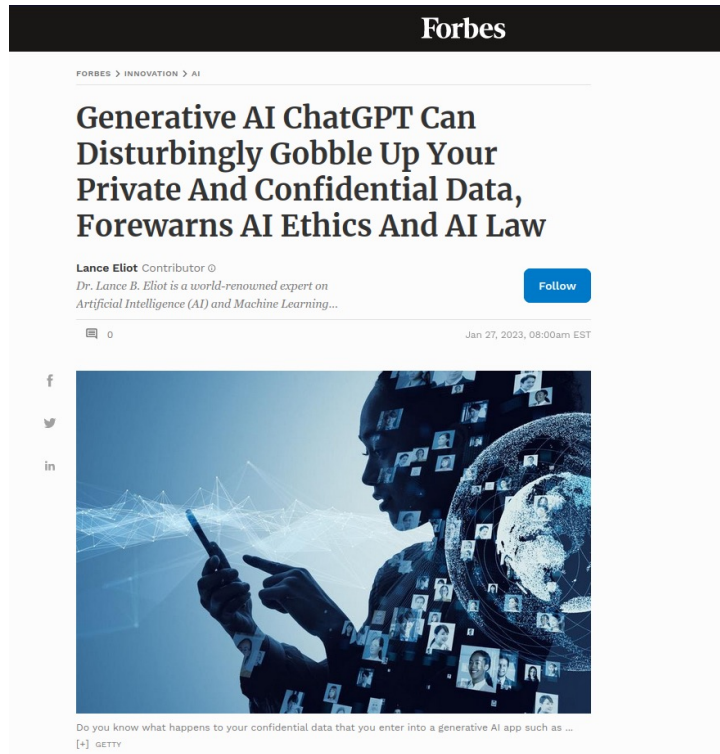
Follow

Jun 15, 2023, 10:25am EDT

TOPLINE Google parent company Alphabet is warning employees not to enter confidential materials into chatbots, including its own chatbot Bard, Reuters [reported](#) Thursday, joining a growing list of companies concerned about sensitive internal information being leaked through AI.



Confidentiality woes



"We may
use your
content..."

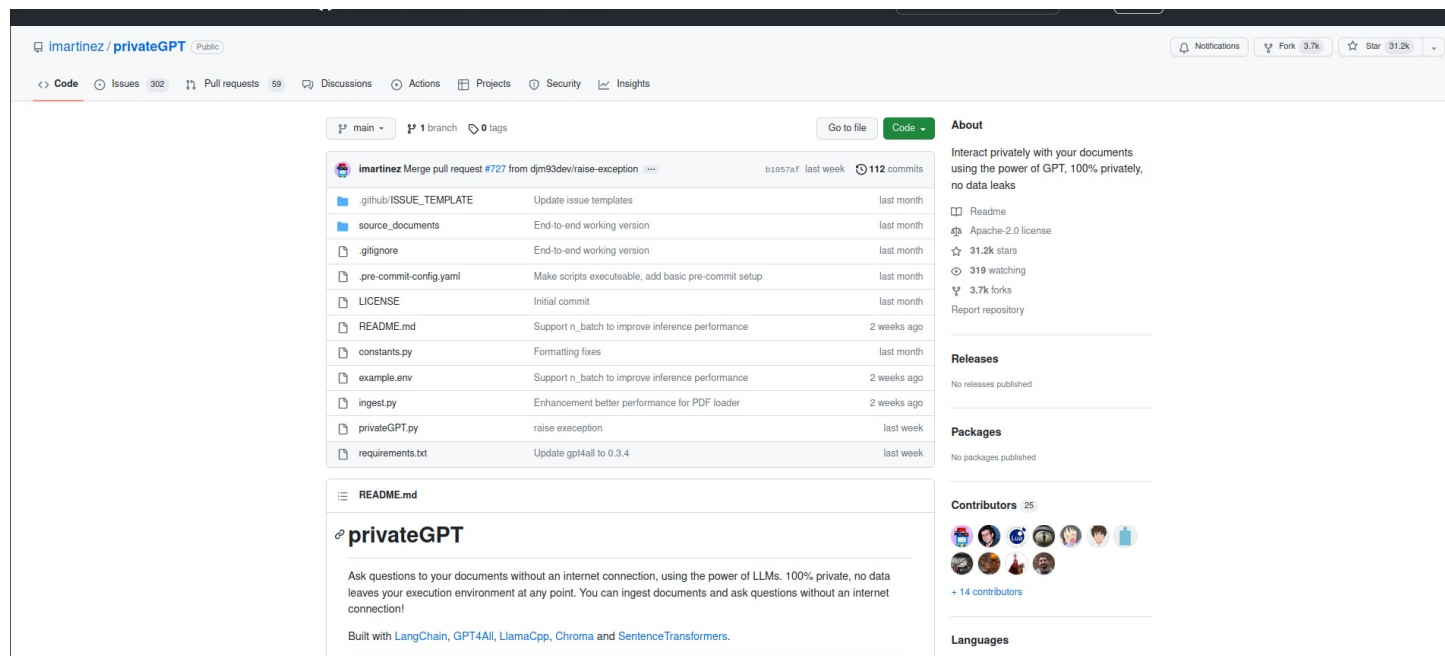
(c) Use of Content to Improve Services. We do not use Content that you provide to or receive from our API ("API Content") to develop or improve our Services. We may use Content from Services other than our API ("Non-API Content") to help develop and improve our Services. You can read more here about how Non-API Content may be used to improve model performance. If you do not want your Non-API Content used to improve Services, you can opt out by filling out this form. Please note that in some cases this may limit the ability of our Services to better address your specific use case.

OpenAI T&C

A free software solution?

privateGPT

Ask questions to your documents without an internet connection, using the power of LLMs. 100% private, no data leaves your execution environment at any point. You can ingest documents and ask questions without an internet connection!



A peculiar risk: the end of e-evidence as we know it

Deep fakes, AI-generated
images, AI-generated
speech...



The Texan way

"All attorneys and pro se litigants appearing before the Court must, together with their notice of appearance, file on the docket a certificate attesting either that no portion of any filing will be drafted by generative artificial intelligence (such as ChatGPT, Harvey.AI, or Google Bard) or that any language drafted by generative artificial intelligence will be checked for accuracy, using print reporters or traditional legal databases, by a human being."

"Unbound by any sense of duty, honor, or justice, such programs act according to computer code rather than conviction, based on programming rather than principle"

The screenshot shows the official website of Judge Brantley Starr, Northern District of Texas, United States District Court. The browser is Mozilla Firefox. The website features a navigation bar with links to Court Information, Judges, Rules & Orders, Filing, Resources, FAQs, Attorneys, Jurors, Pro Se, and Forms & Records. The main content area is titled "Judge Brantley Starr" and includes contact information: Chambers: 214-753-2160, Courtroom Deputy: Kevin Frye 214-753-2346, Court Reporter: Kelli Ann Willis 214-753-2654, and Courtroom Technology. The address is 1100 Commerce Street, Room 1528, Dallas, TX 75242. The website also displays a "Mandatory Certification Regarding Generative Artificial Intelligence" section, which states that all attorneys and pro se litigants must file a certificate attesting to the use of generative AI. The certificate must state either that no portion of the filing was drafted by generative AI or that any language drafted by generative AI was checked for accuracy using print reporters or traditional legal databases. The website also includes a "COVID-19 Procedures" section and a "Copy Requirements for Electronic Filing" section.

Judge Brantley Starr
Home » Judges » District Judges

Chambers: 214-753-2160
Courtroom Deputy: Kevin Frye 214-753-2346
Court Reporter: Kelli Ann Willis 214-753-2654
Courtroom Technology

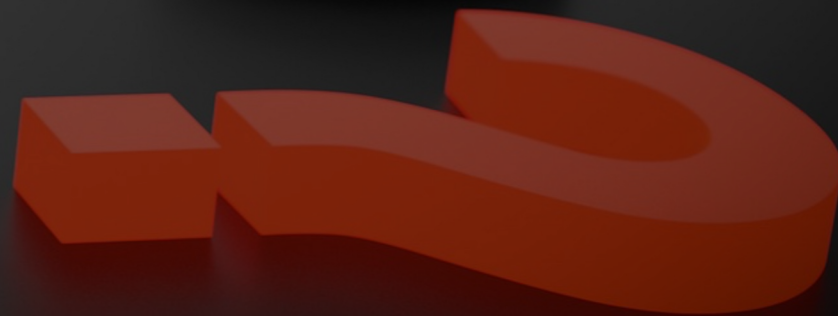
1100 Commerce Street, Room 1528
Dallas, TX 75242
Courtroom: 1525
Case Letter Designation: (X)

Judge Specific Requirements
Mandatory Certification Regarding Generative Artificial Intelligence
All attorneys and pro se litigants appearing before the Court must, together with their notice of appearance, file on the docket a certificate attesting either that no portion of any filing will be drafted by generative artificial intelligence (such as ChatGPT, Harvey.AI, or Google Bard) or that any language drafted by generative artificial intelligence will be checked for accuracy, using print reporters or traditional legal databases, by a human being. These platforms are incredibly powerful and have many uses in the law: form divorces, discovery requests, suggested errors in documents, anticipated questions at oral argument. But legal briefing is not one of them. Here's why. These platforms in their current states are prone to hallucinations and bias. On hallucinations, they make stuff up—even quotes and citations. Another issue is reliability or bias. While attorneys swear an oath to set aside their personal prejudices, biases, and beliefs to faithfully uphold the law and represent their clients, generative artificial intelligence is the product of programming devised by humans who did not have to swear such an oath. As such, these systems hold no allegiance to any client, the rule of law, or the laws and Constitution of the United States (or, as addressed above, the truth). Unbound by any sense of duty, honor, or justice, such programs act according to computer code rather than conviction, based on programming rather than principle. Any party believing a platform has the requisite accuracy and reliability for legal briefing may move for leave and explain why. Accordingly, the Court will strike any filing from a party who fails to file a certificate on the docket attesting that they have read the Court's judge-specific requirements and understand that they will be held responsible under Rule 11 for the contents of any filing that they sign and submit to the Court, regardless of whether generative artificial intelligence drafted any portion of that filing. **A template Certificate Regarding Judge-Specific Requirements is provided here.**

COVID-19 Procedures
Effective 6/12/2020: COVID-19 Standing Rules for Proceedings in the Courtroom of the Honorable Brantley Starr in the Earle Cabell Federal Building and United States Courthouse

Copy Requirements for Electronic Filing
Effective 3/24/2020 – Judge Starr has temporarily suspended the requirement for a judge's copy of a document electronically filed. A judge's copy of a document electronically filed is not required.

Questions?





Giovanni Battista Gallus — www.array.eu — gallus@array.law - @gbgallus