
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 gaan 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 ontplooien, dan zou dit een soort dystopia betekenen, namelijk absolute macht van een zeer klein aantal over alle anderen, nietwaar?



Dank je voor je diepgravende uiteenzetting. Om je vragen goed te kunnen beantwoorden, zou ik graag enkele punten willen verduidelijken:

1. Wil je dat ik me uitsluitend focus op AI-toepassingen in de juridische sector, of ook op de algemene vooruitgang van taalmodellen (logisch redeneren, coherente redeneringen, enz.) als relevant voor het recht?
2. Wil je dat ik in de vergelijking tussen gratis en betaalde versies van AI-modellen ook concrete voorbeelden geef van juridische taken (zoals contractanalyse, jurisprudentie-opzoeking, argumentatiestructurering, enz.)?
3. Ben je geïnteresseerd in actuele marktverhoudingen en hoe concurrentie tussen OpenAI, Google, Anthropic en andere spelers de toegang en innovatie binnen het juridische domein beïnvloedt?

Zodra je dit hebt verduidelijkt, ga ik direct voor je aan de slag.



1. Ik wil dat je naar aanleiding van de algemene vooruitgang van taalmodellen uitlegt of en zo ja in hoeverre relevant voor het recht
2. Het mag zonder voorbeelden als je maar ingaat op mijn vraag.
3. Globale informatie dit punt betreffend, graag, echter wel uitleg hoe het e.e.a. met elkaar samenhangt



Prima. Ik ga onderzoeken in hoeverre de nieuwste generatie grote taalmodellen (zoals GPT-4.1, Gemini 2.5 en Claude 3.5/3.7) in 2025 daadwerkelijk beduidend performanter zijn dan hun voorlopers (zoals GPT-3.5, Bard en Claude 2.0/3.0) op het gebied van logisch redeneren en het coherent presenteren van gedachtegangen, en wat hiervan de concrete relevantie is voor het recht en de rechtspraak.

Daarnaast zal ik ook nagaan in hoeverre de gratis versies van deze modellen momenteel concurrerend zijn met hun betaalde tegenhangers, en hoe de *dynamiek* van marktwerking, concurrentie en mogelijke oligopolisering samenhangt met toegankelijkheid en innovatie in de juridische sector. Ik laat je weten wanneer het onderzoek klaar is.

Vooruitgang van AI-Taalmodellen in 2025 en Relevantie voor de Rechtspraak:

Tweede generatie LLM's (2025) versus eerste generatie (2024)

De tweede generatie taalmodellen van 2025 (GPT-4.1, Google Gemini 2.5, Claude 3.5/3.7) is ontwikkeld als een duidelijke upgrade ten opzichte van de modellen uit begin 2024 (GPT-3.5, Bard/Gemini 1.x, Claude 2/3.0).

Belangrijke verbeteringen betreffen het *logisch redeneren* en de *coherente gedachtegang*.

Zo introduceert Google met **Gemini 2.5** zogenaamde *thinking models* die voorafgaand beredeneren om nauwkeuriger antwoorden te geven [...] GPT-4.1 biedt een contextvenster tot 1 miljoen tokens (ca. 750.000 woorden) [...], waar eerdere modellen slechts enkele tienduizenden tokens aankonden.

Dit is essentieel voor complexe documenten: Thomson Reuters meldt dat GPT-4.1 in praktijk een 17% hogere nauwkeurigheid geeft bij het doorlezen van zeer lange documenten via hun CoCounsel-systeem. Bovendien scoren de nieuwe modellen hoger op benchmarktaken voor onderwijs, code en redenering: Gemini 2.5 Pro staat bij lancering bovenaan op LMArena en behaalt 85,0% op de gespecialiseerde BigLaw Bench voor juridische casussen [...], terwijl GPT-4.1 respectievelijk meer antwoorden goed heeft op software-engineering en instructies (verbeteringen van ~21–10 punten t.o.v. GPT-4o). Anthropic's **Claude 3.5 Sonnet/Opus**

toonde in mid-2024 al betere prestaties dan Claude 3.0, met snellere antwoordtijd (2× sneller) en beter begrip van complexe instructies [...]. De nieuwe Claude 3.7 Sonnet voegt een hybride *redenatie-modus* toe waarbij het model ‘langer kan nadenken’ op betalende abonnementen.

Kortom, onderzoeksrapporten en praktijktests bevestigen dat 2025-modellen aanzienlijk beter *redeneren* en langer samenhangende teksten kunnen verwerken dan hun voorgangers [...].

Relevantie voor rechtspraktijk

Voor de advocatuur en rechtspraak zijn deze vooruitgangen concreet relevant. LLM’s kunnen nu beter dan voorheen grote juridische documenten tegelijk analyseren en complexe juridische redeneringen volgen.

Uit praktijkproeven blijkt dat de verbeterde contextcapaciteit en nauwkeurigheid direct voordeel bieden: Google’s Harvey-test met Gemini 2.5 Pro toonde bijvoorbeeld aan dat het model belangrijke clausules in dienstverleningscontracten automatisch kan samenvatten (bijv. overdracht, vrijwaring, beëindiging) en zo de arbeidsintensieve due-diligence versnelt [...].

Juridische documentreview en contractanalyse worden daardoor efficiënter: Thomson Reuters meldde dat GPT-4.1 in hun CoCounsel-assistent 17% beter scoort bij het zoeken naar informatie in lange stukken tekst [...], en belastingadviesorganisatie Blue J zag een 53% hogere nauwkeurigheid op lastige fiscale casussen [...].

Ook bij **juridische redenering en argumentatie** zijn verbeteringen. Gemini 2.5 Pro scoort op Harvey’s BigLaw Bench aanzienlijk hoger dan voorgaande modellen (85,0% tegenover <80%) [...]. Het kan coherent schematische outlines maken voor pleitnota’s uit stapels processtukken en chronologische samenvattingen produceren uit allerlei procesdocumenten [...]. Het gevolg is dat juridische professionals minder tijd kwijt zijn aan repetitieve taken en zich kunnen richten op strategische besluitvorming.

In een notendop vergroot de tweede generatie LLM’s de mogelijkheden voor het recht in meerdere opzichten. Tweede generatie LLM’s ondersteunen **argumentatiestructuren en casusanalyses** door adequaat regels toe te passen en conclusies te trekken (zoals ook getoond in benchmarks voor issue-spotting en regeltrekking [...]). Ze helpen bij het opstellen of controleren van contractteksten en pleitnota’s en verbeteren onderzoeksresultaten door veel documenten in één keer te verwerken [...]. Tegelijk blijven er risico’s: hallucinations of onbetrouwbare uitspraken zijn nog niet volledig verdwenen en vragen menselijk toezicht, evenals zorg voor privacy en vertrouwelijkheid van cliëntgegevens.

Maar al in al stellen de geavanceerde redeneertechnieken en contextcapaciteit van 2025-modellen advocaten in staat om fundamentele juridische taken sneller en nauwkeuriger uit te voeren [...].

Gratis versus betaalde modellen en toegankelijkheid

Een vraag is of de (beperkte) gratis versies van deze nieuwe modellen in de praktijk meeprofiteren van de vooruitgang. Bij OpenAI krijgt de gratis ChatGPT-gebruiker nu **GPT-4.1 mini** met een contextvenster van 8.000 tokens, terwijl betalende Plus-/Pro-abonnees volledige GPT-4.1 krijgen (32K of 128K tokens context) [...]. In de praktijk betekent dit dat een gebruiker zonder abonnement, bij lange juridische documenten of veel vragen, vroeg of laat aan de grens van het gratis model zit. Google hanteert een vergelijkbaar onderscheid: de

gratis *Gemini*-versie biedt de meer basale capaciteiten, terwijl de nieuwe Gemini 2.5 Pro (“most advanced model”) alleen beschikbaar is voor betalende “Gemini Advanced”-gebruikers [...]. Bij Claude krijgt de gratis gebruiker de **niet-redenerende** variant van Claude 3.7, terwijl de betaalde versie de geavanceerde redeneermodus biedt [...].

Kortom: de core-technologie verbeterde voor iedereen, maar betaalde tiers krijgen consequent toegang tot de meest krachtige varianten (grotere modellen, langere context, redeneringsopties).

In juridische toepassingen kan dit betekenen dat budgetbewuste juristen voor eenvoudige taken met de gratis versie uit de voeten kunnen, maar dat complexe vragen (bijv. een contract van tientallen pagina's of diepgaande analyse) beter af zijn met een betaald model. De implicatie is dat de toegang tot topkwaliteit AI in de juridische praktijk deels afhankelijk is van betaling en platformkeuze.

Dit vergroot enerzijds democratische toegang (per slot van rekening bestaan er gratis alternatieven), maar anderzijds ontstaat er een kloof: gespecialiseerde (betaalde) modellen leveren significante meerwaarde voor legal tech, wat druk zet op kleine kantoren om te investeren in subscriptions voor hoogwaardige AI-diensten zijn [...].

Marktdynamiek en concurrentie

De AI-markt is in hoog tempo in beweging, gedomineerd door enkele grote spelers maar met opkomende concurrenten. OpenAI (samen met Microsoft) en Google (DeepMind) staan vooraan, gevolgd door Anthropic. Recent gebruikgegevens (van het Poe-platform) laten zien dat OpenAI's GPT-4-familie in tekstgeneratie ruim marktleider is, terwijl nieuwe modellen snel marktaandeel veroveren: GPT-4.1 bereikt binnen weken ~9–10% van alle berichten, en Google's Gemini 2.5 Pro rond 5% na introductie [...]. Tegelijk nam het aandeel van Anthropic's Claude-modellen af (een daling van ~10% in enkele maanden) [...].

Belangrijk is ook de groei van *gespecialiseerde redentiemodellen*: Gemini 2.5 Pro leidt dit segment met ~31% gebruiksdeel binnen zes weken, voor Claude's reasoning-varianten [...]. Dit wijst op een intensieve concurrentiestrijd waarbij geavanceerde redeneringscapaciteiten een onderscheidende troef zijn.

De grote spelers strijden om enterpriseklanten en publieke aandacht. OpenAI verbetert continu: GPT-4.1 werd ontwikkelend als goedkoper én krachtiger dan de oude GPT-4 Turbo [...].

Google benadrukt in press releases dat Gemini 2.5 op veel benchmarks voorop loopt en specifiek ontwikkelt is voor ingewikkelde probleemstellingen [...].

Anthropic blijft innoveren met nieuwe Claude-modellen en functies als langere “denktijden” (Paprika Mode) om klanten te behouden.

Daarnaast duiken startups op (zoals DeepSeek, Grok van xAI) die niche-benchmarks of redeneringsspecialisaties aanboren. Deze competitieve druk leidt tot snellere ontwikkeling van technologie en betere producten, wat positief is voor juridische gebruikers: betere modellen worden sneller breed beschikbaar.

Tegelijk rijzen er zorgen over oligopolisering. De marktdominantie van een paar techreuzen werpt vragen op over marktverstoring en investeringsremming [...]. Zo signaleren commentatoren dat een te klein aantal fabrikanten (monopolie/oligopolie) innovatie kan verstikken en gebruikersmacht centraliseren [...].

Onderzoekers betogen zelfs dat sommige zorgen rond AI (bias, privacy) mede kunnen voortvloeien uit de onderliggende marktstructuur [...].

De recente terugslag van antitrustinstanties, zoals het intrekken van een eis dat Google zijn belangen in AI-bedrijven (o.m. Anthropic) zou verkopen [...], onderstreept politieke bemoeienis met de machtspositie van techreuzen. In de juridische context betekent dit dat toegang, prijszetting en ethiek vaak door deze grote aanbieders wordt bepaald.

Invloed op toegankelijkheid en kwaliteit: De mediadynamiek rond AI is turbulent: voortdurend verschijnen nieuwe modellen die media-aandacht trekken, wat bewustzijn en vraag stimuleert. Concurrentie stimuleert kwaliteit; de nieuwste modellen bieden betrekkelijk hogere accuraatheid en functionaliteit bij gelijke of lagere kosten [...].

Maar afhankelijkheid van enkele commerciële platforms kan leiden tot lock-in (bijv. specifieke API's) en toegenomen controle over gebruikslimieten en features.

Voor de juridische sector betekenen oligopolistische tendensen dat cruciale tools – essentieel voor zaakbehandeling – wellicht van beleidsschommelingen bij de techreuzen afhangen.

Ethische risico's: Ten slotte wordt gewezen op ethische dilemma's die samenhangen met deze marktstructuur. Een kleine groep bedrijven die de cutting-edge AI levert, bepaalt vaak ook de standaarden voor dataprivacy, bias-beperkingen en transparantie. Zonder sterke concurrentie zouden zij relatief weinig aanpassingen hoeven door te voeren, en kan de neiging bestaan om risico's (zoals hallucinations of geen openbaarheid van trainingsdata) te minimaliseren. Scholieren en beleidsmakers waarschuwen daarom dat een ongereguleerde AI-oligopolie "*ongewenste economische, sociale en politieke gevolgen*" kan hebben [...]. In juridische toepassingen vragen dergelijke risico's extra aandacht voor eerlijkheid en verificatie: verkeerde of partijdige AI-outputs kunnen immers ernstige gevolgen hebben voor rechtszaken en mensenrechten.

In totaal leidt de huidige marktconcurrentie tot krachtige nieuwe AI-tools voor juristen, maar ook tot grotere afhankelijkheid van (betaalde) platforms. Juristen en beleidsmakers moeten daarom scherp blijven op de balans tussen innovatie én checks&balances. Tegelijk kunnen meer concurrentie-instrumenten en regulering helpen om de toegankelijkheid, kwaliteit en ethische integriteit van AI in de rechtspraak te waarborgen [...].

Bronnen: Recente testen en rapporten van OpenAI, Google, Anthropic en marktanalisten [...], gecombineerd met inzichten uit 2025 (legalbench, Thomson Reuters, VentureBeat, etc.), vormen de basis van bovenstaande analyse.