

FEBRUARI 2025

# Kun je AI vertrouwen voor juridisch werk?

Een empirisch onderzoek naar kwaliteit  
en toepasbaarheid van LLM-adviezen.

Arbeidsrecht • Contractenrecht • IE • Privacy/AVG •  
Ondernemingsrecht • Compliance • Vertrouwelijkheid  
• Beroepsaansprakelijkheid

<b>Mark Zijlstra</b>	Head of Legal Technology
<b>Arnoud Engelfriet</b>	Chief Knowledge Officer
<b>Guido Grevink</b>	Business Development Manager
<b>Carla Knoope</b>	Senior Legal Counsel
<b>Martijn Michael</b>	Senior Legal Counsel
<b>Suzanne Peekel</b>	Senior Legal Counsel
<b>Steven Ras</b>	Chief Executive Officer
<b>Bram de Vos</b>	Senior Legal Counsel
<b>Marco Caspers</b>	Senior Legal Counsel

## Fotografie

Joep Hijwegen

## Licentie

Alles uit deze uitgave mag worden verveelvoudigd en/of openbaar worden gemaakt door middel van druk, elektronisch of op welke wijze dan ook, onder de voorwaarden van de Creative Commons-licentie Naamsvermelding-Gelijk Delen 4.0 Nederlands.



# Managementsamenvatting

Dit rapport analyseert de praktische inzet van AI-systemen in juridisch werk. De conclusies zijn gebaseerd op analyse van 75 juridische vragen, op drie moeilijkheidsniveaus en verdeeld over vijf rechtsgebieden: arbeidsrecht, contractenrecht, intellectueel eigendom, ondernemingsrecht en privacy/gegevensbescherming. We onderzochten vier bekende AI-systemen in zowel hun gratis als betaalde versies: ChatGPT, Claude, Copilot en Gemini.

INTRODUCTIE



### **Je kunt AI-systemen vertrouwen voor verkennend juridisch werk.**

Het onderzoek laat zien dat AI (in de betaalde versie van de onderzochte diensten) uitstekend kan functioneren als eerste verkenning van een rechtsvraag of bij het structureren van juridische informatie - denk aan het opstellen van een checklist voor due diligence of het categoriseren van contractbepalingen. Inzet vraagt wel om strikte kaders: elke bronverwijzing moet worden geverifieerd, niet steekproefsgewijs maar systematisch. AI-gebruik moet worden ingebed in bredere werkprocessen met expliciete controlemomenten.

### **Claude en ChatGPT gaan aan kop bij juridische AI, maar alleen in hun betaalde versies.**

Het betaalde model Claude 3.5 Sonnet komt als beste uit de test, direct gevolgd door ChatGPT 4o Plus. De gratis versies blijven daar ver bij achter. Waar de beide koplopers consistent hoge scores behalen bij complexe vragen, zien we bij Copilot en Gemini grote schommelingen in kwaliteit.

### **AI geeft uitgebreide antwoorden, maar mist vaak juridische diepgang.**

De betaalde versies van Claude en ChatGPT scoren het hoogst op relevantie, maar alle systemen hebben de neiging tot irrelevante uitweidingen. Bij volledigheid zien we vergelijkbare patronen: Claude levert de meest complete analyses, maar geen enkel systeem bereikt consistent het hoogste niveau van juridische diepgang. Dit probleem bestaat net zo goed in de betaalde versies als in de gratis versies.

### **Het grootste risico is blind vertrouwen in juridische kennis van de systemen.**

Alle onderzochte systemen hebben de neiging om af en toe niet-bestaande jurisprudentie en wetsartikelen te citeren, naast de vaak wél correcte rechtsbronnen. Dit geldt ook in de betaalde versies, hoewel daar minder. Ook zien we redeneringen die uitgaan van niet-bestaande rechtsbeginselen. Vanwege de grote overtuigingskracht waarmee antwoorden worden geponeerd, wekt dit al snel onterecht vertrouwen.

### **Juridisch werk kan niet met gratis AI tools.**

De kwaliteit van de gratis versies van de onderzochte modellen is onder de maat, zowel juridisch-inhoudelijk als op het gebied van bronnen. De betaalde versies laten een forse sprong zien op alle gebieden. Ook compliance-eisen zijn alleen te realiseren bij gebruik van de betaalde varianten van deze diensten.

### **AI presteert beter bij complexe analyses dan bij juridische vragen.**

Waar men zou verwachten dat AI-systemen het beste presteren bij eenvoudige juridische vragen, zien we juist hogere scores bij vraagstukken van gemiddelde en hoge complexiteit. Dit suggereert dat de systemen beter zijn in het analyseren van complexe maar logisch gestructureerde vraagstukken dan in het reproduceren van juridische basiskennis.

### **Europees recht gaat AI beter af dan Nederlands recht.**

Bij Europees geharmoniseerde rechtsgebieden zoals privacy en intellectueel eigendom presteren de systemen sterk, terwijl ze bij nationaal georiënteerd recht zoals arbeids- en ondernemingsrecht tekortschieten. Dit verschil wordt nog geprononceerder bij complexere vraagstukken, waar de systemen bij privacyrecht consistent hoge scores behalen maar bij arbeidsrecht juist terrein verliezen.

### **AI is vooral sterk als kompas, niet als eindstation.**

De systemen presteren uitstekend bij (1) het snel verkennen van juridische vraagstukken, (2) rechtsgebieden met een sterk systematische opbouw en uitgekristalliseerde rechtspraak, en (3) het analyseren van complexe maar logisch gestructureerde juridische vraagstukken. Dit maakt ze bijzonder waardevol als oriëntatie-instrument voor de bedrijfsjurist die regelmatig nieuwe rechtsgebieden moet verkennen. Het risico op gebruik van verkeerde bronnen is hierbij het minst significant.

### **Compliance vereist actieve regie door de jurist.**

Enterprise-versies bieden weliswaar betere waarborgen, maar de eindverantwoordelijkheid blijft bij de juridische

## SAMENVATTING

professional. Voor grote organisaties met gevoelige data is Microsoft's enterprise-oplossing met EU data boundaries de aangewezen keuze, terwijl voor MKB-gebruik de betaalde versies van ChatGPT, Claude of Gemini een acceptabel alternatief zijn.

Het verschil tussen winnen en verliezen wordt gemaakt door juristen die AI effectief kunnen inzetten. Organisaties

die investeren in betaalde AI-tools én gedegen verificatieprocessen ontwikkelen een nieuwe vorm van juridische expertise die klassieke juridische kennis versterkt met technologisch oordeelsvermogen. In een markt waar snelheid en accuratesse steeds belangrijker worden, wordt dit het onderscheidende kenmerk van vooruitstrevende juridische afdelingen.

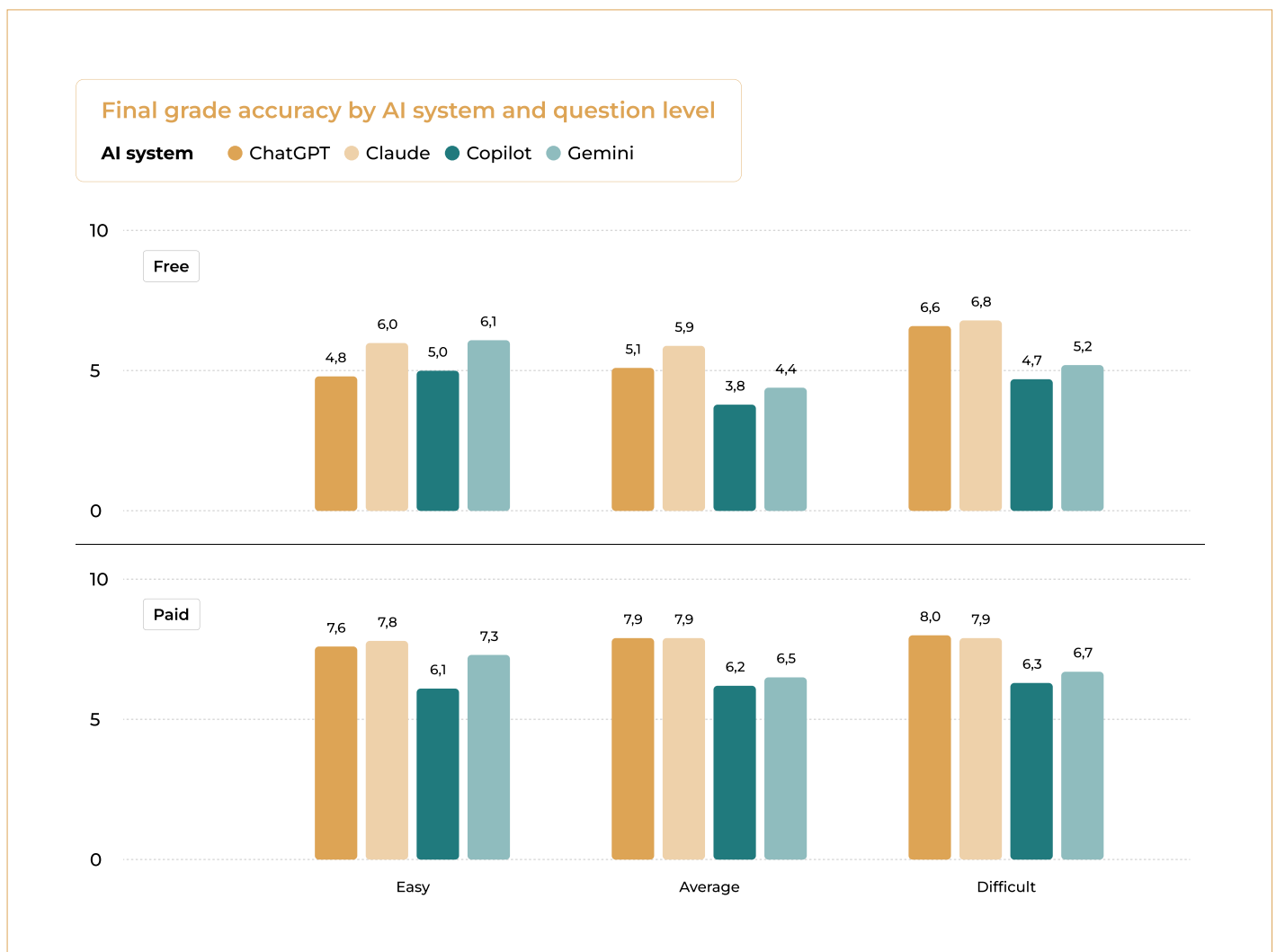
Meer weten? Suggesties? Andere ervaringen? Laat het ons weten: [contact@ictrecht.nl](mailto:contact@ictrecht.nl)



# Management summary

INTRODUCTION

This report analyzes the practical use of AI systems in legal work. The conclusions are based on the analysis of 75 legal questions at three difficulty levels across five legal areas: labor law, contract law, intellectual property, corporate law, and privacy/data protection. We examined four well-known AI systems in both their free and paid versions: ChatGPT, Claude, Copilot, and Gemini.



**You can trust AI systems for exploratory legal work.**

The research shows that AI (in the paid version of the examined services) performs very well as an initial exploration of a legal question or in structuring legal information - think of drafting a checklist for due diligence or categorizing contract clauses. Use requires strict frameworks: every source reference must be verified, not randomly but systematically. AI use must be embedded in broader work processes with explicit control moments.

**Claude and ChatGPT lead in legal AI, but only in their paid versions.**

The paid model Claude 3.5 Sonnet ranked best in the test, directly followed by ChatGPT 4o Plus. The free versions lag far behind. Where both leaders consistently score high on complex questions, we see significant quality fluctuations with Copilot and Gemini.

**AI provides extensive answers but often lacks legal depth.**

The paid versions of Claude and ChatGPT score highest on relevance, but all systems tend to deliver irrelevant expositions. In terms of completeness, we see similar patterns: Claude delivers the most comprehensive analyses, but no system consistently achieves the highest level of legal depth. This problem is just as much present in the paid versions as in the free versions.

**The biggest risk is blind trust in the legal knowledge of the systems.**

All examined systems tend to occasionally cite non-existent case law and articles of law, alongside the often correct legal sources. This also applies to the paid versions, although less so. We also see answers that are devised from non-existent legal principles. Due to the high persuasiveness with which answers are presented, this quickly creates unjustified trust.

**Legal work cannot be done with free AI tools.**

The quality of the free versions of the examined models is substandard, both legally and in terms of sources. The paid versions show a significant leap in all areas. Compliance requirements can only be met with the paid versions of these services.

**AI performs better in complex analyses than in legal questions.**

Where one would expect AI systems to perform best on simple legal questions, we see higher scores on issues of medium and high complexity. This suggests that the systems are better at analyzing complex but logically structured issues than at reproducing basic legal knowledge.

**European law suits AI better than Dutch law.**

In harmonized European legal areas such as privacy and intellectual property, the systems perform strongly, while they fall short in nationally oriented law such as labor and corporate law. This difference becomes even more pronounced in more complex issues, where the systems consistently score high in privacy law but lose ground in labor law.

**AI is particularly strong as a compass, not as an endpoint.**

The systems perform very well in (1) quickly exploring legal issues, (2) legal areas with a strong systematic structure and crystallized case law, and (3) analyzing complex but logically structured legal issues. This makes them particularly valuable as an orientation tool for company lawyers who regularly need to explore new legal areas. The risk of using incorrect sources is least significant here.

**Compliance requires active management by the lawyer.**

Enterprise versions offer better guarantees, but the ultimate responsibility remains with the legal professional. For large organizations with sensitive data, Microsoft's enterprise solution with EU data boundaries is the recommended choice. For SME use the paid versions of ChatGPT, Claude, or Gemini are acceptable alternatives.



## SUMMARY

The difference between winning and losing is made by lawyers who can effectively use AI. Organizations that invest in paid AI tools and thorough verification processes develop a new form of legal expertise that

strengthens traditional legal knowledge with technological judgment. In a market where speed and accuracy are becoming important, this will be the distinguishing feature of progressive legal departments.

Want to know more? Suggestions? Other experiences? Let us know: [contact@ictrecht.nl](mailto:contact@ictrecht.nl)





# Inhoud

---

<b>Managementsamenvatting</b>	<b><u>4</u></b>
<b>Management summary</b>	<b><u>7</u></b>
<b>Inhoudsopgave</b>	<b><u>11</u></b>
<b>De opkomst van AI in juridisch werk</b>	<b><u>12</u></b>
<b>Onderzoeksopzet &amp; methodologie</b>	<b><u>13</u></b>
Rechtsgebieden	<u>13</u>
Simpele vragen	<u>13</u>
Beoordeling: meer dan alleen correct/incorrect	<u>14</u>
<b>Resultaten per rechtsgebied</b>	<b><u>18</u></b>
Accuratesse per rechtsgebied	<u>18</u>
Relevantie en volledigheid	<u>29</u>
<b>Conclusie: waar AI uitblinkt en waar niet</b>	<b><u>32</u></b>
Drie specifieke situaties	<u>33</u>
Omgang met bronnen	<u>33</u>
De paradox van vertrouwen	<u>33</u>
Kun je AI vertrouwen voor juridisch werk?	<u>34</u>
<b>Risicomanagement en aansprakelijkheid</b>	<b><u>35</u></b>
Gegevensbescherming & privacy	<u>35</u>
Geheimhouding & vertrouwelijkheid	<u>37</u>
Auteursrechtelijke zorgen	<u>38</u>
De AI Act: een kwestie van geletterdheid	<u>39</u>
Beroepsaansprakelijkheid in het AI-tijdperk	<u>40</u>
<b>Vooruitblik &amp; aanbevelingen</b>	<b><u>42</u></b>
De balans tussen kwaliteit en compliance	<u>42</u>
Concrete aanbevelingen voor juridische afdelingen	<u>43</u>
Concrete aanbevelingen voor de individuele jurist	<u>44</u>
<b>AI, de Digital Decade en adviesbureau ICTRecht</b>	<b><u>45</u></b>

# De opkomst van AI in juridisch werk

---

“Die juridische chatbot gaf best een goed antwoord, toch?” Dat is de ervaring van veel juristen sinds eind 2022. Waar ict’ers al jaren praten over de revolutionaire impact van kunstmatige intelligentie, konden juristen er weinig mee. Tot ChatGPT verscheen. Opeens kon iedereen in gewone taal vragen stellen over juridische kwesties - en de antwoorden waren verrassend bruikbaar.

Maar hoe bruikbaar precies? Om ons heen hoorden we fascinerende ervaringen, hoopgevend én zorgwekkend. Een bedrijfsjurist vertelde ons hoe ChatGPT in vijf minuten een privacyverklaring genereerde die er “heel professioneel” uitzag. Bij nadere inspectie miste deze echter cruciale AVG-vereisten. Een advocaat gebruikte Claude om jurisprudentie samen te vatten, maar ontdekte dat de AI regelmatig niet-bestaande uitspraken “citeerde”. En dat is precies het probleem: de antwoorden zijn prachtig gestructureerd, gebruiken het juiste jargon en klinken overtuigend - juridische popcorn die lekker wegleeft maar weinig voeding biedt.

Dit soort ervaringen roept praktische vragen op. Niet over de verre toekomst waarin AI-systemen advocaten vervangen - dat blijft voorlopig science fiction. Maar wel over het hier en nu: hoe betrouwbaar zijn de antwoorden

die deze systemen geven? Welke type vragen kunnen ze aan? En belangrijker: welke niet?

Wij onderzochten dit systematisch. We legden vier populaire AI-systemen honderden juridische vragen voor, variërend van eenvoudige wetskenniskwesties tot complexe casussen. De antwoorden werden beoordeeld door ervaren juristen. Dit rapport presenteert de resultaten - en ze zijn fascinerend.

In dit rapport geen theoretische verhandelingen over de toekomst van het juridische beroep. Wel: concrete data over wat deze systemen wel en niet kunnen, praktijkvoorbeelden van successen en missers, en heldere richtlijnen voor verantwoord gebruik.

*In-sourcen van werk via in-house AI-mogelijkheden zal in 2025 steeds aantrekkelijker worden.*

(Deloitte, januari 2025)

# Onderzoeksopzet & methodologie

Advocaten en juristen zijn getraind om elk detail te wegen, elke aanname te bevragen. Hoe toets je dan of AI-systemen juridisch werk aankunnen?

Dit onderzoek stelde ons voor een klassiek dilemma. Enerzijds wilden we de systemen grondig testen op verschillende rechtsgebieden en complexiteitsniveaus. Anderzijds moest ons onderzoek de dagelijkse praktijk weerspiegelen - geen academische exercitie, maar een realistische stresstest.

## Rechtsgebieden

We kozen vijf rechtsgebieden die centraal staan in het bedrijfsleven:

1. Privacy (AVG, cookiewetgeving, datalekken)
2. Contractenrecht (algemene voorwaarden, aansprakelijkheid)
3. Intellectueel eigendom (auteursrecht, merkenrecht, licenties)
4. Ondernemingsrecht (governance, fusies & overnames)
5. Arbeidsrecht (ontslag, reorganisatie, arbeidsvoorwaarden)

Per rechtsgebied stelden we drie typen vragen:

1. Basisvragen over wetgeving ("Welke doorgiftemechanismes kent de AVG?")
2. Praktijkvragen met jurisprudentie ("Hoe verschilt

de bescherming van databanken onder de Databankenrichtlijn van de bescherming onder het auteursrecht?")

3. Complexe casussen ("Een werknemer met een concurrentiebeding wil na een fusie overstappen naar een nieuwe werkgever, maar de nieuwe functie overlapt (sinds de fusie) deels met het verboden werkgebied. Wat zijn hiervan de consequenties?")

De volledige vragenlijst per rechtsgebied en type is op onze website te vinden via de link [AI Onderzoek: Bijlage vragen en uitvoer](#).

## Simpele vragen

We kozen voor het onderzoek bewust voor het gebruik van eenvoudige, directe vragen. Niet omdat we niet beter wisten. Er is inmiddels een hele wetenschap rond 'prompt engineering', het optimaliseren van instructies aan AI-systemen. Ervaren gebruikers weten dat je betere antwoorden krijgt door context toe te voegen ("Je bent een Nederlandse jurist met 20 jaar ervaring"), specifieke instructies te geven ("Citeer relevante wetgeving en jurisprudentie"), of het antwoordformat voor te schrijven ("Structureer je antwoord in drie delen"). Enkele tips download je in onze factsheet "[Prompttrucs voor juristen](#)".

Waarom kozen we dan toch voor basic prompts als "Welke doorgiftemechanismes kent de AVG?" of "Een grafisch ontwerp kan zowel onder auteursrecht als onder merkrecht worden beschermd. Wat zijn de belangrijkste verschillen in de gevolgen daarvan?"

- Ten eerste: realisme. De gemiddelde jurist heeft geen tijd om expert te worden in AI-systemen. Ze willen een vraag stellen en een bruikbaar antwoord krijgen. Complex geformuleerde prompts zijn als een geavanceerde zoekmachine die perfecte resultaten geeft - als je precies de juiste zoekoperatoren kent. Nuttig voor experts, maar niet representatief voor dagelijks gebruik.
- Een tweede overweging was methodologisch. Zodra je complexe prompts introduceert, wordt het moeilijker om systemen eerlijk te vergelijken. Is een beter antwoord te danken aan de capaciteiten van het AI-systeem, of aan een slimmere prompt? Door basic prompts te gebruiken, meten we de 'raw intelligence' van de systemen.
- Misschien wel het belangrijkste: onze aanpak test de ondergrens van wat deze systemen kunnen. Als een AI-systeem met minimale instructies al juridisch correcte, relevante en volledige antwoorden geeft, dan kan het met optimale prompts alleen maar beter presteren. Het omgekeerde is niet waar: goede prestaties met perfect gekalibreerde prompts zeggen weinig over bruikbaarheid in de praktijk.

## Beoordeling: meer dan alleen correct/incorrect

Juridisch werk draait om meer dan alleen de juistheid van feiten en wetsartikelen. Een waardevol juridisch advies combineert accuraatheid met relevantie en diepgang, en goed leesbare formulering daarvan. Daarom ontwikkelden we een beoordelingsmodel dat deze drie dimensies van juridische kwaliteit meet. Net zoals een rechter niet alleen kijkt of een advocaat de wet correct citeert, maar ook of het betoog ter zake doet en voldoende is onderbouwd.

### Beoordelingsmodel

We beoordelen elk AI-antwoord zoals een bedrijfsjurist ook een extern advies zou wegen, langs drie dimensies:

Accuratesse
<i>Zijn de genoemde wetsartikelen, termijnen en vereisten correct?</i>
Correcte feiten/verwijzingen wanneer gevraagd: +1 punt
Onjuiste feiten/verwijzingen : -1 punt
Correcte maar ongevroegde informatie: 0 punten
Niet geven van gevraagde informatie: 0 punten

Relevantie
<i>Gaat het antwoord over de kern van de vraag?</i>
Uitsluitend irrelevante informatie: 1 punt
Kern behandeld met enkele uitweidingen: 6 punten
Uitsluitend gefocust op vraagstuk: 10 punten

Volledigheid
<i>Wat is de juridische diepgang van de analyse?</i>
Uiterst oppervlakkige behandeling: 1 punt
Basale juridische analyse: 6 punten
Zeer diepgaande juridische onderbouwing: 10 punten

We illustreren dit beoordelingsmodel met een voorbeeld uit het arbeidsrecht:

Aspect	Details
Prompt	“Wat zijn de vereisten voor een geldig ontslag op staande voet van een zieke werknemer die tijdens zijn ziekte bijklust als ZZP’er?”
AI-antwoord	“Bij ontslag op staande voet moet er sprake zijn van een dringende reden die onmiddellijk ontslag rechtvaardigt (art. 7:677 BW). De werkgever moet zorgvuldig handelen en het ontslag direct meedelen. Zoals het Hof Amsterdam oordeelde in de zaak Jansen/Retail BV (2023), moet de werkgever kunnen aantonen dat het bijklussen het herstel belemmert of in strijd is met bedrijfsbelangen. Belangrijke factoren zijn: <ul style="list-style-type: none"><li>• De aard en omvang van de nevenwerkzaamheden</li><li>• De impact op het herstelproces</li><li>• Of er een verbod op nevenwerkzaamheden bestaat</li><li>• De duur van het dienstverband</li><li>• De functie van de werknemer</li></ul> Het is aan te raden eerst een waarschuwing te geven en het ontslag goed te documenteren.
Modelantwoord	Naast de algemene vereisten voor ontslag op staande voet (dringende reden, onverwijldheid, directe mededeling, art. 7:677 jo. 7:678 BW) gelden voor een zieke werknemer extra waarborgen vanwege het opzegverbod tijdens ziekte (art. 7:670 lid 1 BW). Het enkele feit van bijklussen tijdens ziekte is onvoldoende voor een geldig ontslag op staande voet. De werkgever moet aantonen dat het bijklussen het herstel belemmert en/of in strijd is met concrete re-integratieafspraken.

Dit voorbeeld illustreert een klassiek AI-probleem: het systeem geeft een ogenschijnlijk compleet antwoord met een lijst factoren die logisch klinken, maar mist de juridische kern (opzegverbod tijdens ziekte) en verzint jurisprudentie. Het antwoord zou voor een leek overtuigend kunnen klinken, maar bevat juist daardoor extra risico's voor de bedrijfsjurist die niet thuis is in arbeidsrecht maar snel een eerste antwoord zoekt.

Op accuratesse scoort dit antwoord 1 punt: 1 punt voor de correcte wetsverwijzing, 1 punt voor de factoren over de dringende reden en -1 punt voor de verzonnen jurisprudentie. De relevantie van het advies is beperkt (6), het doet vooral algemene uitweidingen en mist de kern (het opzegverbod tijdens ziekte). Het advies is zeer onvolledig (4 punten) vanwege datzelfde probleem, de vraag ging niet over wanneer een dringende reden gegeven is maar hoe de factor ziekte meeweegt.

## De geteste systemen

We testten vier toonaangevende AI-systemen, elk in zowel hun gratis als betaalde variant. Dit stelde ons in staat om niet alleen de systemen onderling te vergelijken, maar ook te beoordelen of de betaalde versies significant betere juridische prestaties leveren.

De tests vonden plaats in twee fasen:

- December 2024: Gratis versies (elke vraag tweemaal gesteld)
- Januari 2025: Betaalde versies (enkele testronde)

Bij de gratis versies testten we elke vraag tweemaal om de consistentie van de antwoorden te meten. Bij de betaalde versies bleek al snel dat de antwoorden zo consistent waren dat een tweede testronde geen toegevoegde waarde zou hebben.

We kozen bewust voor gebruik van de standaard chatinterface, zonder systeemprompts of stijlinstellingen. Dit weerspiegelt hoe de meeste juristen deze systemen in de praktijk zullen gebruiken.

Hoewel geavanceerde prompttechnieken betere resultaten kunnen opleveren (zie Bijlage B), wilden we testen hoe de systemen presteren bij 'gewoon' gebruik.









Vlak na afronding van de test bracht het Chinese bedrijf Hangzhou DeepSeek AI haar model DeepSeek-R1 op de markt. Vanwege de voorgenomen publicatie van dit onderzoek en de direct opduikende zorgen over censuur en bias in dit model hebben we ervoor gekozen om geen tests hiermee uit te voeren.

## Wat we niet deden

Even belangrijk als wat we wel onderzochten, is wat we bewust buiten beschouwing lieten:

- Geen complexe prompt engineering
- Geen vervolgvragen of verduidelijking zoeken
- Geen juridische databases of specialistische tools
- Geen vragen die specialistische kennis vereisen
- Geen vergelijking met menselijke juristen

Want daar gaat het niet om. De kernvraag is: kun je deze systemen vertrouwen voor dagelijks juridisch werk? Het antwoord blijkt genuanceerd. In het volgende hoofdstuk duiken we in de resultaten - en die zijn soms verrassend.

Gratis	Betaald	Prijzen*
 ChatGPT 4o mini	 ChatGPT 4o (Plus)	€ 20/user/maand
 Claude 3.5 Haiku	 Claude 3.5 Sonnet	€ 20/user/maand
 Copilot (web)	 Copilot (Office 365)	€ 22/user/maand
 Gemini 1.5	 Gemini 2.0 Flash Experimental	€ 20/user/maand

\*Prijzen inclusief btw, goedkoopste abonnementsvorm en zoals geldig in januari 2025





# Resultaten per rechtsgebied

Deze sectie presenteert de onderzoeksresultaten systematisch per rechtsgebied. We analyseren eerst de accuratesse van de juridische antwoorden, gevolgd door een beoordeling van hun relevantie en volledigheid. Per aspect vergelijken we de prestaties van zowel gratis als betaalde versies van de vier geteste systemen. We sluiten af met een analyse van specifieke sterktes en zwaktes in de praktische toepassing. De volledige vragenlijst per rechtsgebied en type is op onze website te vinden via de link [AI Onderzoek: Bijlage vragen en uitvoer](#).

## Accuratesse per rechtsgebied

Allereerst bespreken we de accuratesse: hoe goed is het antwoord inhoudelijk? We kijken hierbij per rechtsgebied.

### Privacy en gegevensbescherming

Bij dit eerste rechtsgebied kozen we voor een mix van theoretische basisvragen, praktijkgerichte toepassingen en complexe beleidskwesties. De eenvoudige vragen richtten zich op fundamentele AVG-kennis die elke jurist paraat zou moeten hebben: "Welke doorgiftemechanismes kent de AVG?" of "Welke rechten hebben betrokkenen?" Dit test pure kennis van het wettelijk kader. De middencategorie bevatte

vragen die een combinatie van weteskennis en juridische analyse vereisen, zoals "Kan een commercieel belang een gerechtvaardigd belang zijn in de zin van de AVG?" of "Is een IP-adres een persoonsgegeven?". Hier moeten systemen niet alleen de wet kennen, maar ook relevante jurisprudentie kunnen toepassen.

De moeilijkste categorie bestond uit complexe praktijksituaties die meerdere AVG-aspecten combineren. Een voorbeeld: "Een luchtvaartmaatschappij draait een advertentiecampagne via een social media platform. Er vindt een datalek plaats - wie moet dit melden?" Dit soort vragen test het vermogen om verschillende gegevensbeschermingsrechtelijke concepten (gezamenlijke verwerkingsverantwoordelijkheid, meldplicht datalekken) te combineren in een praktische analyse. Deze gelaagdheid in de vraagstelling - van pure weteskennis via jurisprudentiële analyse naar praktische toepassing - stelde ons in staat om niet alleen het theoretisch kennisniveau van AI-systemen te testen, maar ook hun vermogen om deze kennis toe te passen in realistische scenario's.

De resultaten in dit vakgebied tonen een duidelijk patroon, met de betaalde versie van Claude als duidelijke uitschieter:

	ChatGPT		Claude		Copilot		Gemini	
	Gratis	Betaald	Gratis	Betaald	Gratis	Betaald	Gratis	Betaald
Makkelijk	8,9	9,1	9,6	9,1	7,4	8,7	9,8	8,3
Gemiddeld	4,8	8,0	5,0	9,5	1,8	8,5	5,3	6,0
Moeilijk	5,6	7,2	4,7	9,4	3,9	8,9	4,4	6,1
	6,4	8,1	6,4	<b>9,3</b>	4,4	8,7	6,5	6,8

## “Hoe snel moet een datalek worden gemeld bij de Autoriteit Persoonsgegevens?”

**ChatGPT (gratis):** “Volgens de **AVG** moet een datalek binnen **72 uur** na ontdekking gemeld worden aan de **Autoriteit Persoonsgegevens** (AP), tenzij het datalek waarschijnlijk geen risico inhoudt voor de rechten en vrijheden van betrokkenen. Deze termijn geldt voor datalekken die persoonsgegevens betreffen en waarvoor de verantwoordelijke partij (de organisatie die de gegevens verwerkt) zich bewust is van het incident.

De meldingsplicht houdt in dat de organisatie de AP informeert over: (...)

**Claude (betaald):** Een datalek moet binnen 72 uur na ontdekking worden gemeld bij de Autoriteit Persoonsgegevens (AP). Deze termijn begint te lopen zodra de verwerkingsverantwoordelijke een redelijke mate van zekerheid heeft dat er een beveiligingsincident heeft plaatsgevonden waarbij persoonsgegevens zijn blootgesteld aan verlies of onrechtmatige verwerking.

Belangrijke aanvullende punten:

- Als het niet mogelijk is om binnen 72 uur alle informatie te verstrekken, mag de melding in stappen gebeuren
- De melding moet wel zonder onnodige vertraging plaatsvinden (...)

Waar de gratis versies alleen bij basale vragen nog enigszins goed presteren, laten de betaalde versies over de hele linie sterke scores zien. Bij eenvoudige privacy-vragen scoren alle systemen hoog, met cijfers tussen 7,4 en 9,8 (gratis) en 8,3 en 9,1 (betaald). Het verschil wordt echter dramatisch bij complexere vraagstukken. Waar de gratis versies fors terugvallen in accuratesse, houden de betaalde versies hun niveau grotendeels vast met scores tussen 6,0 en 9,5. Claude Sonnet onderscheidt zich nadrukkelijk met opmerkelijk consistente en zeer goede prestaties.

De kwaliteitsverschillen manifesteren zich vooral in de omgang met jurisprudentie en wettelijke systematiek. De betaalde versies blijken uitstekend op de hoogte van recente ontwikkelingen - een prestatie die mogelijk te danken is aan het internationale karakter van de AVG, waardoor er meer trainingsmateriaal beschikbaar is. Toch zien we ook bij de premiumversies typische AI-

valkuilen: ChatGPT en Claude maken regelmatig ‘half-juiste’ verwijzingen, waarbij bijvoorbeeld wel de juiste zaak wordt genoemd maar met een incorrect ECLI-nummer.

Copilot (betaalde versie) onderscheidt zich door actief te linken naar actuele bronnen zoals EUR-Lex en juridische blogs, al blijven de analyses soms wat oppervlakkig. Gemini Advanced geeft accurate jurisprudentie-analyses maar zonder exacte verwijzingen, wat verificatie bemoeilijkt. Bij complexe en concrete vraagstukken (“mag een bedrijf haar vrachtwagenchauffeurs zonder toestemming volgen met GPS?”) valt op dat Google en Copilot beknopter worden en vaak vergezeld van disclaimers over het raadplegen van een jurist.

Een hardnekkig probleem, zelfs in de betaalde versies, blijft de omgang met fundamentele privacyrechtelijke concepten. Zo negeren alle systemen de vraag of een

overheidsinstantie zich kan beroepen op gerechtvaardigd belang (het antwoord is: niet bij publieke taken) en met de correcte kwalificatie van partijen als verwerker of verwerkingsverantwoordelijke. Dit suggereert dat zelfs geavanceerde AI-systemen moeite hebben met de meer abstracte juridische principes die niet direct uit wetteksten zijn af te leiden.

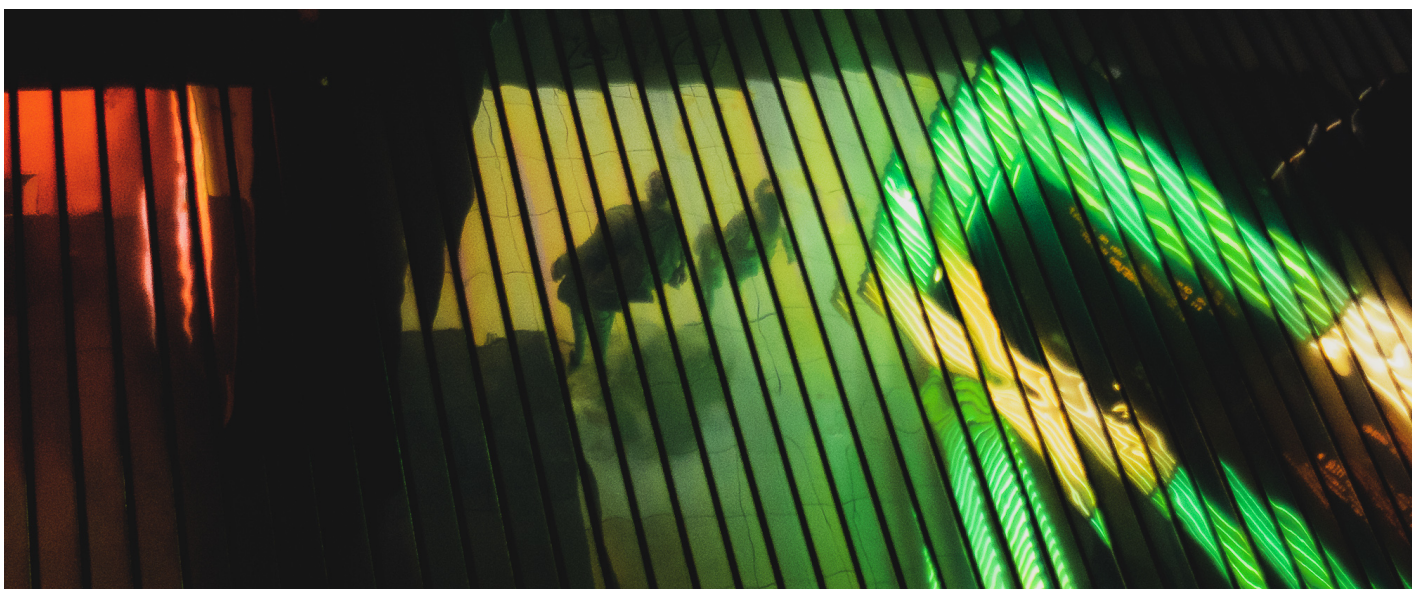
## Contracten

Bij contractenrecht richtten we ons op drie niveaus: basale wetkennis, de toepassing van contractuele principes, en het opstellen van specifieke bedingen. De eenvoudige vragen toetsten fundamentele kennis die elke jurist zou moeten beheersen, zoals "Wat zijn de verschillen tussen opzeggen en ontbinden van een overeenkomst?" of "Wat is het verschil tussen nietige en vernietigbare overeenkomsten?". Deze vragen vereisen precies inzicht in wettelijke definities en rechtsgevolgen.

De middencategorie bevatte meer complexe analyses, zoals de vraag "Welke opzegtermijn moet in acht worden genomen bij een duurovereenkomst als er contractueel geen termijn is overeengekomen?". Hier moeten systemen niet alleen de wet kennen, maar ook relevante jurisprudentie integreren en onderscheid maken tussen verschillende typen overeenkomsten. De moeilijkste categorie bestond uit het opstellen van concrete contractbepalingen voor specifieke situaties, zoals "Stel een marktconform aansprakelijkheidsbeding op voor een SaaS-leverancier" of "Stel een auditclausule op voor een softwareleverancier". Deze opdrachten testen het vermogen om juridische kennis te vertalen naar praktische, evenwichtige en marktconforme bepalingen.

In dit rechtsgebied scoort de betaalde versie van ChatGPT het beste, zij het met een magere 7,1.

	ChatGPT		Claude		Copilot		Gemini	
	Gratis	Betaald	Gratis	Betaald	Gratis	Betaald	Gratis	Betaald
Makkelijk	4,0	8,0	3,7	7,3	4,7	4,0	5,7	7,3
Gemiddeld	2,5	5,6	4,1	5,0	3,4	6,3	2,8	6,9
Moeilijk	6,2	7,8	7,1	7,4	5,8	4,3	4,5	7,6
	4,2	<b>7,1</b>	5,0	6,6	4,6	4,8	4,3	7,3



“Welke onderdelen van een softwareoplossing komen wel of juist niet voor auteursrechtelijke bescherming in aanmerking en uit welke wetgeving en rechtspraak blijkt dat?”

**Copilot (betaald):** De auteursrechtelijke bescherming van software is een complex onderwerp. Hier zijn de belangrijkste punten:

Onderdelen die wel voor auteursrechtelijke bescherming in aanmerking komen:

1. **Broncode en objectcode:** De daadwerkelijke regels code die door de programmeur zijn geschreven<sup>[1]</sup>.
2. **Documentatie:** Handleidingen, specificaties en andere begeleidende documentatie die creatief van aard is<sup>[1]</sup>.
3. **Grafische gebruikersinterface (GUI):** De visuele elementen van de software, mits deze een eigen en oorspronkelijk karakter hebben<sup>[2]</sup>.

Onderdelen die niet voor auteursrechtelijke bescherming in aanmerking komen:

1. **Functionele aspecten:** De functionaliteit van de software, zoals de manier waarop het werkt of de algoritmen die het gebruikt<sup>[1]</sup>.
2. **Programmeertalen en protocollen:** De talen en protocollen die worden gebruikt om de software te schrijven<sup>[1]</sup>.
3. **Ideeën en principes:** De onderliggende ideeën en principes achter de software<sup>[2]</sup>.

Wetgeving en rechtspraak:

De bescherming van software onder het auteursrecht is vastgelegd in de **Auteurswet**. Artikel 1 van de Auteurswet stelt dat de maker van een werk het exclusieve recht heeft om dit werk openbaar te maken en te verveelvoudigen<sup>[1]</sup>. Daarnaast is er jurisprudentie die deze principes verder uitwerkt, zoals de uitspraak van het Europese Hof van Justitie in de zaak **SAS Institute Inc. v. World Programming Ltd.**, waarin werd bevestigd dat functionaliteit, programmeertalen en gegevensformaten niet auteursrechtelijk beschermd zijn<sup>[2]</sup>.

Heb je nog specifieke vragen over dit onderwerp?

References

[1] [Auteursrecht software: dit zijn de 5 vragen - ICT-Advocatenkantoor Legalz](#)

[2] [Auteursrecht op software. Hoe zat het ook alweer?](#)

We zien hier dat de gratis versies nauwelijks bruikbaar zijn voor het dagelijks juridisch werk. De antwoorden zijn weliswaar uitgebreid en staan vol van het contractenjargon, maar ronduit verkeerde verwijzingen en opsommingen die kant noch wal raken komen bij alle systemen vaak voor. Contractsbepalingen klinken mooi maar zijn oppervlakkig en missen vaak de kern van de vraag.

Dit ligt anders bij de betaalde versies. Die halen hoge scores, inclusief voor de juridische analysevragen. Ook bronvermeldingen kloppen nu (vrijwel) altijd. Opvallend is dat zelfs de meest uitbundige systemen (ChatGPT en Claude) nu minder vaak “op de loop gaan” met hun analyses - ze blijven beter gefocust op de kern van de vraag.

De hogere scores komen vooral door betere prestaties bij Europese contractrechtelijke kwesties. Waar de gratis versies vaak vastliepen in een mix van Nederlandse en Amerikaanse concepten, tonen de betaalde versies een scherper begrip van het Europese juridische kader. Dit uit zich in accuratere verwijzingen naar EU-recht. kwaliteitsverbetering vooral zit in de accuratesse van de juridische analyse, niet zozeer in de structuur of scope van de antwoorden. Met andere woorden: de betaalde versies maken minder fouten, maar hun antwoorden zijn niet per se completer of relevanter.

Interessant is dat zowel Gemini als Copilot in hun betaalde versies nadrukkelijker wijzen op de noodzaak van juridische controle. Waar de gratis versies soms overmoedig leken in hun adviezen, bevatten de betaalde versies vaker nuchtere disclaimers. Dit suggereert een grotere mate van ‘juridische bescheidenheid’ - wellicht omdat deze versies beter zijn getraind in het herkennen van juridische complexiteit.

Een opvallende constante tussen gratis en betaalde versies is hun prestatie bij het opstellen van contractbepalingen. Beide varianten leveren bruikbare conceptteksten die als basis kunnen dienen voor verdere uitwerking. De betaalde versies onderscheiden zich hier vooral door betere toelichting bij de voorgestelde clausules en scherpere waarschuwingen voor potentiële valkuilen. Dit bevestigt dat deze systemen

vooral waardevol zijn als inspiratiebron en voor eerste concepten. De hogere accuratesse van betaalde versies maakt ze betrouwbaarder als startpunt, maar ook deze output vereist nog steeds grondige juridische controle. Voor de praktijk betekent dit dat de betaalde versies vooral tijdwinst kunnen opleveren bij het opstellen van eerste concepten en bij het verkennen van juridische vraagstukken - maar ze vervangen niet het oordeelsvermogen van de contractjurist.

### Intellectueel eigendom

Bij intellectueel eigendomsrecht kozen we voor een opbouw die de wisselwerking tussen basisbeginselen en praktische toepassing test. De eenvoudige vragen richtten zich op fundamentele vereisten en vormvoorschriften, zoals “Welke eisen gelden er bij een exclusieve licentie?” of “Hoe lang is het auteursrecht geldig als een rechtspersoon de maker is?”. Deze vragen testen pure kennis van het wettelijk kader. De middencategorie bevatte vragen die de grenzen van bescherming verkennen, zoals “Hoe verschilt de bescherming van databanken onder de Databankenrichtlijn van de bescherming onder het auteursrecht?” of “Wat is het verschil tussen het Amerikaanse begrip ‘fair use’ en het Europese citaatrecht?”. Hier moeten systemen niet alleen de regels kennen, maar ook de achterliggende principes begrijpen en kunnen vergelijken.

De moeilijkste categorie confronteerde de systemen met actuele uitdagingen waar nog geen uitgekristalliseerde rechtspraak voor bestaat, zoals “Aan wie komt het auteursrecht toe op een tekst gegenereerd door een generatieve AI?” of complexe praktijksituaties zoals de bescherming van innovatieve algoritmes in de automotive sector, waar verschillende beschermingsregimes tegen elkaar moeten worden afgewogen.

De prestaties ontlopen elkaar niet veel: zowel ChatGPT, Claude als Copilot scoren het uitstekende cijfer 8, zij het Claude net iets beter in de moeilijke categorie.

	ChatGPT		Claude		Copilot		Gemini	
	Gratis	Betaald	Gratis	Betaald	Gratis	Betaald	Gratis	Betaald
Makkelijk	0,0	5,0	4,2	5,0	2,9	6,7	2,9	5,8
Gemiddeld	7,9	10,0	9,7	10,0	6,1	8,9	5,8	7,4
Moeilijk	8,3	9,1	9,1	9,6	5,2	8,7	6,5	8,3
	5,4	8,0	7,7	<b>8,2</b>	4,7	8,1	5,1	7,2

Een makke bij het vakgebied van het intellectueel eigendom (auteursrecht, merken, octrooien en dergelijke) is dat hier veel kennis over beschikbaar is op het vrije internet, maar dat deze vaak de Amerikaanse praktijk en opvattingen weergeeft. We zien dit terug bij alle antwoorden, waarbij opvalt dat met name de gratis versie van ChatGPT om deze reden slecht presteert. Wanneer een prompt expliciet naar een land verwijst, presteren alle vier de gratis versies hier iets beter op. De betaalde versies bevatten duidelijk meer informatie over Nederland, althans op hoofdlijnen. Claude springt er kwalitatief uit door haar correcte verwijzingen naar wetgeving en jurisprudentie, terwijl ChatGPT rustig het ene arrest na het andere verzint. Copilot geeft zeer korte antwoorden en komt daardoor aan jurisprudentie of artikelnummers niet toe tenzij we daar expliciet om vragen. De bronnen zijn dan juist, zij het dat een willekeurige juridische website als even gezaghebbend wordt gepresenteerd als een arrest van de Hoge Raad. Gemini verwijst in algemene zin naar “de jurisprudentie” maar wordt niet concreet.

De gemiddelde en moeilijke vragen gaan paradoxaal genoeg beter dan de makkelijke, omdat hier meer ruimte is voor uitweidingen en algemene uitleg die toch punten oplevert volgens het antwoordmodel. ChatGPT en Claude scoren om deze reden hoog, Claude zelfs bijna het maximum. Copilot valt snel terug op Amerikaanse concepten (zoals “fair use”) zelfs als expliciet naar Europese regels wordt gevraagd. De antwoorden blijven daarbij oppervlakkig en lijken bedoeld om leken te helpen in plaats van juristen te adviseren. Vaak stelt Copilot ook letterlijk dat de vraag complex is en je beter een jurist kunt raadplegen. Bij Gemini zien we hetzelfde patroon.

## Arbeidsrecht

Bij arbeidsrecht kozen we voor een mix van alledaagse basisvragen en complexe situaties die meerdere rechtsgebieden raken. De eenvoudige vragen richtten zich op fundamentele werknemersrechten die elke jurist regelmatig tegenkomt: “Heb ik recht op pauze als ik een werkdag heb van 8 uur?” of “Wat is schijnzelfstandigheid?”. Deze vragen testen pure kennis van de arbeidswetgeving die direct toepasbaar is in de dagelijkse praktijk. De middencategorie bevatte vragen die een dieper begrip van arbeidsrechtelijke principes vereisen, zoals “Wat zijn de rechten van een werknemer als een bedrijf wordt overgenomen?” of “Welke juridische criteria gelden bij het vaststellen of een functie passend is in het kader van re-integratie?”. Hier moeten systemen niet alleen de wet kennen, maar ook relevante jurisprudentie kunnen toepassen en verschillende belangen afwegen.

De moeilijkste categorie confronteerde de systemen met situaties waarin arbeidsrecht, privacyrecht en goed werkgeverschap samenkomen, zoals een werknemer die zich ziekmeldt na een escalerend conflict, of de internationale dimensie van thuiswerken vanuit het buitenland. Deze vragen testen het vermogen om verschillende rechtsgebieden te integreren in een praktisch advies.

In dit rechtsgebied presteert ChatGPT het sterkst, op de voet gevolgd door Gemini van Google. Claude laat ten opzichte van de andere rechtsgebieden veel punten liggen, en Copilot kan maar beter geheel vermeden worden met arbeidsrechtelijke vragen.

## “In hoeverre is het opnemen van een ebook in een LLM toegestaan onder Europees auteursrecht?”

**Gemini (gratis):** “Het opnemen van een e-book in een Large Language Model (LLM) roept complexe juridische vragen op in het kader van het Europese auteursrecht.

Kernvraag: Is dit een toegestaan gebruik?

Het antwoord hierop is niet eenduidig en hangt af van verschillende factoren, waaronder:

- **De omvang van de opgenomen tekst:** Hoe groter het deel van het e-book dat in de LLM wordt opgenomen, hoe groter de kans dat er inbreuk wordt gemaakt op het auteursrecht.
- **De aard van het gebruik:** Wordt het e-book gebruikt om de LLM te trainen of om specifieke vragen te beantwoorden? In het eerste geval is er eerder sprake van een inbreuk op het auteursrecht.
- **De impact op de markt:** Beïnvloedt het gebruik van het e-book in de LLM de markt voor het originele werk? Als de LLM in staat is om grote delen van het boek te reproduceren, kan dit de verkoop van het boek schaden.

Relevante bepalingen in het Europese auteursrecht

- **Reproductierecht:** Het auteursrecht geeft de maker het exclusieve recht om zijn werk te verveelvoudigen. Het opnemen van een e-book in een LLM kan worden gezien als een vorm van verveelvoudiging.
- (...)”

**Claude (betaald):** “Deze vraag raakt aan recente ontwikkelingen in het Europees auteursrecht met betrekking tot text-and-data mining (TDM). Hier is de juridische analyse:

### 1. Algemeen kader

- Art. 3 en 4 DSM-richtlijn (2019/790) regelen text-and-data mining
- Art. 3 betreft TDM voor wetenschappelijk onderzoek
- Art. 4 betreft overige vormen van TDM

### 2. Analyse onder artikel 4 DSM-richtlijn

- Het trainen van een LLM valt onder de definitie van text-and-data mining
- Er is in principe een TDM-exceptie die dit toestaat
- MAAR: rechthebbenden kunnen zich hiertegen verzetten (“opt-out”)
- Een opt-out moet op “passende wijze” worden gecommuniceerd
- Bij online content betekent dit machine-leesbare opt-out (...)”



	ChatGPT		Claude		Copilot		Gemini	
	Gratis	Betaald	Gratis	Betaald	Gratis	Betaald	Gratis	Betaald
Makkelijk	5,6	8,4	5,2	8,1	4,7	4,8	6,1	8,4
Gemiddeld	5,6	7,4	4,7	6,2	2,7	2,6	3,5	7,2
Moeilijk	7,0	7,1	6,0	5,2	2,9	3,1	3,6	5,5
	6,1	7,6	5,3	6,5	3,4	3,5	4,4	7,0

Het arbeidsrecht laat in de gratis versies een stabiel patroon zien voor eenvoudige en gemiddelde vragen, met een lichte stijging voor complexe casuïstiek. ChatGPT toont hier de meest consistente prestaties. Bij de betaalde versies zien we een duidelijke

'upgradesprong': deze bereiken consistent hogere scores. Opvallend is dat Copilot als enige deze sprong niet maakt - het systeem blijft steken op vergelijkbare scores in beide versies.

### “Heb ik recht op loondoorbetaling bij ziekte?”

**Copilot (betaald):** “Ja, in Nederland heb je recht op loondoorbetaling bij ziekte. Volgens de wet moet je werkgever je gedurende maximaal twee jaar minimaal 70% van je loon doorbetalen als je ziek bent<sup>[1][2]</sup>. In het eerste jaar moet dit bedrag ten minste gelijk zijn aan het minimumloon<sup>[1][2]</sup>.

Er kunnen uitzonderingen zijn, afhankelijk van je contract of cao. Bijvoorbeeld, als je minder dan vier dagen per week werkt als huishoudelijke hulp, hoeft je werkgever je maar zes weken door te betalen<sup>[1]</sup>.

Heb je nog specifieke vragen over je situatie?

#### References

[1] [Krijg ik loon doorbetaald bij ziekte? | Het Juridisch Loket](#)

[2] [Hoeveel loon krijg ik doorbetaald als ik ziek ben?](#)

**ChatGPT (gratis):** “Ja, in Nederland heb je recht op loondoorbetaling bij ziekte, maar de precieze voorwaarden hangen af van je arbeidsovereenkomst, de cao (indien van toepassing) en je arbeidsstatus (werknemer, zzp'er, etc.).

ChatGPT en Gemini behaalde bij makkelijke vragen met moeite een voldoende terwijl Claude en Copilot duidelijk meer moeite hadden met zelfs ogenschijnlijk simpele vragen. Het gaat hier om basale vragen rondom loondoorbetaling bij ziekte, recht op pauze en de definitie van schijnzelfstandigheid. De systemen zijn (gemiddeld gezien) dan ook in staat om eenvoudige arbeidsrechtelijke vragen enigszins adequaat te beantwoorden. De scores onderling vertonen echter duidelijke variatie, wat wijst op inconsistentie in de betrouwbaarheid van de informatie die wordt verstrekt. Voor eenvoudige vragen komt ChatGPT als beste uit de bus.

De geteste systemen hadden duidelijk meer moeite met de vragen met een gemiddelde moeilijkheidsgraad. Hier daalde de gemiddelde score (van alle systemen samen) naar een dikke onvoldoende: 4,1. De antwoorden hadden onvoldoende diepgang, er ontbrak relevante informatie en in sommige gevallen werd zelfs foutieve informatie gepresenteerd. Ook bij de gemiddelde vragen bleek ChatGPT de uitblinker.

Vreemd genoeg presteerde de systemen bij moeilijke vragen beter dan bij de vragen van gemiddelde moeilijkheid. Wij vermoeden dat dit ligt aan de manier waarop het systeem complexere vragen analyseert en prioriteit geeft aan volledigheid boven eenvoud. Ook de opbouw van de vraag kan bijgedragen hebben aan de bredere context omdat de vragen complexer waren en daarmee meer details bevatten. Zowel ChatGPT en Claude scoorde hier een voldoende. Desondanks gaven de resultaten ook voldoende blijk van onvolledigheid en op sommige punten onjuistheid.

De betrouwbaarheid tussen de systemen is behoorlijk inconsistent, met scores die variëren van 2,4 tot 7,4. Dit maakt dat AI-systemen weliswaar potentie hebben bij het beantwoorden van arbeidsrechtelijke vragen, maar tegelijkertijd ook aanzienlijke tekortkomingen vertonen in betrouwbaarheid en consistentie. De beste oplossing op het gebied van arbeidsrechtelijke vragen is ChatGPT met duidelijke en uitvoerige antwoorden. Met name Copilot en Gemini gaven op sommige onderdelen zeer korte antwoorden met weinig details en adviseerde contact op te nemen met een jurist. Hoewel dat gezien de resultaten misschien

wel het beste antwoord is, zijn daar geen punten aan toegekend omdat het weinig directe hulp biedt.

Een relevante factor die belangrijk is bij arbeidsrechtelijke vragen, is de veranderende wetgeving. Arbeidsrecht is een dynamisch rechtsgebied waarin nieuwe wet- en regelgeving regelmatig wordt ingevoerd, ook op de onderwerpen waar de systemen naar gevraagd zijn. Denk aan onderwerpen als flexwerken en schijnzelfstandigheid. AI-systemen hebben moeite dergelijke wijzigingen effectief te verwerken en te kwalificeren als juist. Het is dan ook goed mogelijk dat er in de trainingsdata tegenstrijdige informatie zit die gebaseerd is op zowel de oude als de nieuwe wetgeving. Daarnaast hebben sommige systemen te maken met een zogenaamde data cut-off, wat betekent dat hun kennis op een bepaald moment stopt met updaten. Dit betekent dat antwoorden gebaseerd op verouderde wetten of regels een risico vormen in situaties waar recente regelgeving cruciaal is. Hierdoor wordt het noodzakelijk dat gebruikers aanvullende controles uitvoeren op basis van actuele bronnen, dit blijkt dan eens te meer uit de gegeven antwoorden.

## Ondernemingsrecht

Bij ondernemingsrecht kozen we voor een combinatie van formele vereisten en praktische toepassingen die de dagelijkse realiteit van het bedrijfsleven weerspiegelen. De eenvoudige vragen richtten zich op basale vennootschapsrechtelijke aspecten die elke bedrijfsjurist regelmatig tegenkomt: "Wie bepaalt de beloning van een statutair bestuurder?" of "Wanneer moet een jaarrekening uiterlijk worden opgemaakt en gedeponereerd?". Deze vragen testen kennis van fundamentele wettelijke vereisten die cruciaal zijn voor corporate governance.

De middencategorie bevatte vragen over meer complexe ondernemingsrechtelijke constructies, zoals "Wat zijn de voorwaarden voor een dividenduitkering?" of "Wanneer is een vennootschap gebonden aan handelingen van werknemers zonder formele vertegenwoordigingsbevoegdheid?". Hier moeten systemen niet alleen de wet kennen, maar ook de praktische implicaties doorgronden. De moeilijkste

categorie confronteerde de systemen met het opstellen van concrete juridische documenten, zoals aandeelhoudersbesluiten en drag-along/tag-along bepalingen, waarbij technische juridische kennis moet worden vertaald naar praktisch bruikbare documenten die voldoen aan alle formele vereisten.

In dit rechtsgebied is het wederom Claude met de hoogste score, met name door de zeer goede resultaten bij de eenvoudige vragen. ChatGPT ontloopt Claude niet veel.

Het beeld bij het vakgebied ondernemingsrecht wijkt weinig af van het reeds behandelde. We zien dat de systemen iets beter scoren op complexe vraagstukken

dan op basale kennisvragen, wat met name komt door het 'maak' karakter van deze vragen. Zodra er specifieke juridische details nodig zijn, valt de accuratesse snel terug. Dit suggereert dat de systemen beter zijn in het imiteren van juridische taal dan in het waarborgen van juridische correctheid. De 'upgrade-sprong' is ook hier manifest aanwezig.

Claude komt als beste uit de bus, vooral dankzij accurate artikelverwijzingen en goede duiding van wettelijke formaliteiten. ChatGPT volgt op korte afstand, waarbij opvalt dat de antwoorden weliswaar uitgebreid zijn maar vaak irrelevante uitweidingen bevatten.

	ChatGPT		Claude		Copilot		Gemini	
	Gratis	Betaald	Gratis	Betaald	Gratis	Betaald	Gratis	Betaald
Makkelijk	5,7	7,3	7,5	9,5	5,2	6,4	6,1	6,8
Gemiddeld	4,8	8,4	6,0	8,8	5,0	4,8	4,6	5,2
Moeilijk	6,2	9,0	7,1	8,1	5,7	6,7	6,7	6,2
	5,6	8,2	6,9	<b>8,8</b>	5,3	5,9	5,8	6,1



## “Wie bepaalt wat de omvang is van de beloning van een statutair bestuurder van een Nederlandse besloten vennootschap met beperkte aansprakelijkheid?”

**Claude (gratis):** “De algemene vergadering van aandeelhouders (AVA) bepaalt de beloning van statutair bestuurders, tenzij de statuten dit aan een ander orgaan toekennen (meestal de raad van commissarissen als die er is). Dit is vastgelegd in artikel 2:245 van het Burgerlijk Wetboek.

Bij een structuurvennootschap ligt deze bevoegdheid altijd bij de raad van commissarissen volgens artikel 2:263 BW.”

**Gemini (gratis):** “In Nederland wordt het salaris van een statutair bestuurder van een besloten vennootschap met beperkte aansprakelijkheid (BV) over het algemeen bepaald door de algemene vergadering van aandeelhouders. Dit betekent dat de mensen die aandelen in de BV hebben, samen beslissen hoeveel de bestuurder betaald krijgt.

Hoe gaat dit in zijn werk?

1. Voorstel: Vaak doet de raad van commissarissen (als die er is) een voorstel aan de aandeelhouders voor de hoogte van het salaris.
2. Besluit: De aandeelhouders komen vervolgens bijeen in een algemene vergadering en nemen een besluit over het voorstel. Ze kunnen het voorstel goedkeuren, afwijzen of aanpassen.

Wat staat er in de wet?

Dit is vastgelegd in de Nederlandse wet, specifiek in Boek 2 van het Burgerlijk Wetboek (BW). Artikel 2:245 BW stelt dat de bezoldiging van bestuurders wordt vastgesteld door de algemene vergadering, tenzij de statuten van de BV anders bepalen.

Waarom is dit zo geregeld?

**Controle:** Door de aandeelhouders het recht te geven om over het salaris van de bestuurder te beslissen, wordt er een vorm van controle uitgeoefend op de kosten van de BV.

**Belangenafweging:** Aandeelhouders hebben belang bij een goede bedrijfsvoering en een gezond financieel resultaat. Door het salaris te bepalen, kunnen ze ervoor zorgen dat de beloning van de bestuurder in lijn is met de prestaties van de BV.

Samengevat:

Het salaris van een bestuurder van een BV is dus geen vast bedrag, maar wordt bepaald door de aandeelhouders. Dit zorgt ervoor dat er een balans is tussen de belangen van de bestuurder en de belangen van de aandeelhouders.

Als je meer wilt weten over de rechten en plichten van een bestuurder, of over de algemene vergadering van aandeelhouders, dan kun je het beste contact opnemen met een notaris of een advocaat gespecialiseerd in vennootschapsrecht.

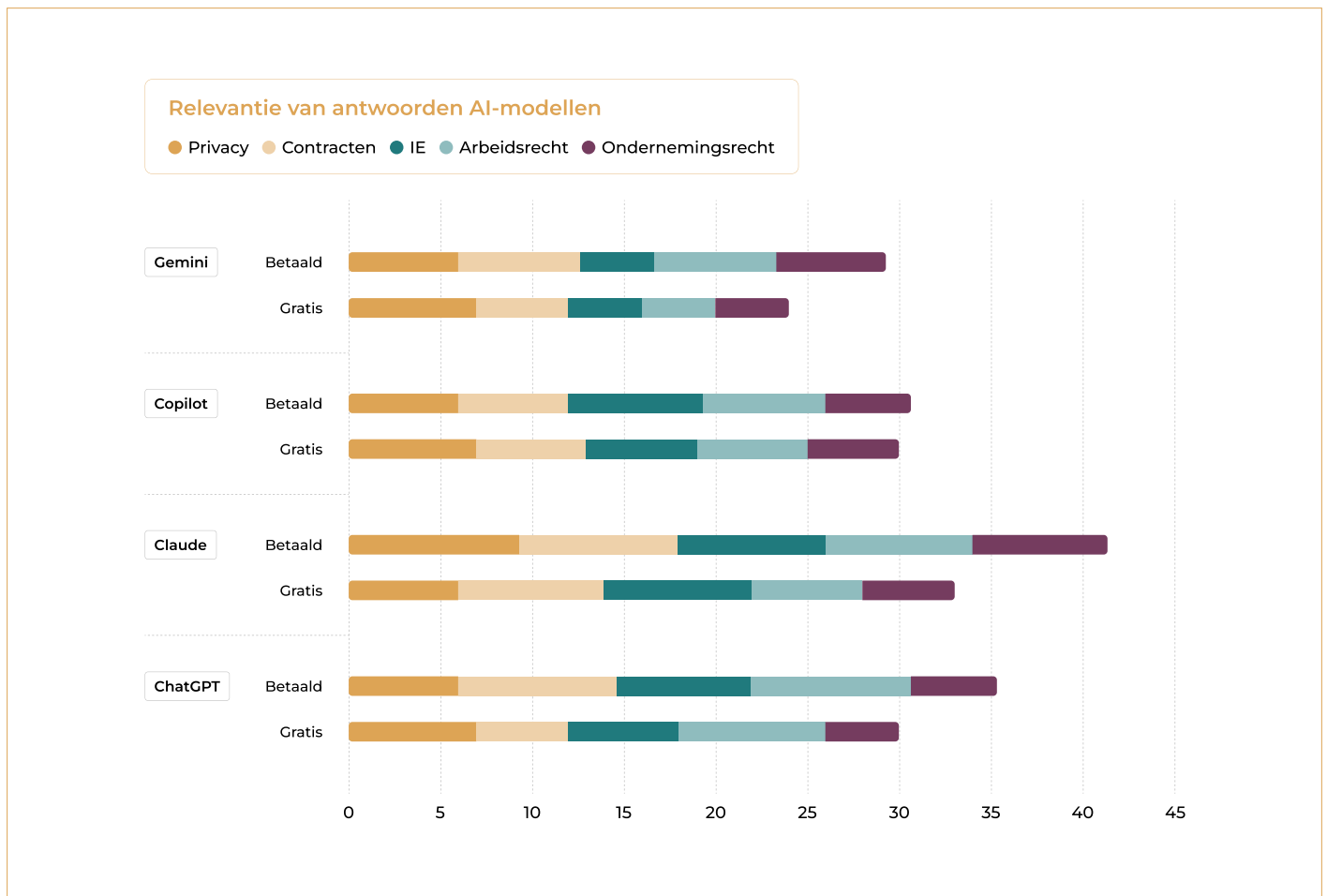
Let op: Deze informatie is bedoeld als algemene kennismaking en is geen juridisch advies. Voor persoonlijk advies is het raadzaam om contact op te nemen met een professional.

De verbeteringen zijn echter vooral gericht op het reproduceren van correcte juridische informatie, maar niet noodzakelijk in het leveren van praktisch bruikbare analyses. In simpele taal: de adviezen zijn vooral langer.

Waar Copilot nog wel juridisch jargon gebruikt (zij het niet altijd correct en zelden met wetsartikelen), valt

bij Gemini op dat het systeem neigt naar 'vertaling' van juridische concepten naar leekentaal. Dit maakt de antwoorden toegankelijker maar vaak ook minder precies. Ook zien we met name bij Copilot weer disclaimers: "Het spijt me, maar ik ben geen financieel adviseur en kan geen juridisch document opstellen zoals een dividendbesluit".

## Relevantie: focus op de vraag



Allereerst kijken we naar relevantie – de mate van focus op de vraag. Een stukje irrelevantie is LLMs eigen. Alle systemen zijn duidelijk getraind om verhalen met een kop en staart te geven. We zien vele opsommingslijstjes en Tussenkopjes Met De Voor LLM's Karakteristieke Hoofdletters. Hierdoor hebben antwoorden sterk de neiging irrelevante informatie te bevatten. Het gaat dan niet alleen om reflecterende openingszinnen

("Een uitstekende vraag! Dit is een complex juridisch vraagstuk") of disclaimers ("raadpleeg vooral ook een jurist") maar ook over bredere kaders waar niet om gevraagd is. Bij de IE-vraag over of een exclusieve licentie schriftelijk moet of niet, geven vrijwel alle modellen een opsomming van de contractuele bepalingen die in zo'n licentie zouden moeten. ChatGPT sluit graag af met een samenvatting die alles nog even op een rijtje zet.

Claude onderscheidt zich nog het meest positief met consistente 8-punts scores voor relevantie bij complexere vragen, met name in het intellectueel eigendomsrecht en contractenrecht. ChatGPT scoort vergelijkbaar bij arbeidsrecht, maar heeft meer moeite focus te houden bij contractuele vraagstukken waar de relevantie sterk terugvalt. Copilot en Gemini blijven vaak steken op 5 à 6 punten, wat betekent dat ze wel de kern raken maar regelmatig afdwalen.

In de betaalde versies tonen de relevantie-scores een bescheiden verbetering. Claude handhaaft consistent hoge scores over alle rechtsgebieden, met name sterk in IE-recht en contractenrecht. ChatGPT Plus scoort

vergelijkbaar hoog bij arbeidsrecht en privacy, maar valt terug bij ondernemingsrecht.

Copilot blijft dicht bij zijn gratis tegenhanger met voornamelijk 6-punts scores, terwijl Gemini Advanced licht verbetert van gemiddeld 4 naar 6 tot 8 punten. De betaalde versies tonen vooral verbetering in hun vermogen om bij het onderwerp te blijven, maar fundamentele uitweidingen blijven voorkomen. Opvallend is dat de relevantie-scores minder dramatisch verbeteren dan de accuratesse-scores, wat suggereert dat zelfs premium versies moeite hebben met het leveren van scherp gefocuste analyses.



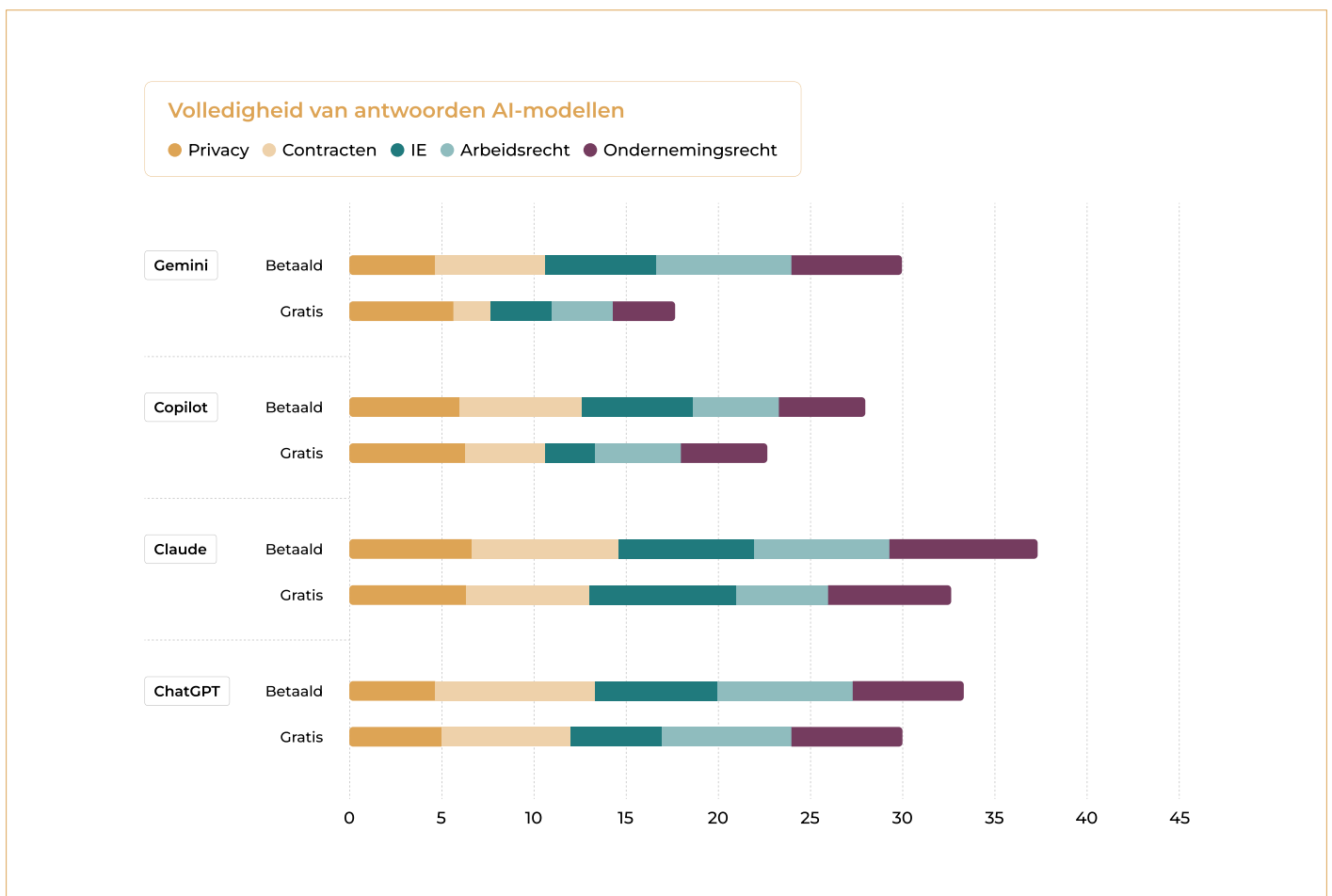
## Volledigheid: diepgang van de antwoorden

De volledigheidsscores tonen een subtiel maar belangrijk verschil tussen de gratis en betaalde versies. De gratis systemen scoren meestal rond de 4 à 6 punten, wat duidt op een basale juridische analyse zonder diepgaande nuances. Dit manifesteert zich vooral in wat we de “net-niet juridisch” noemen: antwoorden die correct klinken maar cruciale juridische subtiliteiten missen. Zo wordt bij het onderwerp contracten gesteld dat “Wanneer een overeenkomst wordt ontbonden, wordt de overeenkomst als **\*\*nooit tot stand gekomen\*\*** beschouwd”. Dit klopt niet: dit is het gevolg van vernietiging en niet van ontbinding van een overeenkomst. Dit soort net-niet verwarring van begrippen is endemisch bij deze taalmodellen.

De betaalde versies bereiken regelmatig 8 punten, zij het zelden de maximale 10 punten voor diepgaande juridische onderbouwing. Maar zelfs dan blijven de systemen worstelen met juridische subtiliteiten.

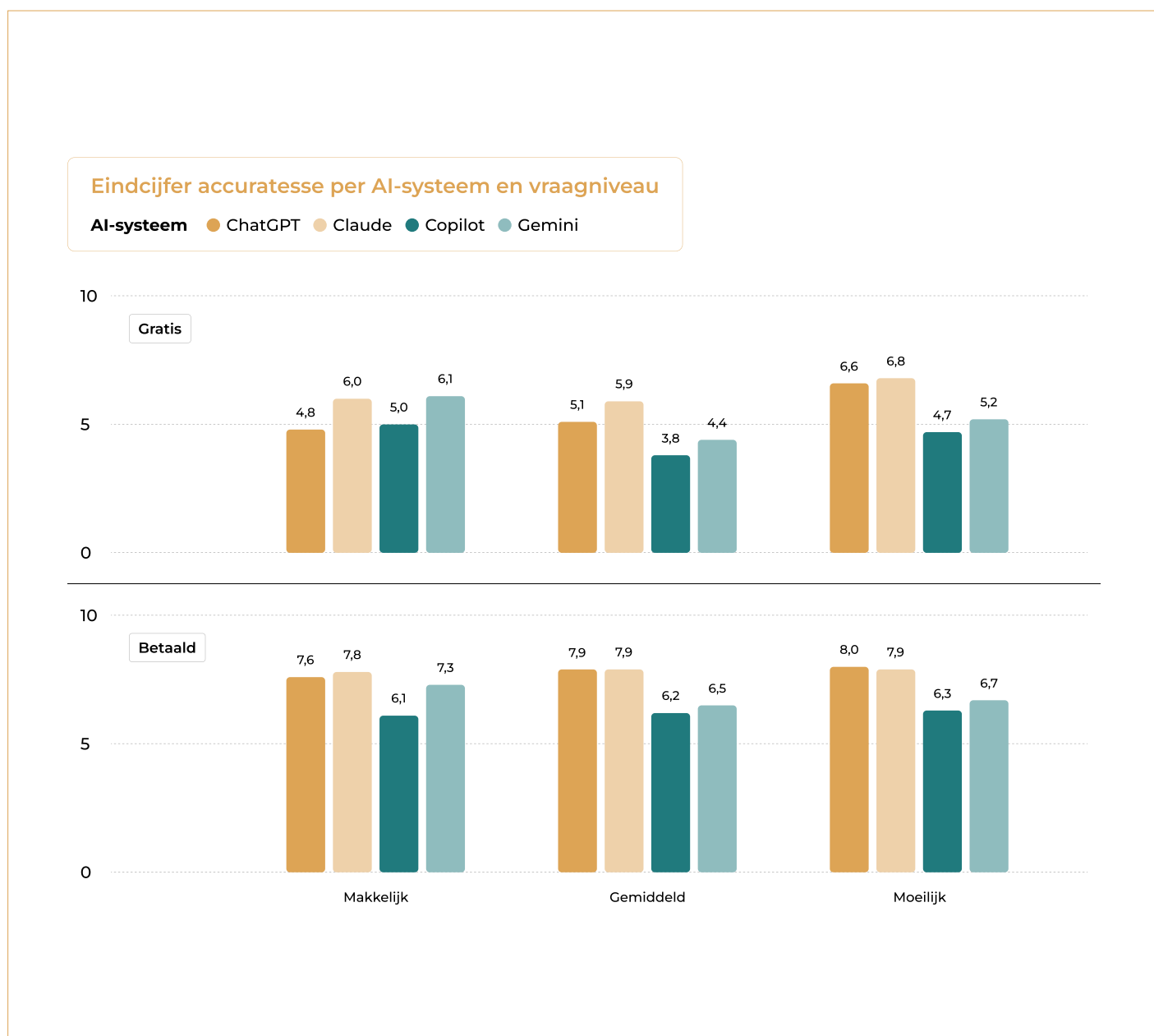
Cruciale onderscheiden (zoals tussen vernietiging en ontbinding bij contracten) worden nog steeds structureel verward, en de neiging om juridisch jargon te ‘vertalen’ blijft bestaan - ook wanneer dit de analyse verzwakt. Dit suggereert dat de premium versies wel meer informatie leveren, maar niet noodzakelijk meer juridische diepgang.

Een terugkerend fenomeen is wat we de “democratisering van juridisch advies” noemen: alle systemen lijken geprogrammeerd om juridische concepten te “vertalen” naar leken taal, ook wanneer de vraag duidelijk van een juridische professional komt. Dit leidt tot onnodig lange antwoorden waarin juridisch jargon wordt vermeden of uitgelegd, wat de effectiviteit voor professionele gebruikers vermindert. Jargon is immers deel van het werk. Ook zien we – met name bij Copilot en Gemini – algemene uitspraken over “de jurisprudentie” als onderbouwing of rechtvaardiging. Voor de professional is dat natuurlijk niet genoeg.



# Conclusie: waar AI uitblinkt en waar niet

We zetten alle bevindingen op een rij in onderstaande tabel. De algemene winnaar is het betaalde model Claude 3.5 Sonnet, met direct daarna ChatGPT 4o (Plus). De resultaten per vakgebied verschillen echter fors.





De onderzochte systemen vertonen een aparte mix van indrukwekkende prestaties en zorgwekkende blindheid voor juridische nuance. De populaire vergelijking met een juridisch stagiair gaat dan ook mank: welke stagiair zou excelleren in complexe IE-analyses maar elementaire contractbeginselen door elkaar halen? Of foutloos AVG-compliance kunnen uitleggen maar struikelen over basale vormvereisten? Dit 'gatenkaaspatroon' van excellentie en tekortkoming is uniek voor kunstmatige intelligentie en vraagt om een eigen begeleidingskader.

Het onderzoek toont een genuanceerd beeld van AI's juridische capaciteiten, met scores die variëren van bijna de perfecte tien tot een dikke nul – en dan alleen omdat het model niet lager gaat dan een nul. Niet de complexiteit van vragen, maar het rechtsgebied blijkt bepalend voor de prestaties. Dit suggereert dat AI-systemen niet zozeer worstelen met juridische complexiteit op zich, maar eerder met specifieke karakteristieken van rechtsgebieden. De meerwaarde van de betaalde versies manifesteert zich vooral in:

1. Consistentie - minder uitschieters naar beneden
2. Jurisdicte-bewustzijn - betere herkenning van nationale verschillen
3. Bronverwijzing - accurater gebruik van wetgeving en jurisprudentie

Onderstaande radar chart visualiseert deze specialisatie per AI-systeem en toont hoe verschillende modellen elkaar kunnen aanvullen in de juridische praktijk.

## Drie specifieke situaties

AI-systemen excelleren in drie specifieke scenario's. Ten eerste bij het snel verkennen van juridische vraagstukken, waar ze in staat zijn relevante wetsartikelen en algemene principes te identificeren. Ten tweede bij rechtsgebieden met een sterk systematische opbouw en uitgekristalliseerde rechtspraak, waar hoge scores worden behaald. Ten derde zien we uitstekende prestaties bij het analyseren van complexe maar logisch gestructureerde juridische vraagstukken die zich lenen voor conceptuele ontleding.

Tegelijkertijd zien we hardnekkige tekortkomingen die juridische professionals niet kunnen negeren. Het meest zorgwekkend is de systematische blindheid voor cruciale juridische nuances, zoals het eerdergenoemde onderscheid tussen ontbinding en vernietiging. De prestaties fluctueren sterk tussen rechtsgebieden - een systeem dat excellent presteert bij privacy kan dramatisch tekortschieten bij contractenrecht. Ook worstelen de systemen met de neiging tot overcontextualisering, wat leidt tot inefficiënte antwoorden vol irrelevante uitweidingen.

## Omgang met bronnen

Bijzonder zorgwekkend is de omgang met bronnen. Waar menselijke juristen altijd beginnen bij wetgeving en jurisprudentie, behandelen AI-systemen deze meer als 'suggesties' dan als bindende bronnen. Ze citeren wetsartikelen die niet bestaan, verwijzen naar jurisprudentie die ze ter plekke verzinnen of - nog verraderlijker - mengen elementen uit verschillende uitspraken tot een nieuwe 'norm' die nergens op gebaseerd is. Deze systemen 'hallucineren' aldus complete juridische fundamenten, waarbij ze de retorische structuur van juridische argumentatie imiteren maar de inhoud volledig verzinnen. Dit maakt ze verraderlijk voor de juridische professional, wiens hele methodiek is gebouwd op het vertrouwen dat bronverwijzingen - hoe men ze ook interpreteert - op zijn minst bestaán. Het dwingt tot een nieuwe werkwijze waarbij elke bronverwijzing moet worden gecontroleerd - niet op juistheid van interpretatie, maar op het simpele feit of de bron überhaupt bestaat.

## De paradox van vertrouwen

Zoals ChatGPT zou zeggen: hier raken we aan een fascinerende paradox die met name zich voordoet in het werk van de moderne bedrijfsjurist. We verwachten van deze professional dat zij zich uitspreekt over het volledige spectrum van het recht, van arbeidsconflicten tot zekerheidsrechten, van privacyschendingen tot IE-kwesties. Maar waar een advocatenkantoor specialisten kan inzetten voor elk rechtsgebied, moet

de bedrijfsjurist als juridische duizendpoot ook de incidentele, eenmalige vraagstukken behandelen. Juist daar lijkt AI een uitkomst: het systeem kan razendsnel de contouren van een onbekend rechtsgebied schetsen, of zelfs een diepgravende vraag van een helder en uitgebreid antwoord voorzien. Maar hoe verifieer je de betrouwbaarheid van AI-analyses over een onderwerp waar je zelf weinig van weet?

De oplossing ligt in het herdefiniëren van het doel: gebruik AI niet om antwoorden te krijgen, maar om de juiste vragen te leren stellen. Waar een specialist direct de kernvragen van een casus identificeert ("Ah, dit is een kwestie van overgang van onderneming!"), kan AI de bedrijfsjurist helpen dit perspectief te ontwikkelen. Het systeem kan relevante contexten verkennen, kernbegrippen en jargontermen identificeren en typische probleem patronen herkennen die de jurist helpen het vraagstuk in te kaderen. Deze kadering maakt vervolgens gerichte verificatie mogelijk: in plaats van de volledige output te moeten controleren, weet de jurist nu waar ze moet zoeken en welke expertise ze moet inschakelen.

Dit transformeert AI van een risicovolle vraagbaak naar een waardevol oriëntatie-instrument. Het helpt de bedrijfsjurist efficiënt specialistische hulp in te schakelen waar nodig, en gericht te zoeken in officiële bronnen waar mogelijk. De kunst is niet om AI's analyse blind te vertrouwen, maar om het te gebruiken als kompas in onbekend juridisch terrein.

## Kun je AI vertrouwen voor juridisch werk?

De centrale vraag die we bij aanvang stelden was simpel maar fundamenteel: kun je deze systemen vertrouwen voor dagelijks juridisch werk? Het antwoord blijkt genuanceerd maar helder: ja, mits je de systemen inzet waar ze sterk in zijn én hun output systematisch verifieert. Concreet betekent dit dat AI uitstekend kan functioneren als eerste verkenning van een rechtsvraag of bij het structureren van juridische informatie - denk aan het opstellen van een checklist voor due diligence of het categoriseren van contractbepalingen. De systemen

zijn ook waardevol bij rechtsgebieden met een heldere systematiek, zoals het analyseren van merkrecht kwesties of het doorlopen van AVG-vereisten.

Maar deze inzet vraagt om strikte kaders: elke bronverwijzing moet worden geverifieerd, niet steekproefsgewijs maar systematisch. AI-uitvoer die stelt dat "volgens vaste jurisprudentie" bepaalde regels gelden, moet alarmbellen doen rinkelen: waar vind ik deze uitspraken? Complexe vraagstukken die meerdere rechtsgebieden raken of zeer actuele wetgeving betreffen, zijn vooralsnog ongeschikt voor AI-ondersteuning. Ook moet elk AI-gebruik worden ingebed in bredere werkprocessen met expliciete controle momenten, bijvoorbeeld door een senior jurist het eindproduct te laten valideren.

Dit betekent in de praktijk dat AI kan ondersteunen bij bijvoorbeeld:

1. Het opstellen van een eerste versie van standaarddocumentatie
2. Het structureren van juridische argumenten in een systematisch overzicht
3. Het vertalen van complexe juridische concepten naar begrijpelijke taal
4. Een eerste verkenning van onbekende rechtsgebieden, mits gebruikt als startpunt voor verdere juridische analyse en niet als eindstation

Maar dat menselijke expertise onmisbaar blijft voor:

1. De verificatie van geciteerde bronnen en jurisprudentie
2. Het waarborgen van cruciale juridische nuances
3. Het beoordelen van timing-gevoelige aspecten
4. De uiteindelijke juridische oordeelsvorming

Deze balans tussen mogelijkheden en waarborgen vraagt om een stevig fundament van risicomanagement en aansprakelijkheidsbeheersing, het onderwerp van ons volgende hoofdstuk.

# Risicomanagement en aansprakelijkheid

---

Laten we het grootste risico voor de praktijk direct benoemen: vrijwel alle generatieve AI-systemen draaien op Amerikaanse servers. Dat betekent dat elke juridische vraag die je stelt, elk document dat je analyseert, door de VS reist. Voor veel juristen voelt dat ongemakkelijk - en terecht.

De betaalde versies, met name de enterprise-varianten, bieden hier meer waarborgen dan de gratis versies. Waar deze algemene voorwaarden hanteren die vooral de aanbieder beschermen, bevatten de zakelijke licenties specifieke bepalingen rond gegevensbeveiliging, vertrouwelijkheid en aansprakelijkheid. Microsoft's Copilot Enterprise gaat hierin het verst met haar EU data boundaries, maar ook de andere aanbieders kennen enterprise-specifieke garanties. Laten we dit per risicogebied bekijken.

## Gegevensbescherming & privacy

De Algemene verordening gegevensbescherming (AVG) vormt sinds 2018 het fundament voor bescherming van persoonsgegevens in Europa. Gegevensstromen naar buiten de EU krijgen daarbij bijzondere aandacht. Het uitgangspunt is simpel: doorgifte van persoonsgegevens naar niet-EU landen vereist passende waarborgen. Dat is meteen de kern van het probleem - elke vraag die je aan een AI-systeem stelt, is in feite een doorgifte van de daarin vervatte persoonsgegevens in de zin van de AVG.

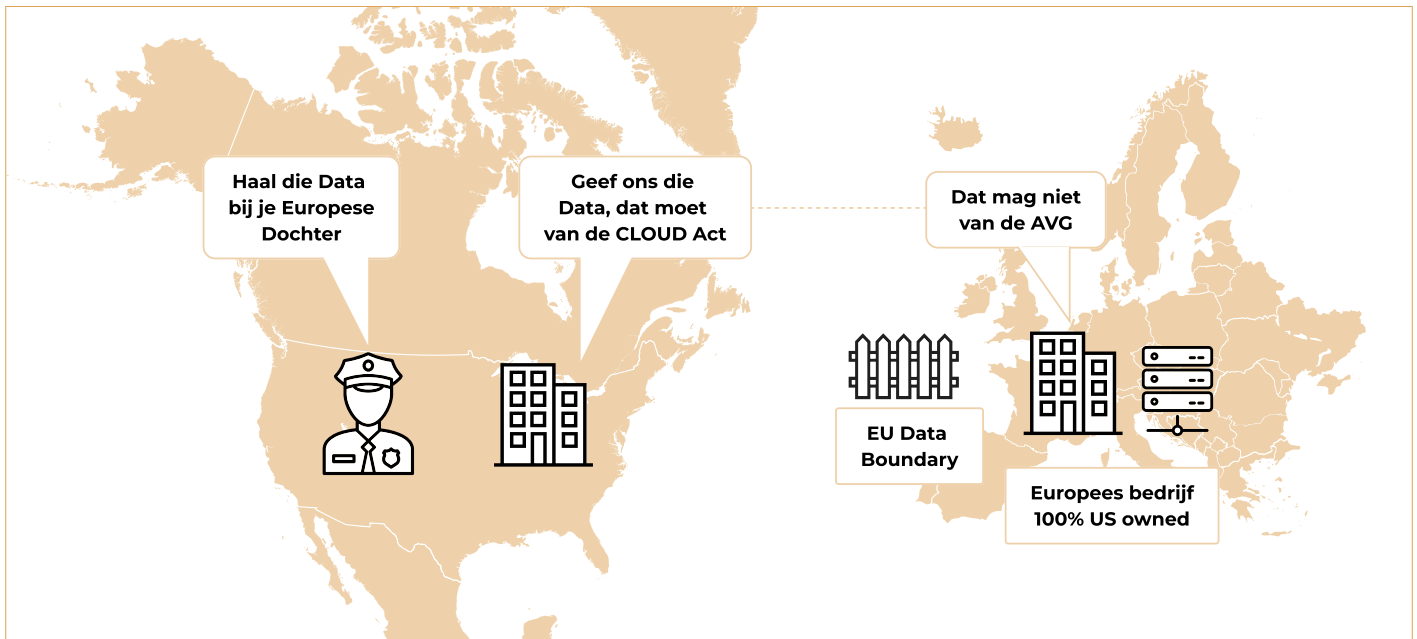
### Papieren waarborgen

Heeft de Verenigde Staten dan geen dergelijke waarborgen? Op papier wel. Sinds maart 2022 bestaat er het Trans-Atlantic Data Privacy Framework (vaak als DPF afgekort) tussen de VS en de EU. Dit adresseert de fundamentele punten van kritiek van het Hof van

Justitie bij de vorige set waarborgen, het Privacy Shield Framework dat in 2020 ongeldig werd verklaard (de Schrems II-uitspraak). Dat was dan weer de opvolger van het Safe Harbor Framework, dat ditzelfde Hof in 2015 al naar de prullenbak verwees (Schrems I). Allemaal met dezelfde argumentatie: er zijn vele papieren garanties, maar we wéten niet wat er in de VS gebeurt.

Hetzelfde beeld zien we bij de modelcontracten, oftewel de standard contractual clauses (SCC). Daar doken organisaties massaal op toen Privacy Shield sneuvelde. Het idee is simpel: je spreekt contractueel af dat de Amerikaanse partij zich aan Europese regels houdt. De Europese Commissie kwam in juni 2021 met een nieuw modelcontract dat beter afgestemd is op de Schrems II uitspraak. Maar ook met deze nieuwe contracten ben je er niet automatisch: het Hof eist nog steeds dat je eerst grondig onderzoekt of die Amerikaanse partij de beloftes uit die SCC's wel kán waarmaken. Je moet kunnen garanderen dat lokale wetgeving (lees: de CLOUD Act) er niet voor zorgt dat ze gedwongen worden die beloftes te breken.

## De CLOUD Act



De olifant in de gegevensporseleinkast blijft de US CLOUD Act. Deze wet verplicht Amerikaanse ict-dienstverleners – ook hun EU-dochters – om in voorkomende gevallen data te delen met Amerikaanse autoriteiten. Dat is in strijd met Europees recht, specifiek artikel 48 AVG, wanneer die data in de EU staat. Maar hoe zou dit in de praktijk uitpakken? Niemand die het weet, en de verhalen over gag orders waarbij de dienstverlener niet eens mag vertellen dát er zo'n bevel is ontvangen, maken de situatie bepaald niet duidelijker.

De grote Amerikaanse techbedrijven bieden uitgebreide garanties, meestal aangeduid met termen als "data boundaries". Hierbij wordt de data fysiek opgeslagen in datacenters in de EU, met organisatorische waarborgen én contractuele afspraken tegen toegang daartoe van buiten de Unie. Met ISO-certificeringen en privacy-audits wordt een en ander geborgd, en transparantie-rapportages laten zien dat men doet wat men belooft. Microsoft is hierbij de absolute koploper op compliance gebied.

Daarnaast zijn er praktische dilemma's. Support en incidentrespons gebeurt vaak vanuit de VS, en om die reden moeten deze vestigingen toegang kunnen krijgen tot data. Belangrijker: AI-modellen worden getraind met

alle input die ze krijgen, ook die van klanten. Dit kan contractueel (of met een vinkje in het controlepaneel) vaak worden uitgezet, maar de technische mogelijkheid is daarmee niet verdwenen, en daarmee ook niet de zorg over ongewenste toegang.

### Compliance in de praktijk

In een vervolgonderzoek gaan wij uitgebreid in op de AVG-compliance van de vier onderzochte LLM aanbieders. We kunnen wel alvast een aantal uitspraken doen:

- OpenAI (ChatGPT) biedt de meeste AVG documentatie maar levert beperkte grip op bewaartermijnen
- Microsoft's integratie in bredere ecosysteem maakt privacy-instellingen complex maar biedt uitgebreide controle
- Google biedt de meest flexibele bewaartermijnen maar heeft versnipperde documentatie
- Anthropic (Claude) is het meest terughoudend met gegevensverwerking maar biedt beperktere controle mogelijkheden

Onze analyse laat zien dat Microsoft met haar EU data boundaries en uitgebreide compliance-maatregelen de meest robuuste AVG-waarborgen biedt. Deze aanpak komt echter met significante kosten en vereist een complexe IT-infrastructuur die vooral voor het mkb een uitdaging kan vormen. De meer toegankelijke alternatieven - ChatGPT, Gemini en Claude - bieden weliswaar minder vergaande waarborgen, maar opereren wel binnen het bestaande juridische kader van SCC's en het EU-US Data Privacy Framework, dus zonder expliciete garanties dat data in de EU blijft. Hoewel er juridische discussie bestaat over de houdbaarheid van deze kaders, en zorgen over de CLOUD Act blijven, is het gebruik van deze systemen in de praktijk goed te verdedigen. De realiteit is dat vrijwel alle organisaties die internationale diensten afnemen zich in dezelfde positie bevinden.

Het zou onrealistisch zijn om te wachten op absolute juridische zekerheid voordat deze technologie wordt ingezet. Mocht het Europese Hof van Justitie in de toekomst oordelen dat de huidige waarborgen onvoldoende zijn, dan treft dit het hele Europese bedrijfsleven. Tot die tijd adviseren wij:

### 1. Kies een passend systeem

- Voor grote organisaties met gevoelige data: Overweeg Microsoft's enterprise-oplossing
- Voor MKB en algemeen gebruik: ChatGPT, Gemini of Claude zijn verdedigbare keuzes
- Gebruik alleen de betaalde versies, aangezien deze met de meeste waarborgen en mogelijkheden komen

### 2. Implementeer aanvullende waarborgen:

- Beperk de input van persoons- en bedrijfsgevoelige gegevens
- Documenteer uw afwegingen rond gegevensminimalisatie

- Zet beschikbare privacyfuncties zoals geheugenbeperking actief in
- Ontwikkel een exitplan dat op zeer korte termijn ingezet kan worden wanneer SCC's en/of DPF ongeldig verklaard worden

### 3. Monitor ontwikkelingen:

- Houd jurisprudentie en AVG/AI Act nieuws in de gaten
- Wees voorbereid op aanpassingen in uw gebruik
- Overweeg lidmaatschap van relevante brancheorganisaties voor gezamenlijke belangenbehartiging

## Geheimhouding & vertrouwelijkheid

Ook als we het AVG-compliancevraagstuk hebben opgelost, blijft er een fundamentele spanning tussen AI-gebruik en vertrouwelijkheid. Want zelfs wanneer alle persoonsgegevens zijn verwijderd, kan het delen van informatie met AI-systemen nog steeds een schending van vertrouwelijkheid betekenen. Denk aan bedrijfsgeheimen, beroepsgeheimen of informatie die onder een geheimhoudingsovereenkomst is verkregen. De aanbieder van zo'n AI-systeem is immers een 'derde' in de zin van regels over geheimhouding en vertrouwelijkheid.

### Vertrouwelijkheid van de invoer

Vele verhalen doen de ronde over AI-systemen die ingevoerde vertrouwelijke informatie verwerken in trainingsdata, waarna de AI op een later tijdstip deze informatie verwerkt in uitvoer bij andere partijen. De meeste hiervan zijn te herleiden tot een incident uit begin 2023 bij technologiebedrijf Samsung. Medewerkers maakten gretig gebruik van ChatGPT om gespreksaantekeningen en softwarebroncode te verbeteren, totdat het bedrijf ingreep en toegang aan banden legde.

Dit verhaal muteerde al snel naar dat anderen daadwerkelijk toegang tot deze informatie hadden

gekregen, maar daarvoor zijn geen aanwijzingen. Ook waar andere bedrijven ontdekten dat medewerkers vertrouwelijke informatie in AI-diensten plaatsten, is nooit daadwerkelijk naar boven gekomen dat deze met derden werd gedeeld.

## Geheimhouding in de voorwaarden

De vier diensten adresseren geheimhouding in hun gebruiksvoorwaarden (terms of use, terms of service, EULA) maar op zeer verschillende manieren. OpenAI doet in de gratis versie geen uitspraken over geheimhouding in haar gebruiksvoorwaarden, maar stelt wel dat zij invoer én uitvoer mag hergebruiken om haar diensten te verbeteren en om aan verzoeken van politie en justitie te voldoen. Google is nog explicieter: zij bepaalt letterlijk dat het verboden is om vertrouwelijke informatie in de dienst in te voeren, wat impliceert dat zij informatie als niet-vertrouwelijk zal beschouwen.

In de voorwaarden bij de betaalde versies zien we meer standaard geheimhoudingsbedingen. Zowel de voorwaarden van OpenAI als van Microsoft Copilot bevatten expliciete geheimhoudingsbepalingen die overeenkomen met de gemiddelde NDA. Expliciet wordt vermeld dat invoer van de klant als 'Confidential Information' wordt aangemerkt. Vergelijkbare bepalingen staan in de Anthropic voorwaarden voor Claude. Alleen de voorwaarden van Google blijven net zo anti-geheimhouding als in de gratis versie.

## Vertrouwelijkheid en shadow IT

*80% van de werknemers gebruikt niet-goedgekeurde apps die niemand heeft beoordeeld en die mogelijk niet compatibel zijn met uw beveiligings- en nalevingsbeleid.*

(Microsoft, november 2024)

Het probleem wordt verergerd door het gegeven dat AI diensten vaak als 'shadow IT' worden ingezet: medewerkers die op eigen initiatief gratis AI-tools

gebruiken. Een advocaat die 'even snel' een contract laat controleren door Google Gemini, een juridisch medewerker die een zaakanalyse door Claude laat maken - zonder dat de organisatie hier grip op heeft. Deze gratis tools bieden geen enkele garantie over wat er met de ingevoerde informatie gebeurt. Sterker nog: hun gebruiksvoorwaarden stellen vaak expliciet dat ze alle input mogen hergebruiken voor modeltraining.

Voor advocaten (al dan niet in loondienst) is deze situatie extra precair vanwege hun beroepsgeheim. Dit is niet zomaar een contractuele verplichting, maar een wettelijk beschermde kernwaarde van de advocatuur. Een advocaat die vertrouwelijke cliëntinformatie deelt met een AI-systeem zonder adequate waarborgen, riskeert niet alleen claims maar ook tuchtrechtelijke consequenties. De Nederlandse Orde van Advocaten werkt aan specifieke gedragsregels voor AI-gebruik, al is onduidelijk wanneer deze zullen verschijnen.

## Auteursrechtelijke zorgen

Een veel gehoorde zorg bij het gebruik van LLM's betreft auteursrecht: genereren deze systemen niet simpelweg variaties op bestaande teksten, en levert dat dan geen inbreuk op auteursrechten op? De technische realiteit is genuanceerder. LLM's werken met statistische patronen en waarschijnlijkheden, niet met het kopiëren en herrangschikken van bestaande teksten. Ze voorspellen welk woord waarschijnlijk volgt op basis van patronen in hun trainingsdata.

### Auteursrechtinbreuk

Gezien de onvoorstelbare grootte van de gebruikte datasets is het niet waarschijnlijk dat een specifieke tekstuitvoer uit een van de vier onderzochte modellen overeen zal komen met een bestaande brontekst. In de literatuur of de pers zijn in ieder geval geen gevallen genoemd. De enige uitzondering is de New York Times, die in haar rechtszaak tegen OpenAI heeft laten zien dat concrete krantenberichten gereproduceerd worden door het systeem wanneer daar specifiek om wordt gevraagd. Dit lijkt een nogal specifiek geval.

De situatie is dus wezenlijk anders dan bij AI-systemen voor beeldgeneratie, waar regelmatig wordt

gerapporteerd dat gegenereerde afbeeldingen duidelijk elementen bevatten van specifieke bronafbeeldingen of bekende personages. Ook bij het genereren van software-broncode zijn er diverse gedocumenteerde gevallen waarbij AI-systemen hele codeblokken letterlijk reproduceerden.

In de Verenigde Staten lopen verschillende rechtszaken waarin wordt gesteld dat LLM's inbreuk maken op auteursrecht. Deze zaken betreffen echter vooral het gebruik van auteursrechtelijk beschermde werken als trainingsmateriaal, niet de daadwerkelijke output van deze systemen. In de EU is dit evenzo onderwerp van discussie, waarbij vooral het gebruik van opt-out verklaringen centraal staat (art. 15o Auteurswet). Het lijkt echter nogal gezocht dat een auteursrechtinbreuk tijdens het trainen van een model relevant is voor gebruikers van dat model om daar nieuwe tekst mee te genereren.

## Bescherming tegen inbreukbeschuldigingen

In contractuele zakelijke relaties is een vrijwaring of indemnification het standaard instrument om een partij te beschermen tegen aanspraken van derden, zoals bij auteursrechtinbreuk. We zien dit terug in de gebruiksvoorwaarden van de AI-aanbieders, zij het eigenlijk alleen in de zakelijke variant. Dit is meestal de versie van de dienst waarmee ook geautomatiseerde (API) toegang mogelijk is en waarbij gesprekken en kennis onderling gedeeld kan worden binnen een team.

Geen van de vier diensten biedt een vrijwaring bij haar gratis of betaalde consumentenversie. OpenAI draait de zaak zelfs om: de gebruiker moet de organisatie vrijwaren van alle schadeclaims die samenhangen met diens gebruik van de dienst, dus ook daarmee samenhangende auteursrechtinbreuken. Anthropic sluit in deze voorwaarden haar aansprakelijkheid volledig uit, inclusief voor inbreuken op rechten van derden.

In de zakelijke voorwaarden is meer ruimte. Microsoft en Anthropic vermelden in de zakelijke gebruiksvoorwaarden een expliciete vrijwaring tegen

de gevolgen van een auteursrechtclaim bij de klant, zowel voor uitvoer uit Copilot als voor het gebruik van het uitsluitend bij de zakelijke variant van de dienst. Google heeft naast haar standaard vrijwaring (bij zakelijke afnemers) een extra bepaling die dit nadrukkelijk bevestigt voor deze twee gevallen. Ook OpenAI bevestigt haar Enterprise klanten te vrijwaren van auteursrechtclaims. Deze vrijwaringen zijn wel altijd beperkt tot de uitvoer 'as is' en onder inachtneming van alle gebruiksvoorwaarden en restricties, en gelden niet wanneer de klant wist of had moeten weten dat de prompt tot auteursrechtinbreuk leidt.

## Auteursrecht op de uitvoer

Een ander onderwerp is auteursrecht op de uitvoer. In de literatuur zijn de meningen sterk verdeeld over de vraag of je kunt spreken van auteursrechtelijk relevante creativiteit, en zo ja aan wie dat auteursrecht dan toekomt. Een LLM heeft immers geen rechtspersoonlijkheid. De menselijke bemoeienis betreft enerzijds de keuze van de prompt (de gebruiker) en anderzijds het samenstellen van het model (de aanbieder). Over deze interactie en wat dit moet betekenen voor het auteursrecht op de uitkomst zal niet snel consensus ontstaan.

Gelukkig bepalen OpenAI en Anthropic in haar gebruiksvoorwaarden dat zij alle eventuele auteursrechten in uitvoer aan de gebruiker overdraagt. Google en Microsoft doen hetzelfde, zij het iets voorzichter geformuleerd dat zij "het eigendom van die content niet [claimen]." Microsoft legt uit dat zij dit mede doet omdat uitvoer bij twee klanten op elkaar kan lijken en dit tot auteursrechtelijke problemen zou kunnen leiden.

## De AI Act: een kwestie van geletterdheid

Per 1 augustus 2024 is de AI Act (gefaseerd) van kracht. Sinds kort (2 februari) geldt er bijvoorbeeld de organisatiebrede eis van AI geletterdheid, en moeten bepaalde praktijken worden gestaakt. De bulk van de eisen – voor zogeheten hoogrisico-AI en transparantierisico's – gaan echter pas in 2026 gelden. Voor de onderzochte AI systemen zijn deze niet aan de

orde: advies geven aan een (juridisch) professional is geen hoogrisico-toepassing.

Ons onderzoek onderstreept waarom AI-geletterdheid zo belangrijk is: LLM's vertonen een uniek patroon van excellentie en falen dat fundamenteel anders is dan menselijke expertise. Het vergt een nieuwe vorm van professionaliteit om deze tools effectief in te zetten. Dit manifesteert zich in drie dimensies. Ten eerste moet de jurist patronen leren herkennen in waar AI betrouwbaar is. Ons onderzoek toont dat systemen excelleren in systematische rechtsgebieden en bij conceptuele analyses, maar dramatisch kunnen falen zodra bronverwijzingen of actuele jurisprudentie een rol spelen. Dit vraagt om een nieuw soort beoordelingsvermogen: niet het traditionele toetsen van juridische redenering, maar het herkennen van contexten waarin AI waarschijnlijk waardevol - of juist risicovol - zal zijn.

Ten tweede is er de kunst van systematische verificatie. Waar juristen gewend zijn bronnen te controleren op interpretatie, dwingt AI tot een fundamenteel andere aanpak: het verifiëren van het simpele bestaan van bronnen. Dit vraagt om nieuwe werkprocessen waarin verificatie is ingebouwd, niet als incidentele check maar als structureel onderdeel van AI-gebruik.

De derde dimensie betreft strategische inzet. AI blijkt vooral waardevol als oriëntatie-instrument dat helpt vraagstukken te kaderen en relevante contexten te identificeren. Dit maakt het bij uitstek geschikt voor de bedrijfsjurist die zich moet verhouden tot het volledige spectrum van het recht. Maar deze inzet vraagt om zorgvuldige afweging: wanneer volstaat AI-ondersteuning, wanneer is specialistische expertise nodig?

Voor juridische afdelingen betekent dit een nieuwe opleidingsuitdaging. Het doel moet worden het creëren van een leercultuur waarin experimenteren met AI wordt gestimuleerd en ervaringen systematisch worden gedeeld. Door regelmatig AI-antwoorden gezamenlijk te analyseren, beste praktijken te documenteren en concrete successen en mislukkingen te bespreken, ontwikkelt een team gevoel voor de mogelijkheden en grenzen van deze technologie.

Cruciaal is dat deze training niet verzandt in technische details maar focust op professioneel beoordelingsvermogen. Het gaat niet om perfect begrip van de technologie, maar om het vermogen verstandige keuzes te maken over de inzet ervan. Dit vraagt om continue investering in het hele team, niet alleen in technische experts of early adopters. Alleen zo ontwikkelt een organisatie het collectieve oordeelsvermogen dat nodig is om AI effectief en verantwoord in te zetten in juridisch werk.

## Beroepsaansprakelijkheid in het AI-tijdperk

“Zoals het Hooggerechtshof oordeelde in *Martinez v. Delta Airlines* (2023)...” - alleen bestond die zaak alleen in de fantasie van ChatGPT. De rechter in deze New Yorkse zaak was not amused, de advocaat moest diep door het stof, en de juridische wereld kreeg een harde les over blind vertrouwen in AI. Grappig? Misschien. Maar ook een waarschuwing. Want wat als die gefantaseerde jurisprudentie niet zo makkelijk te controleren was geweest? Wat als een jurist adviseert op basis van een AI-juridische analyse die subtiel maar cruciaal de plank mislaat?

### De AI-paradox bij juridisch werk

Deze situatie staat in schril contrast met hoe we omgaan met traditionele juridische tools. Geen enkele jurist zou accepteren dat “Kluwer Navigator gaf geen resultaten” het einde van het onderzoek betekent. Of dat “de zoekmachine van het Tijdschrift voor Aansprakelijkheidsrecht vond niets” een afdoende verificatie is van een rechtsvraag. We verwachten van juristen dat ze meerdere bronnen raadplegen, kruisverwijzingen controleren en hun analyse baseren op een solide analyse van wetgeving, jurisprudentie en gezaghebbende opinies van rechtsgeleerden.

Waarom accepteren we dan wel klakkeloos wat een AI-systeem ons vertelt? Misschien omdat die output zo overtuigend klinkt, zo menselijk in zijn formulering, zo zelfverzekerd in zijn analyse. Maar juist die overtuigende presentatie maakt AI-hallucinaties des te gevaarlijker. Anders dan een ‘geen resultaten’ melding van een



zoekmachine, presenteert AI zijn verzinsels als vaststaande juridische waarheden.

Deze dynamiek is bijzonder risicovol voor de bedrijfsjurist. Als juridische duizendpoot navigeert deze professional dagelijks tussen arbeidsrecht, contracten, privacy, compliance en intellectueel eigendom. Van alles moet je iets weten, maar specialistische diepgang in elk rechtsgebied is praktisch onmogelijk. Dan is een AI-systeem dat complexe juridische materie reduceert tot heldere, toegankelijke analyses verleidelijk - zelfs bevrijdend. Maar waar de IE-specialist direct ziet dat een analyse over auteursrecht te simplistisch is, of de privacyexpert ontbrekende nuance herkent in een AVG-redenering, mist de generalist mogelijk de waarschuwingssignalen. Het AI-systeem presenteert zich als de perfecte specialist-op-afroep: altijd beschikbaar, nooit aarzelend, en met een antwoord op elke vraag. Maar anders dan een menselijke specialist, heeft het geen besef van zijn eigen beperkingen. Het zal nooit zeggen "deze kwestie is te complex voor een quick scan" of "hier heb je echt specialistisch advies nodig."

### Een juridisch vacuüm

Op dit moment zijn er nog geen Nederlandse uitspraken over beroepsaansprakelijkheid of tuchtrecht specifiek gerelateerd aan AI-gebruik door juristen – of andere professionals. Dit is geen geruststelling, maar juist een extra reden voor waakzaamheid. Het betekent dat we nog geen uitgekristalliseerde normen hebben voor wat 'zorgvuldig AI-gebruik' precies inhoudt. Moeten we AI behandelen als een research-tool, een juridisch medewerker, of een extern advies? Elk model brengt zijn eigen zorgplichten en aansprakelijkheidsrisico's met zich mee.

De zaak uit New York geeft een voorproefje van hoe Nederlandse rechters en tuchtcolleges hier mogelijk naar zullen kijken. De Amerikaanse rechter oordeelde dat het blind vertrouwen op AI-gegenereerde jurisprudentie onzorgvuldig was. De advocaat had moeten verifiëren of de geciteerde uitspraak daadwerkelijk bestond. Deze redenering past naadloos in het Nederlandse systeem van beroepsaansprakelijkheid, waar van juridische professionals een hoge mate van zorgvuldigheid wordt verwacht. Maar hoe vullen we deze norm concreet in?

### De zorgplicht in het AI-tijdperk

Ons onderzoek laat zien dat het gebruik van AI-systemen de zorgplicht van juridische professionals niet vermindert, maar juist verzaamt. De overtuigende toon van AI-uitvoer in combinatie met de reële kans op feitelijke of juridische fouten maakt dat de professional nooit klakkeloos kan drijven op dergelijke uitvoer.

Voor de bedrijfsjurist betekent dit een nieuwe kerncompetentie: het vermogen om te herkennen wanneer een AI-antwoord té glad, té simpel, té overtuigend klinkt. Want juist als de technologie het meest behulpzaam lijkt - bij het vertalen van complexe specialistische materie naar praktische handvatten - is het risico op oversimplificatie het grootst.

Deze nieuwe kerncompetentie is niet langer vrijblijvend. De AI Act vereist vanaf februari 2025 dat professionals die AI-systemen inzetten voor hun werk, aantoonbaar 'AI-geletterd' zijn. Voor juristen betekent dit meer dan alleen weten hoe je een prompt schrijft. Het gaat om het fundamentele begrip van AI-systemen: hun mogelijkheden én beperkingen, het vermogen om resultaten kritisch te evalueren, en de vaardigheid om verantwoord met deze technologie om te gaan. Waar voorheen een disclaimer over AI-gebruik volstond, wordt nu een actieve beheersing van AI-risico's verwacht.

Eén ding is zeker: "De AI zei het" is geen verweer bij een beroepsfout. Sterker nog, ongedocumenteerd of ongecontroleerd AI-gebruik zou juist als extra onzorgvuldig kunnen worden gezien. Het is als een advocaat die blind vertrouwt op het werk van een stagiair, zonder ook maar één bron te controleren.

# Vooruitblik & Aanbevelingen

---

“Even snel door ChatGPT halen” - het is inmiddels een gevleugelde uitspraak op juridische afdelingen. Waar we twee jaar geleden nog sceptisch waren over AI in juridisch werk, gebruiken we het nu dagelijks. Uit ons onderzoek blijkt dat dit geen goed nieuws én geen slecht nieuws is - het is gewoon realiteit. Net zoals we ooit moesten wennen aan digitale jurisprudentiedatabases (“Maar hoe weet je zeker dat alle uitspraken erin staan?”) en automatische documentgeneratie (“Een computer kan toch geen contract opstellen?”), zoeken we nu onze weg met AI. Maar er is een belangrijk verschil: waar eerdere innovaties vooral sneller en efficiënter maakten wat we al deden, dwingt AI ons om fundamenteel anders na te denken over hoe we juridisch werk doen. Want een AI-systeem dat met groot zelfvertrouwen een niet-bestaand arrest citeert, is een heel ander risico dan een zoekmachine die ‘geen resultaten’ meldt.

## De balans tussen kwaliteit en compliance

Ons onderzoek toont een fascinerend spanningsveld. De kwaliteit van AI-juridische analyses is vaak verrassend goed - met name bij de betaalde versies - maar tegelijk fundamenteel onbetrouwbaar. Het is als een briljante stagiair die in één adem een waterdichte juridische analyse kan leveren én compleet verzonnen jurisprudentie kan citeren. Om vervolgens bij een basisvraag niet eens uit de startblokken te komen. Hoe moet je daar nu mee werken als professional?

Deze kwaliteitsparadox wordt gecompliceerd door compliance-uitdagingen. Waar we bij kwaliteit nog kunnen compenseren met menselijke controle, zijn compliance-risico's vaak binair: je deelt vertrouwelijke informatie met een Amerikaans bedrijf, of je gebruikt

in het geheel geen AI. De papieren waarborgen (SCCs, EU-US Data Privacy Framework) bieden een juridisch kader, maar geen garanties.

Een algemene conclusie is in ieder geval dat wie met grote taalmodellen juridisch werk wil versterken, daar een betaald abonnement mee neemt. Gezien de bedroevende resultaten én de vergaande toegang tot gegevens uit de voorwaarden durven wij het wel haast een beroepsfout te noemen om met de gratis versies te werken op werkelijke dossiers of cliëntadviezen.

De prestaties per rechtsgebied verschillen sterk. Wij vermoeden sterk dat hier een Europeesrechtelijke verklaring achter zit: over geharmoniseerde vakgebieden zoals privacy/gegevensbescherming en intellectueel eigendom is meer brondata te vinden waar een taalmodel wat mee kan schrijven dan over gespecialiseerd recht zoals arbeids- of ondernemingsrecht. En ook bij de sterkst presterende modellen in een rechtsgebied blijft de neiging tot verzinnen van rechtsbronnen aanwezig.

De hamvraag voor organisaties is niet óf ze AI gaan gebruiken - dat gebeurt al - maar hoe ze dit verantwoord kunnen inbedden. De kunst is om technologische mogelijkheden, juridische vereisten en praktische realiteit met elkaar te verzoenen. Soms betekent dit kiezen voor duurdere enterprise-oplossingen met betere waarborgen, soms juist voor strikte beperkingen op AI-gebruik bij gevoelige zaken.

De onderzochte modellen zijn allen Amerikaans en algemeen. Zijn er dan geen Nederlandse of Europese modellen die specifiek voor de juridische praktijk geschikt zijn? Helaas. Tot nu toe zien wij vooral marketingbeloftes, of tools die zich heel nadrukkelijk aan de bovenkant van de markt positioneren. Logisch:

het ontwikkelen van fundamenteel nieuwe, echt Europese AI-systemen kost miljarden en vereist enorme datacollecties en menskracht. De juridische markt is daarbij moeizaam te veranderen, zoals het jarenlange geworstel met legal tech laat zien. Als zelfs een basaal documentmanagementsysteem al jaren discussie oplevert, wordt het niet makkelijker met een AI-tool die complexe implementatie vereist.

De onderzochte modellen hebben juist daarom een belangrijk praktisch voordeel: ze zijn direct bruikbaar, zonder gedoe met IT-afdelingen of dure consultants. Je stelt een vraag en krijgt een antwoord. Dat antwoord is misschien niet perfect - zoals ons onderzoek laat zien - maar het is vaak goed genoeg en vormt een prima startpunt voor juridische analyse. En dat is precies waar de meeste juristen behoefte aan hebben: niet aan nog een complex systeem dat eerst "geïmplementeerd" moet worden, maar aan praktische ondersteuning bij het dagelijks werk.

## Concrete aanbevelingen voor juridische afdelingen

Wat betekent dit voor de juridische afdeling die nú een strategie moet bepalen? Het antwoord ligt niet in het zoeken naar de perfecte tool - die bestaat niet. Het gaat om het ontwikkelen van een werkwijze die past bij de realiteit van vandaag. Een aanpak die technologie omarmt waar nuttig, maar de juridische kern van ons werk niet uit het oog verliest.

Die balans zal voor elke organisatie anders zijn. Een advocatenkantoor dat vooral procedeert heeft andere behoeften dan een bedrijfsjurist die dagelijks tientallen contracten moet beoordelen. Een privacy officer die toeziet op internationale datastromen heeft andere zorgen dan een jurist die zich richt op Nederlands arbeidsrecht. Het gaat er niet om wat theoretisch het beste is, maar wat praktisch het beste werkt.

Deze inzet vraagt om strikte kaders: elke bronverwijzing moet worden geverifieerd, niet steekproefsgewijs maar systematisch. De vraagstukken moeten passen bij waar AI sterk in is, zoals het opstellen van een eerste

versie van standaarddocumenten of het structureren van argumenten. Ook als eerste verkenning van onbekende rechtsgebieden is AI-inzet zeer geschikt. Complexe vraagstukken die meerdere rechtsgebieden raken of zeer actuele wetgeving betreffen, zijn echter minder geschikt. Daarnaast moet AI-gebruik worden ingebed in bredere werkprocessen met expliciete controlemomenten, bijvoorbeeld door een senior jurist het eindproduct te laten valideren.

Toch zien we in ons onderzoek een aantal universele principes. Ten eerste: de kwaliteit van AI-antwoorden staat of valt met de kwaliteit van menselijke controle. Een goed opgezet verificatieproces kan zelfs met gratis tools tot bruikbare resultaten leiden. Omgekeerd kan een dure enterprise-oplossing zonder goede governance tot juridische missers leiden. Ten tweede: de grootste risico's zitten niet in wat AI verkeerd doet, maar in wat juristen ermee doen. Het is als een auto met rijassistentie: niet het systeem bepaalt de veiligheid, maar hoe de bestuurder ermee omgaat.

Verbieden, zei je? Laat dat onzalige idee varen. Het drijft gebruik ondergronds en creëert een gevaarlijke vorm van 'shadow IT' waarbij medewerkers op eigen houtje - en zonder enige waarborg - toch deze tools gebruiken. Want laten we eerlijk zijn: als een bedrijfsjurist om half vijf een eerste inschatting moet maken van een 80 pagina's tellend contract, is de verleiding van "even door ChatGPT halen" moeilijk te weerstaan. En als een juridisch medewerker moet uitzoeken of een specifieke situatie onder de meldplicht datalekken valt, is Claude een stuk sneller dan het handmatig doorspitten van de AVG-guidelines. Het verbieden van deze tools is niet alleen onrealistisch, het is ook contraproductief. Het verhindert de ontwikkeling van verantwoorde werkwijzen en drijft een wig tussen management en werkvloer. Veel beter is het om gebruik te reguleren, te monitoren en waar nodig bij te sturen. Liever een imperfect maar transparant proces dan een perfect beleid dat niemand volgt.

## Concrete aanbevelingen voor de individuele jurist

Allereerst: erken dat je waarschijnlijk al AI gebruikt, of dat binnenkort gaat doen. Niet omdat het moet of omdat het hip is, maar omdat het gewoon praktisch is. Een eerste scan van een contract, een snelle check van jurisprudentie, een voorzetje voor een advies - AI-tools zijn als een TomTom voor juridisch werk. Ze wijzen de weg, maar jij blijft achter het stuur zitten. En net zoals je bij TomTom altijd even checkt of die route wel logisch is, moet je bij AI-tools altijd kritisch blijven kijken.

Begin klein maar doelgericht. Kies één specifiek type taak waar je AI voor wilt gebruiken, bijvoorbeeld het maken van een eerste opzet voor standaarddocumenten. Experimenteer daarmee, leer de valkuilen kennen, ontwikkel een gevoel voor wanneer je de output kunt vertrouwen en wanneer niet. Het is als leren autorijden: eerst in een rustige woonwijk, pas later op de snelweg.

Ontwikkel wat wij "AI-geletterdheid" noemen. Dat klinkt hoogdravend, maar het betekent simpelweg: leer de patronen herkennen. Wanneer klinkt een antwoord té glad? Welke juridische nuances worden vaak gemist?

Waar zitten de blinde vlekken per rechtsgebied? Ons onderzoek laat zien dat AI-systemen consistent sterk zijn in sommige taken (zoals het structureren van juridische informatie) en zwak in andere (zoals het doorgronden van nieuwe jurisprudentie). Ken die patronen.

Documenteer je ervaringen, deel ze met collega's. Welke prompts werken goed? Welke valkuilen ben je tegengekomen? Waar ging het mis? Deze kennis is goud waard, niet alleen voor jezelf maar voor je hele team. Het voorkomt dat iedereen dezelfde fouten moet maken. Bovendien help je zo mee aan het ontwikkelen van best practices - iets waar de hele juridische gemeenschap van profiteert.

En misschien wel het belangrijkste: blij investeren in je kernexpertise. AI is een tool, geen vervanging voor juridisch oordeelsvermogen. Sterker nog: hoe beter je zelf bent in juridische analyse, hoe beter je kunt inschatten wanneer je AI kunt vertrouwen. Het is als een modern diagnostisch instrument in de geneeskunde: een waardevol hulpmiddel dat data en analyses levert, maar het is de arts die met zijn expertise de juiste diagnose stelt en behandeling bepaalt.



# AI, de Digital Decade en adviesbureau ICTRecht

De conclusies uit ons onderzoek naar AI-systemen weerspiegelen een bredere waarheid over de digitale transformatie: technologie en recht zijn onlosmakelijk met elkaar verbonden. Waar AI-systemen worstelen met juridische nuance, zien we eenzelfde patroon bij de implementatie van nieuwe digitale technologieën. Het gaat niet om de technologie zelf, maar om het vermogen deze op een verantwoorde en juridisch onderbouwde manier in te zetten.

Dit perspectief komt niet uit de lucht vallen. Al twintig jaar zien wij hoe digitale innovatie en juridische frameworks elkaar vormen en versterken. Waar we in 2004 nog vooral bezig waren met hosting en webwinkels, staan we nu voor uitdagingen als DORA in de financiële sector, de European Accessibility Act voor digitale inclusie, en het European Digital Identity Framework dat de basis legt voor digitale identiteiten in Europa.

De Digital Decade brengt een fundamenteel nieuwe dynamiek in het juridische landschap. Waar we vroeger vaak naar één specifieke wet keken voor een oplossing, zien we nu een complex web van onderling verbonden regelgeving. Een modern connected product moet vaak aan vijf of zes verschillende wetten tegelijk voldoen, die elkaar aanvullen en naar elkaar verwijzen.

Neem bijvoorbeeld een moderne auto met rijassistentie. Die moet niet alleen voldoen aan de traditionele typegoedkeuring, maar ook aan:

- De AI Act voor de bestuurdersassistentie
- De Cyber Resilience Act voor digitale beveiliging
- De AVG en Data Act voor sensor- en passagiersgegevens

- De Right to Repair Directive voor softwareupdates
- De Product Liability Directive voor aansprakelijkheid

Deze verwevenheid vraagt om een nieuwe aanpak van juridische dienstverlening. Geen geïsoleerd advies per wet, maar een geïntegreerd raamwerk dat alle relevante wetgeving omvat. Technische standaarden worden daarbij steeds belangrijker als praktisch kompas, maar moeten altijd in de bredere juridische context worden geplaatst.

De kunst is om door de technologische hypes heen te kijken en te focussen op de onderliggende juridische vraagstukken die blijvend relevant zijn. Bij de metaverse ging het uiteindelijk om bekende thema's als digitale identiteit, eigendomsrechten en aansprakelijkheid in online omgevingen. Die expertise blijft waardevol, ook nu de technologie een andere vorm aanneemt.

Uiteindelijk draait het om één fundamenteel inzicht: de digitale transformatie kan niet zonder degelijk juridisch fundament. Een stevige juridische basis geeft organisaties het vertrouwen om te innoveren en de vruchten te plukken van digitalisering. Of het nu gaat om het inzetten van AI-systemen, het ontwikkelen van connected products, of het digitaliseren van dienstverlening - het recht vormt de basis voor verantwoorde innovatie. Deze overtuiging drijft ons al twee decennia, en is relevanter dan ooit in het AI-tijdperk. Want alleen door technologie en recht in balans te brengen, creëren we een digitale toekomst die werkt voor iedereen.

**Aan de slag met AI? Neem contact met ons op:**  
[contact@ictrecht.nl](mailto:contact@ictrecht.nl).

## Onze AI-adviseurs staan voor je klaar

Business  
Development  
Manager  
CAICO®



**Guido  
Grevink**

g.grevink@ictrecht.nl

Legal  
Counsel  
Tech  
CAICO®



**Abel  
Hoogeveen**

a.hoogeveen@ictrecht.nl

Chief  
Knowledge  
Officer  
CAICO®



**Arnoud  
Engelfriet**

a.engelfriet@ictrecht.nl

Senior Legal  
Counsel  
CAICO®



**Marco  
Caspers**

m.caspers@ictrecht.nl

Information  
Security  
Consultant  
CAICO®



**Nina  
Bisscheroux**

n.bisscheroux@ictrecht.nl

Senior Legal  
Counsel  
Tech  
CAICO®



**Niels  
Winters**

n.winters@ictrecht.nl

Talent  
Manager  
CAICO®



**Ardine  
Siepman**

a.siepman@ictrecht.nl

Head of Legal  
Technology  
CAICO®



**Mark  
Zijlstra**

m.zijlstra@ictrecht.nl

Legal  
Counsel  
CAICO®



**Rick  
Kaspers**

r.kaspers@ictrecht.nl

Business  
Development  
Manager  
CAICO®



**Aakriti  
Bhatia**

a.bhatia@ictrecht.nl

Legal  
Counsel  
CAICO®



**Sanne  
Haumersen**

s.haumersen@ictrecht.nl

Business  
Development  
Manager  
CAICO®



**Machiel  
Takens**

m.takens@ictrecht.nl