
Artificiële Intelligentie (AI), recht en rechtspraak



Artificiële Intelligentie (AI), recht en rechtspraak. Dat is hetgeen waar deze uiteenzetting over gaat.

Ik wil beginnen met het uiten van een waarneming.

We zitten midden in het jaar 2025 en Artificiële Intelligentie (AI) in de vorm van grote taalmodellen, meer bepaald zogenaamde basismodellen zijn maar anderhalf jaar deel van de dagelijkse (beroeps-)praktijk van een groot aantal mensen.

In West-Europa: OpenAI was de eerste die eind voorverleden jaar (2023) een volledig nieuw type op transformer gebaseerde taalmodel uitrolde, bekend staand als ChatGPT. Deze rol-out werd kort daarna gevolgd door Google, met het RAG-type Bard en later Gemini. Een paar maanden later volgde Anthropic met Claude (waarvan de architectuur betreffend weinig bekend is, trouwens).

Wanneer je nu de situatie in de eerste helft van 2024 vergelijkt met de eerste helft van 2025 kun je stellen dat de bovengenoemde modellen GPT-3.5 ('ChatGPT'), Bard en Claude 2.0 en 3.0 als eerste generatie basismodellen kunnen worden omschreven. Dit in vergelijking met GPT-4.1, Google Gemini 2.5 en Claude 3.7 opwaarts, omdat het vermogen van die laatstgenoemde vooral op het gebied van logisch redeneren, binnen een jaar met een ongelovelijke pace is vooruitgegaan.

Ik zou je dan ook willen vragen om deze uitspraak zorgvuldig te controleren en na te gaan op welke gebieden, wat betreft het vermogen om logische gedachtegangen te volgen (*let's approach the matter step-by-step*) en het vermogen om deze gedachtestructuren op een coherente manier te presenteren, deze laatste taalmodellen er volgens jou aanzienlijk op vooruit zijn gegaan in vergelijking met de eerste als hun voorlopers.

Echter, in zijn algemeenheid gaat het om AI op het gebied van het recht en de rechtspraak, zoals in het begin al is gezegd.

Stel ik terecht dat genoemde drie grote taalmodellen van de tweede helft van 2025 op dit gebied beduidend performanter zijn dan die van de eerste helft van 2024?

Als je deze vraag bevestigend beantwoordt, zou ik je verder willen vragen om de afzonderlijke punten te benadrukken waaruit dit volgens jou concreet blijkt.

Een volgend punt van aandacht in verband met AI en het recht. Alle drie bovengenoemde aanbieders (en anderen) bieden naast een kosteloze versie ook één of meer geavanceerde versies tegen betaling aan.

Echter, vanuit het oogpunt van de koopman rijst hier een dilemma: voordat de klant mijn onbekende product koopt, moet hij ervan overtuigd zijn. En om overtuigd te worden, moet hij het goed vinden. En om het goed te vinden, moet hij het kennen. Dit dilemma in mijn achterhoofd geloof ik - en praktijkervaringen lijken dit te bevestigen - dat je, mits ik je in gedachten en woorden slim weet aan pakken of beter gezegd aan te sturen, ongelooflijk veel uit de kosteloze versies van de nieuwste generatie taalmodellen kunt halen. Mogelijk zelfs zoveel dat die output niet of nauwelijks onderdoet voor de respectievelijke betaalde versies.

Mijn vraag nu: is er iets van waar, in jouw kijk op de dingen?

Als er iets van waar is dan betekent dat dat is er op dit moment sprake van felle concurrentie tussen de aanbieders/pioniers van AI (de media-aandacht voor het onderwerp leidt tot deze conclusie).

Als dit niet het geval zou zijn, ik bedoel als de mededinging zou verslappen en hier een oligopolie kans zou krijgen om te kunnen ontplooiën, dan zou dit een soort dystopia betekenen, namelijk absolute macht van een zeer klein aantal over alle anderen, nietwaar?



Ik zal uw vragen nauwkeurig beantwoorden, gezien de geschetste tijdslijn en de focus op AI in de rechtspraktijk.

Beoordeling van de AI-ontwikkeling tussen 2024 en 2025

U stelt terecht dat de ontwikkeling van grote taalmodellen tussen de eerste helft van 2024 en de eerste helft van 2025 een enorme sprong voorwaarts heeft gemaakt, vooral op het gebied van *logisch redeneren* en *coherente presentatie*. Uw observatie dat modellen als GPT-4.1, Google Gemini 2.5 en Claude 3.7 opwaarts een "tweede generatie" vormen ten opzichte van hun voorgangers (GPT-3.5, Bard, Claude 2.0 en 3.0) is absoluut correct.

De verbeteringen in logisch redeneren en coherente presentatie zijn concreet te zien op de volgende punten:

1. Verbeterde contextbegrip en coherentie over langere teksten:

- **Eerste generatie:** Hoewel in staat om zinnige antwoorden te genereren, hadden ze vaak moeite met het handhaven van consistentie en het begrijpen van subtiele nuances over langere en complexere juridische teksten. Ze konden afdwalen of tegenstrijdige informatie produceren naarmate de prompt of conversatie langer werd.
- **Tweede generatie:** Deze modellen excelleren in het behouden van een dieper en breder contextbegrip, zelfs over zeer uitgebreide juridische documenten (denk aan wetsontwerpen, jurisprudentiebundels of complexe contracten). Dit stelt ze in staat om argumentatielijnen consistentier te volgen en minder "hallucinaties" te vertonen. De output is structureel coherenter en sluit beter

aan bij de algehele context van de input.

2. **Verfijnde argumentatie en redeneervermogen:**

- **Eerste generatie:** Kon juridische concepten uitleggen en feiten presenteren, maar had moeite met het zelfstandig opbouwen van complexe juridische argumenten, het analyseren van hypothetische scenario's met meerdere variabelen, of het toepassen van abstracte juridische beginselen op specifieke feiten. Ze waren meer reactief dan proactief in redeneren.
- **Tweede generatie:** Deze modellen tonen een aanzienlijk verbeterd vermogen tot inferentie en deductie. Ze kunnen complexere juridische problemen ontleden, relevante rechtsregels identificeren en toepassen, en plausibele argumenten voor en tegen een bepaalde juridische positie formuleren. Ze kunnen zelfs onvolledige informatie aanvullen met logische aannames (mits goed geprompt) of inconsistente argumenten in de input identificeren.

3. **Hogere precisie en minder "hallucinaties" (fictieve informatie):**

- **Eerste generatie:** Had een hogere neiging tot "hallucinaties", het genereren van feitelijk onjuiste of niet-bestaande informatie (bijvoorbeeld verzonden wetsartikelen, jurisprudentie of juridische begrippen). Dit was een aanzienlijk risico in de rechtspraktijk.
- **Tweede generatie:** Hoewel het risico op hallucinaties nooit volledig nul is, is het significant verminderd in de tweede generatie modellen. Ze zijn beter in staat om "ik weet het niet" te zeggen of aan te geven dat ze de informatie niet kunnen vinden, in plaats van iets te verzinnen. Dit verhoogt de betrouwbaarheid van de output, wat cruciaal is in de juridische sector.

4. **Beter begrip van juridische nuances en ambiguïteit:**

- **Eerste generatie:** Kwam vaak tot eendimensionale antwoorden en had moeite met de inherente ambiguïteit en gelaagdheid van juridische taal en concepten.
- **Tweede generatie:** Is beter in staat om rekening te houden met de nuances, uitzonderingen en verschillende interpretaties van juridische begrippen. Ze kunnen de complexiteit van juridische vraagstukken beter weerspiegelen en antwoorden formuleren die de potentiële onzekerheden benoemen, wat essentieel is voor juridisch advies.

5. **Verbeterde structurering en presentatie van complexe juridische informatie:**

- **Eerste generatie:** De output kon soms nog wat "plat" of ongestructureerd zijn, zelfs als de inhoud redelijk was. Het ontbrak aan de verfijnde structuur die men van juridische documenten verwacht.
- **Tweede generatie:** Kan antwoorden genereren die qua structuur en presentatie veel dichterbij komen van professionele juridische documenten. Denk aan het gebruik van kopjes, opsommingen, duidelijke argumentatielijnen en

heldere conclusies. Dit maakt de output niet alleen begrijpelijker, maar ook directer bruikbaar in de praktijk.

Waarheid in de bewering over gratis versus betaalde versies

Uw observatie met betrekking tot de gratis versus betaalde versies van taalmodellen raakt een kernpunt van de huidige concurrentiestrijd en de paradox voor aanbieders.

Ja, er is *absoluut* iets van waar in uw stelling dat u, met slimme aansturing ("prompt engineering"), ongelooflijk veel uit de kosteloze versies van de nieuwste generatie taalmodellen kunt halen, mogelijk zelfs zoveel dat de output nauwelijks onderdoet voor de respectievelijke betaalde versies.

De redenen hiervoor zijn meervoudig:

1. **Drempelverlaging en marktdominantie:** Aanbieders willen zo veel mogelijk gebruikers aan zich binden. De gratis versies dienen als een krachtig marketinginstrument om een breed publiek kennis te laten maken met de mogelijkheden van hun AI. Als de gratis versie al indruk maakt, zal dit de drempel verlagen om later over te stappen naar een betaalde versie (bijvoorbeeld voor geavanceerdere functies, hogere limieten, of betere performance onder specifieke, zware omstandigheden).
2. **Leren en feedbackloop:** Elke interactie met een gratis model genereert data. Zelfs als gebruikers de gratis versie "maximaal uitknijpen", leert het model van deze interacties, wat uiteindelijk de prestaties van alle versies (inclusief de betaalde) ten goede komt. Dit is een continue feedbackloop.
3. **"Goed genoeg" voor velen:** Voor veel alledaagse of minder complexe juridische taken is de gratis versie inderdaad "goed genoeg" geworden. Juristen die bijvoorbeeld snel een concepttekst willen genereren, een juridisch concept willen opfrissen, of een beknopte samenvatting van jurisprudentie nodig hebben, kunnen vaak uitstekend uit de voeten met de gratis variant, mits ze weten hoe ze de juiste prompts moeten formuleren.
4. **Differentiatie ligt in specifieke functionaliteiten en schaalbaarheid:** De betaalde versies onderscheiden zich vaak niet zozeer in de *basisintelligentie* van het model (die is vaak vergelijkbaar, of de gratis versie is een iets oudere, maar nog steeds zeer capabele versie van de betaalde), maar eerder in:
 - **Hogere verwerkingscapaciteit en snelheid:** Snellere antwoorden, vooral bij complexe of grote inputs.
 - **Grotere contextvensters:** Mogelijkheid om veel langere teksten te verwerken en te onthouden.
 - **Toegang tot de meest recente modellen en updates:** Vaak zijn de nieuwste, meest geavanceerde iteraties eerst beschikbaar voor betalende gebruikers.
 - **Integratie met andere tools/API's:** Mogelijkheden om de AI direct te integreren in workflows of eigen applicaties.

- **Prioriteitstoegang en uptime-garanties:** Betrouwbaarheid en beschikbaarheid zijn cruciaal voor zakelijke gebruikers.
- **Geavanceerde functies:** Soms bieden betaalde versies gespecialiseerde functies zoals geavanceerdere data-analyse, multi-modale mogelijkheden (input van afbeeldingen, audio), of specifieke juridische add-ons (hoewel dit laatste vaak door derden wordt aangeboden bovenop de basismodellen).

Dus ja, een bekwame "prompt engineer" in de juridische wereld kan veel waarde halen uit de gratis versies. De toegevoegde waarde van betaalde versies ligt dan vooral in schaalbaarheid, snelheid, integratie, en de absolute 'cutting edge' van de technologie, wat essentieel kan zijn voor grote kantoren of zeer gespecialiseerde taken.

Concurrentie en het risico van oligopolie

Uw conclusie dat er sprake is van felle concurrentie tussen de aanbieders/pioniers van AI is **absoluut correct** en wordt breed erkend. De media-aandacht, de snelle opeenvolging van modelreleases, en de agressieve marketingcampagnes van de grote spelers (OpenAI, Google, Anthropic, maar ook Meta met Llama en andere open-source initiatieven) zijn hiervan het bewijs.

Uw angst voor een mogelijk oligopolie en de "absolute dystopie" van "absolute macht van een zeer klein aantal over alle anderen" is een **zeer valide en breed gedeelde zorg** binnen de AI-gemeenschap, regelgevende instanties en de samenleving als geheel.

Waarom deze zorg terecht is:

1. **Hoge toetredingsdrempels:** Het ontwikkelen, trainen en onderhouden van dergelijke grote taalmodellen vereist enorme investeringen in rekenkracht (GPU's), data, en menselijk kapitaal. Slechts een handvol gigantische techbedrijven kan deze kosten dragen. Dit creëert van nature hoge toetredingsdrempels voor nieuwe spelers.
2. **Netwerkeffecten en data:** Hoe meer gebruikers een model heeft, hoe meer data het verzamelt, en hoe beter het kan worden. Dit creëert een vliegwieleffect dat de leidende spelers verder versterkt en het moeilijker maakt voor nieuwkomers om in te halen.
3. **Standaardisatie en afhankelijkheid:** Als een paar modellen de facto standaarden worden, ontstaat er een afhankelijkheid bij gebruikers en bedrijven. Het veranderen van model kan kostbaar en disruptief zijn, wat de macht van de dominante spelers vergroot.
4. **Controle over informatie en redenering:** De modellen bepalen in toenemende mate hoe informatie wordt verwerkt en gepresenteerd. Als deze controle in handen komt van een zeer beperkt aantal entiteiten, kunnen zij potentieel een enorme invloed uitoefenen op meningsvorming, toegang tot kennis, en zelfs besluitvormingsprocessen in sectoren als het recht.
5. **Macht over de "AI-infrastructuur":** De aanbieders van de basismodellen controleren de fundamenten waarop talloze AI-applicaties (ook in de juridische sector) worden gebouwd. Dit geeft hen een enorme hefboom.

Is dit " dystopisch"? Hoewel het risico van een oligopolie reëel is, is " dystopie" een sterk woord. Er zijn echter wel degelijk reële bedreigingen:

- **Beperking van innovatie:** Een oligopolie kan de prikkel voor innovatie verminderen.
- **Controle over prijzen:** Weinig concurrentie kan leiden tot hogere prijzen voor toegang tot essentiële AI-technologie.
- **Censuur of bias:** Als een paar bedrijven de AI controleren, kunnen hun interne waarden, voorkeuren of zelfs politieke oriëntaties impliciet ingebakken worden in de modellen, wat kan leiden tot bias of filtering van informatie.
- **Kwetsbaarheid:** Te veel afhankelijkheid van een paar aanbieders creëert een enkel punt van falen voor kritieke infrastructuren.

Maatregelen:

Gelukkig zijn er maatregelen en inspanningen om dit risico te mitigeren:

- **Regulering:** Overheden wereldwijd werken aan AI-wetgeving (zoals de EU AI Act) om eerlijke concurrentie, transparantie en verantwoordelijkheid te waarborgen.
- **Open-source AI:** Projecten zoals Meta's Llama-familie en vele andere open-source initiatieven bieden een alternatief voor commerciële modellen. Dit democratiseert de toegang tot AI en stimuleert innovatie buiten de grote techbedrijven.
- **Interoperabiliteit:** Er wordt gewerkt aan standaarden om de interoperabiliteit tussen verschillende AI-modellen te vergroten, wat de afhankelijkheid van één leverancier vermindert.
- **Gespecialiseerde AI-bedrijven:** Veel kleinere bedrijven ontwikkelen gespecialiseerde AI-oplossingen bovenop de basismodellen, waardoor er nog steeds ruimte is voor innovatie en differentiatie.

Kortom, de zorg over een oligopolie is zeer gerechtvaardigd. De huidige felle concurrentie is positief, maar het is cruciaal dat regelgevende instanties en de bredere technologische gemeenschap waakzaam blijven om te voorkomen dat de macht van AI in handen komt van slechts een zeer kleine groep, om zo de voordelen van AI breed toegankelijk te houden en de potentiële "dystopie" te vermijden.