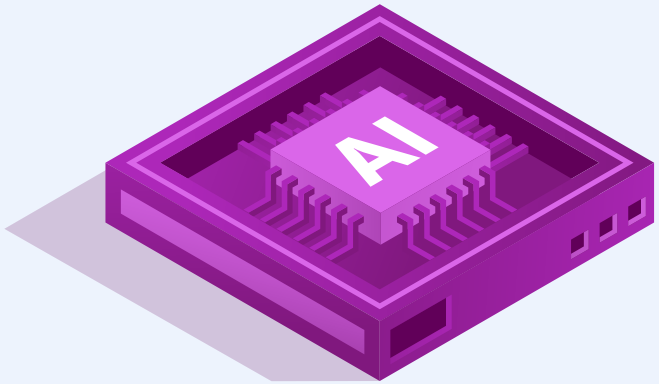
Verantwoord vooruit

AP-visie op generatieve AI

AP | bescherming in een digitale wereld

Februari 2026

Inhoud



Managementsamenvatting

1. **Generatieve artificiële intelligentie (AI) is een impactvolle technologie die bijdraagt aan de maatschappelijke transformatie van vandaag.** Er liggen grote welvaarts- en welzijnsverhogende kansen in bijvoorbeeld de zorg, het onderwijs, onderzoek, en innovatie binnen het bedrijfsleven. Daarbij is het cruciaal om generatieve AI zo in te zetten dat het bijdraagt aan het beschermen en versterken van grondrechten en fundamentele waarden.

2. **De AP ziet mogelijkheid tot rechtmatige ontwikkeling en inzet van generatieve AI onder de Algemene verordening gegevensbescherming (AVG), maar identificeert daarbij uitdagingen en wijst op het belang van technische mitigerende maatregelen.** In 2026 publiceert de AP een juridische analyse op dit onderwerp.

3. **Bij te grote afhankelijkheid van buitenaf kan de Europese Unie (EU) onvoldoende sturen op een verantwoord toekomstbeeld.** Door gebruik van generatieve AI-toepassingen vindt momenteel een ongekend snelle centralisatie van gevoelige data plaats. Deze maakt personen en organisaties kwetsbaar en de EU afhankelijk.

4. **Op samenlevingsniveau pleit de AP voor een aantal maatschappijbrede uitgangspunten om generatieve AI verantwoord te kunnen omarmen.** Hieronder vallen Europese digitale autonomie, maatschappelijke weerbaarheid, democratische sturing, een goed functionerende markt voor verantwoorde oplossingen en het vermogen om door de AI-keten heen te corrigeren. Via maatschappelijke instrumenten kunnen deze uitgangspunten actief worden nagestreefd. Hoofdstuk 8 bespreekt deze onderwerpen.

5. **Ook voor de specifieke inzet van generatieve AI door Nederlandse organisaties ziet de AP een aantal aandachtspunten.** Toepassingen moeten een helder doel hebben en worden voor dat doel gebruikt na een gedegen risico-afweging met passende (AVG-)waarborgen. Andere kernvoorwaarden zijn bijvoorbeeld voldoende AI-geletterdheid bij de inzet en beheersing van systemen en data. Deze kenmerken worden besproken in hoofdstuk 8 en komen ook terug in figuur 1, aandachtspunten inzet generatieve AI.

6. **Nederland heeft een sterke positie om richting te geven aan verantwoorde ontwikkeling van generatieve AI; breed maatschappelijk en politiek debat is hiervoor essentieel.** Keuzes zoals het al dan niet de voorkeur geven aan open-source modellen zijn nu relevant. Dit vereist betrokkenheid van alle lagen van de samenleving: van onderwijs en onderzoek tot ontwikkelaar, van burger tot bestuurder en van consument tot directeur.

7. **De AP draagt met concrete stappen bij aan de verantwoorde ontwikkeling en inzet van generatieve AI.** We zijn bereikbaar via een e-mailloket voor generatieve AI, we organiseren periodieke bijeenkomsten en gaan in gesprek over de juiste vormen van guidance op het onderwerp. Zo creëren we een gedeeld perspectief op de rol van generatieve AI in de samenleving. De AP draagt in Europees verband bij aan verdere normering van de technologie.

FIGUUR 1 | Aandachtspunten inzet generatieve AI

Inzet generatieve AI: aandachtspunten voor organisaties

Dit figuur is gepubliceerd in AP's visie op generatieve AI (Januari 2026).

In het figuur staan aandachtpunten. De AP adviseert met deze punten rekening te houden bij de inzet van generatieve AI. Dit overzicht dient als steun en is geen volledig overzicht voor compliance.

Prioriteit voor bestuurders

Ontwerp

Doel: Veel keuzes rondom de inzet van generatieve AI zijn afhankelijk van het doel, zorg dat deze helder en afgebakend is. Ga na of het wel nodig is om generatieve AI in te zetten voor dit doel, mogelijk volstaat een minder complexe aanpak.

Informatiehuishouding: Overzicht en controle over data en verwerkingen is essentieel voordat componenten die gebruik maken van generatieve AI worden toegevoegd aan een systeem. Denk aan veilige informatieopslag, toegangsbeheer en reactievermogen op nieuwe cyberrisico's.

Wetgeving: Bepaal de relevante wetgeving. Als er persoonsgegevens worden verwerkt is de AVG van toepassing, en afhankelijk van de context kan het product als hoog-risico of verboden classificeren onder de AI-verordening.

AI-geletterdheid: Generatieve AI maakt ontwikkelingen door, met implicaties voor de risico's. Kennis over hoe generatieve AI werkt is vereist om hier met common sense mee om te gaan, en weerbaarheid op te bouwen voor wat je niet ziet aankomen.

Tijdens inzet

AI-geletterdheid: Zorg dat er bij bestuurders en op de werkvloer voldoende kennis is over de werking en risico's van het systeem voor verantwoord handelen bij de inzet. Toets ook of de eindgebruiker zich voldoende bewust is van de gevolgen van het gebruik van generatieve AI in deze toepassing.

Monitoring: Meet of het gebruik van de toepassing verloopt zoals verwacht, zowel gemiddeld als in specifieke gevallen per steekproef.

Incidenten afhandelen: Zorg voor een werkwijze en cultuur waarbij incidenten snel opgemerkt worden. Dit stelt in staat de schade te beperken en te leren. Houd rekening met verplichte rapportering van incidenten.

Modellaag

Verantwoord en rechtmatig: Kies een model dat op een verantwoorde en rechtmatige manier kan worden ingezet.

Open-weights: Neem in de keuze voor een model mee of de parameters van het model openbaar zijn gemaakt.

Evaluatie: Neem de kwaliteiten en testuitkomsten mee in de keuze voor een model. Hierbij kan onder andere gebruik worden gemaakt van modelkaarten, openbare benchmarks of openbare impact analyses. Wees ook bewust van de ecologische impact van modelkeuzes.

Toepassingslaag

Inzichtelijk en transparant: Zorg dat het voor de gebruiker duidelijk is dat deze een generatieve AI-toepassing gebruikt. Hierbij kan op aanvraag meer informatie worden gegeven voor verdere analyse. Dit creëert een basis voor vertrouwen.

Risico management: Risico's dienen op toepassingsniveau te zijn geïdentificeerd, afgewogen en gemitigeerd tot een acceptabel niveau. Maak hierbij gebruik van bestaande overzichten en methoden. Besteed hierbij bewust aandacht aan risico's die typerend zijn voor generatieve AI, zoals:

- *Bias.* Onderzoek of groepen of individuen benadeeld worden door deze toepassing.
- *Hallucinaties.* Onderzoek de gevolgen van uitkomsten die geloofwaardig overkomen maar inhoudelijk onwaar zijn.

Systeemlaag

Controle over ingevoerde data: Gebruikers voeren mogelijk meer gegevens in dan noodzakelijk, en hier dient rekening mee te worden gehouden. Voor de opslag en verwerking van deze data geldt normale wet en regelgeving.

Controle over meta data: Gegevens zoals logging of gebruikersprofielen op de achtergrond dienen ook zorgvuldig behandeld te worden, en zijn van belang voor monitoring, interoperabiliteit en gebruikers gegevenstoegang.

Afhankelijkheid: De snelste oplossingen voor GenAI systemen komen vaak in een vorm die een lock-in positie opbouwt bij de leverancier. Keuzes rondom open-weight modellen en cloudintegratie spelen hierin mee.

Ga hier bewust mee om en weeg aspecten mee zoals:

- Eigenaarschap van ingevoerde data en metadata
- Keuzevrijheid in onderliggende model en modelupdates
- Proces voor kritieke ondersteuning, zoals beveiligingsupdates
- Positie van invloed leverancier over kritieke processen toepassing

Filters: Input filters kunnen relevant zijn om de vrijheden van de gebruiker in te perken, en output filters kunnen ongewenste output blokkeren. Denk hierbij aan het voorkomen van de invoer van persoonsgegevens of haatdragende teksten als output.

Totstandkoming en recente ontwikkelingen

Dit visiedocument is de afgelopen periode door de AP ontwikkeld in een iteratief proces. De eerste bevindingen zijn begin 2025 besproken met een selecte groep experts. Vervolgens is op 23 mei 2025 een consultatieversie van de visie gepubliceerd. Op 28 mei organiseerde de AP in samenwerking met ECP | Platform voor de InformatieSamenleving het evenement *Generative AI: moving forward responsibly* om feedback op te halen onder een brede doelgroep. Per e-mailloket zijn verder schriftelijke bijdragen en feedback ontvangen. De input is verwerkt in de huidige visiedocumenten. Een gedetailleerde beschrijving van opvolging per inputitem is te vinden op de consultatiepagina. De AP houdt de ontwikkelingen rondom generatieve AI nauwlettend in de gaten. Dit visiedocument is definitief en geeft de huidige kijk weer. Wanneer nodig zal de AP aanscherpingen van de zienswijze blijven communiceren, bijvoorbeeld via de periodieke risicorapportages of nieuwe guidance documenten.

Tijdens deze periode zijn de ontwikkelingen verder gevorderd. Deze publicatie van het definitieve visiedocument komt zo'n 8 maanden na de publicatie van de consultatieversie. Hierbij een greep uit de meest in het oog springende recente ontwikkelingen:

- **Kleine, open, specialistische modellen.** Met behulp van *reinforcement learning* (zie hoofdstuk 2) worden modellen op steeds meer diverse doelen getraind, waardoor modellen met een sterker specialisme ontstaan. De trend om de geleerde patronen in zo weinig mogelijk parameters op te slaan zet door, en het aanbod van open-weight modellen (zie hoofdstuk 4) groeit. Deze eigenschappen dienen als aanjagers voor agentic AI.

- **Verantwoorde initiatieven.** De General Purpose AI (GPAI) Code-of-Practice is aangenomen als ondersteunend instrument van de AI-verordening in de EU. De meeste grote generatieve AI-aanbieders hebben toegezegd hieraan te gaan voldoen. Er worden modellen uitgebracht waarbij Europese waarden centraal staan, zoals het Zwitserse Apertus[1] en de TildeOpen LLM.[2] De training van GPT-NL is in volle gang[3], en voor gevoelige toepassingen neemt de populariteit toe van het in eigen beheer draaien van generatieve AI-modellen.[4]

- **Commercialisering.** Bij concrete toepassingen valt een steeds verdere commercialisering van de technologie op. Dit is een te verwachten ontwikkeling bij nieuwe technologieën. Dit heeft echter ook maatschappelijke impact. Tools die in eerste instantie gratis beschikbaar waren, gaan steeds meer op zoek naar businessmodellen, zoals gebruik op rekening, integratie van advertenties en uitvoering van online aankopen.

- **Geopolitiek.** China en de Verenigde Staten (VS) hebben een AI-strategie gepresenteerd en nemen politieke keuzes met betrekking tot generatieve AI. Dit heeft onder andere betrekking op investeringen, regelgeving en internationale handel. De VS moedigt het aanbieden van open-source generatieve AI aan omdat dit mogelijk de wereldwijde standaard wordt, en de VS vindt het belangrijk dat deze modellen op Amerikaanse opvattingen gebaseerd zijn. De huidige regering van de VS adviseert om referenties naar misinformatie, diversiteit, gelijkheid, inclusiviteit en klimaatverandering te verwijderen uit het AI-risicomanagement framework van de NIST.[5] Deze handeling strookt niet met Europese waarden.

De AP publiceert in 2026 een handreiking met een juridische analyse op AVG aspecten van de ontwikkeling en inzet van generatieve AI-modellen. Gelijktijdig met de consultatie van dit visiedocument is de inhoud van de handreiking geconsulteerd.

---

1 https://ethz.ch/en/news-and-events/eth-news/news/2025/09/press-release-apertus-a-fully-open-transparent-multilingual-language-model.html
2 https://digital-strategy.ec.europa.eu/en/library/eu-funded-tildeopen-llm-delivers-european-ai-breakthrough-multilingual-innovation
3 https://gpt-nl.nl/nieuws/start-training-gpt-nl/
4 https://www.rijksorganisatieodi.nl/actueel/nieuws/2025/05/08/de-impact-van-ai-op-de-rijksoverheid
5 https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf

1. Inleiding

Generatieve AI heeft de afgelopen 3 jaar een periode doorgemaakt van snelle ontwikkeling en maatschappelijke adoptie. De technologie is onderdeel geworden van de digitale basisvoorzieningen van onze samenleving. Dat brengt een hele nieuwe set aan kansen en uitdagingen met zich mee: op economisch, technologisch, juridisch en maatschappelijk vlak.

Een impactvolle ontwikkeling zoals de opkomst van generatieve AI verdient de volle aandacht van de samenleving. De AP stuurt op een toekomst waarin digitale technologie ten dienste staat van de samenleving en het maatschappelijk belang. Hiervoor draagt de AP bijvoorbeeld bij aan een brede waaier van grondrechten, zoals het recht op onderwijs, de vrijheid van ondernemerschap, het recht op sociale zekerheid en sociale bijstand, of het recht op gezondheidsbescherming. Een uitgekiende en verantwoorde omarming van generatieve AI maakt het mogelijk om op al deze terrein stappen te zetten.

Tegelijkertijd maakt de technologische werking van generatieve AI dat grondrechten juist extra waarborging en bescherming verdienen. Denk aan het recht op non-discriminatie en het recht op een eerlijk proces, het recht op eigendom — zoals auteursrechten — of het recht op de bescherming van persoonsgegevens. Op het gebied van gegevensbescherming ziet de AP in dat er verschillende dillema's bestaan wat betreft de rechtmatigheid van de ontwikkeling en inzet van generatieve AI. De European Data Protection Board (EDPB) heeft vastgesteld dat de inzet van deze foundation modellen door Nederlandse en Europese partijen niet per definitie onrechtmatig is. Verdere standpunten over de rechtmatigheid van de inzet van generatieve AI moeten op Europees niveau nog worden uitgewerkt. De AP acht de rechtmatige ontwikkeling en inzet van generatieve AI mogelijk, maar ziet ook uitdagingen op andere gebieden dan gegevensbescherming.

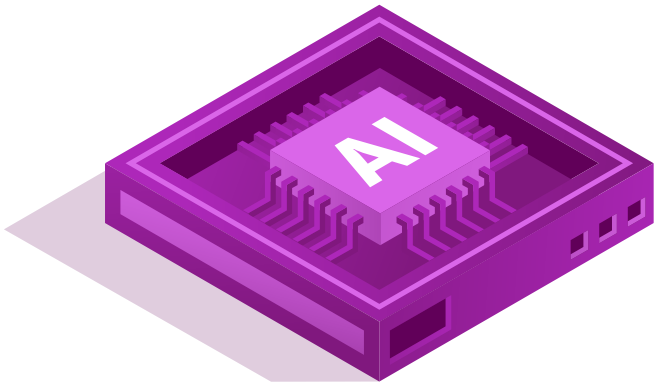
In dit visiestuk kijkt de AP met een toekomstgerichte blik naar generatieve AI: wat verstaan wij eronder, welke kansen biedt de technologie ons, met welke risico's gaat het gepaard en wat moet er gebeuren om generatieve AI verantwoord te omarmen? We kijken hierbij vanuit een helikopterperspectief naar de bescherming van onze grondrechten. Deze visie hanteert een brede scope, zowel voor de technologie zelf als de impact op de maatschappij. Chatbots, maar ook generatieve AI voor molecuulstructuren, vallen binnen de reikwijdte van de visie voor verantwoorde inzet van deze generatieve AI. Dit doen we niet alleen vanuit onze rol als toezichthouder op de AVG en Richtlijn voor gegevensbescherming bij rechtshandhaving (RGR), maar ook als coördinerend AI- en algoritmetoezichthouder, en om een bijdrage te leveren aan de voorbereiding van het toezicht op de AI-verordening.

Deze visie is in de eerste plaats gericht op professionals die in hun werk te maken krijgen met generatieve AI. Dit bedoelen we in brede zin: ontwikkelaars die dagelijks beslissingen maken over wat generatieve AI wel en niet kan, maar ook bestuurders en managers die moeten besluiten over de inzet en het gebruik van generatieve AI in hun organisatie.

De AP beoogt met deze visie bij te dragen aan het maatschappelijk debat over generatieve AI en de contouren te schetsen van wat er nodig is voor een veilige en verantwoorde inzet hiervan. Daarbij realiseert de AP zich maar al te goed dat de technologische ontwikkelingen elkaar razendsnel opvolgen, en de kennis van morgen alweer anders is dan de kennis van vandaag. Deze visie is daarom geen statische analyse, of een uitputtend overzicht van voorbeelden, kansen en risico's. De visie moet worden gezien als een uitnodiging om hierover met elkaar in gesprek te gaan.. Als samenleving moeten we vorm geven aan de rol die generatieve AI speelt in de maatschappij.

2. Generatieve AI: beeld van de technologie

Generatieve AI is niet meer weg te denken uit ons dagelijks leven. Maar wat we precies verstaan onder generatieve AI kan per gesprek verschillen, waardoor soms verwarring ontstaat. In dit hoofdstuk lichten we toe hoe we generatieve AI in deze visie benaderen. Daarnaast schetsen we een beeld van de laatste technologische ontwikkelingen.



Generatieve AI

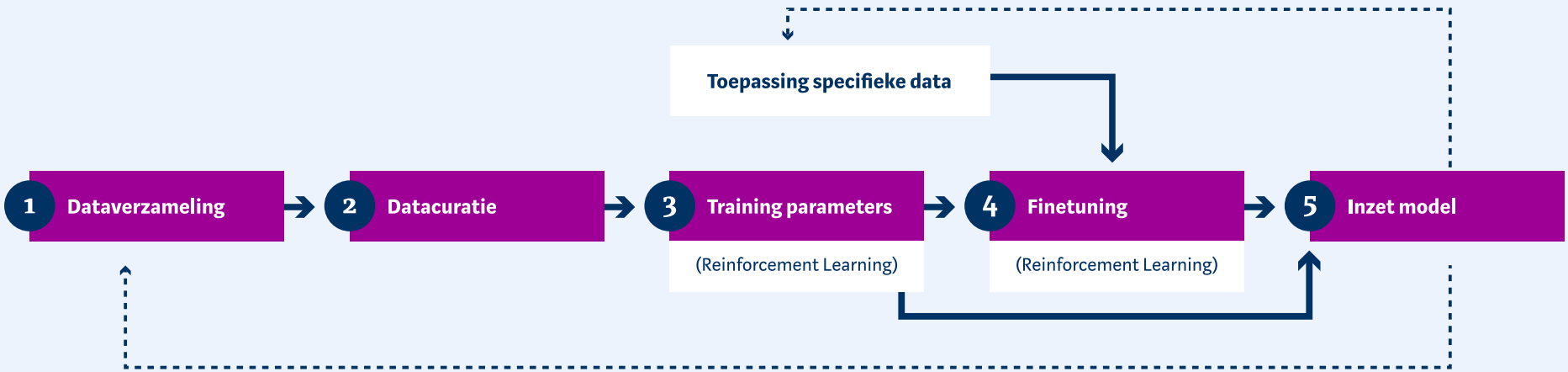
Generatieve AI is een vorm van AI die in staat is om nieuwe data te genereren. De meest populaire toepassingen maken teksten en afbeeldingen waar niet meer aan af is te zien dat deze door AI gegenereerd zijn. Deze visie gaat over AI-modellen die in staat zijn realistische data te genereren en over de systemen en applicaties waar die modellen onderdeel van zijn.

Generatieve AI-modellen kunnen allerlei verschillende vormen van output genereren en daarin gestuurd worden. Hiermee kunnen deze modellen als fundament dienen voor veel verschillende specialistische toepassingen en ingezet worden voor allerlei doeleinden. Dit type modellen wordt vaak *foundation models* of *general purpose AI-models* genoemd. In de gekozen benadering vallen modellen met die noemers onder generatieve AI.

Output van generatieve AI is gebaseerd op kansrekening en inherent vatbaar om onwaarheden te bevatten. Dit is een gevolg van hoe de technologie werkt. Generatieve AI werkt op basis van patronen en voorbeelden in andere gegevens. Keuzes tijdens de training hebben invloed op de uitkomsten. Daarnaast hebben generatieve AI-modellen geen wereldbeeld met feiten, maar zijn in eerste instantie geoptimaliseerd om data te genereren die realistisch overkomen. Als je vraagt naar de naam van de pizzeria om de hoek, is er kans dat een generatief AI-model dit met de beste voorspelling invult, ook al bestaat er helemaal geen pizzeria om de hoek.

Buiten de scope van deze visie liggen tal van andere vormen van AI. AI-technologie is al lange tijd in ontwikkeling en wordt bijvoorbeeld veel gebruikt voor classificatie. De AI is dan alleen getraind om patronen te herkennen, in plaats van ook te genereren.

FIGUUR 2: Training van een generatief AI-model



Simpel gezegd kan het trainen van een generatief AI-model worden gezien als een set opeenvolgende stappen. Omdat de gegenereerde data in veel gevallen weer worden verzameld voor training, is er sprake van een feedbackloop:

Stap 1: Dataverzameling. Voorbeelddata worden verzameld, vaak door scraping.

Stap 2: Datacuratie. Ongewenste voorbeelden (zoals persoonsgegevens of haatdragende inhoud) worden verwijderd in een curatiestap.

Stap 3: Training parameters. De parameters worden getraind op basis van patronen in voorbeelddata en mogelijk aangescherpt door *reinforcement learning*.

Stap 4: Finetuning. Door finetuning wordt het model afgestemd op een specifieke toepassing of begrenzing.

Stap 5: Inzet van het model. De uitkomst van deze stappen is een getraind generatief AI-model dat wordt ingezet in een toepassing.

Technologische ontwikkelingen

De technologie achter generatieve AI is nog volop in ontwikkeling. In onderstaande figuur geven we een overzicht van een aantal belangrijke ontwikkelingen. Dit overzicht is niet compleet, maar geeft een indruk van ontwikkelingen binnen verschillende lagen van de technologie. We maken onderscheid tussen ontwikkelingen in de modellen zelf (modellaag), ontwikkelingen over de systemen waarin de modellen worden ingezet (systeem-laag), en de toepassingen waarlangs die systemen contact maken met de omgeving (toepassingslaag).

FIGUUR 3| schematische weergave van lagen en ontwikkelingen van generatieve AI

In alle 3 de lagen van de technologie zien we de afgelopen jaren veel ontwikkelingen. Binnen de modellaag, die uiteindelijk de kern van de technologie vormt, experimenteren wetenschappers en ontwikkelaars met verbeteringen in trainingsalgoritmes en aanpassingen op modelarchitectuur.[6] Doordat veel – ook geavanceerde – modellen openbaar beschikbaar zijn om te downloaden, kan iedereen hierop voortbouwen. Binnen de systeemlaag dragen innovaties bij om de modelcapaciteiten verder te benutten. Het verbinden van een model met een database kan er bijvoorbeeld voor zorgen dat een systeem ook geheugen gebruikt, in plaats van alleen te reageren op de laatste input.[7] In de toepassingslaag zien we veel vernieuwing in de manier waarop de systemen interacteren met de omgeving. Zo is er veel aandacht voor autonome *AI Agents*[8], een type applicatie dat autonoom taken uitvoert, waarover meer in de beschreven trends in hoofdstuk 4.

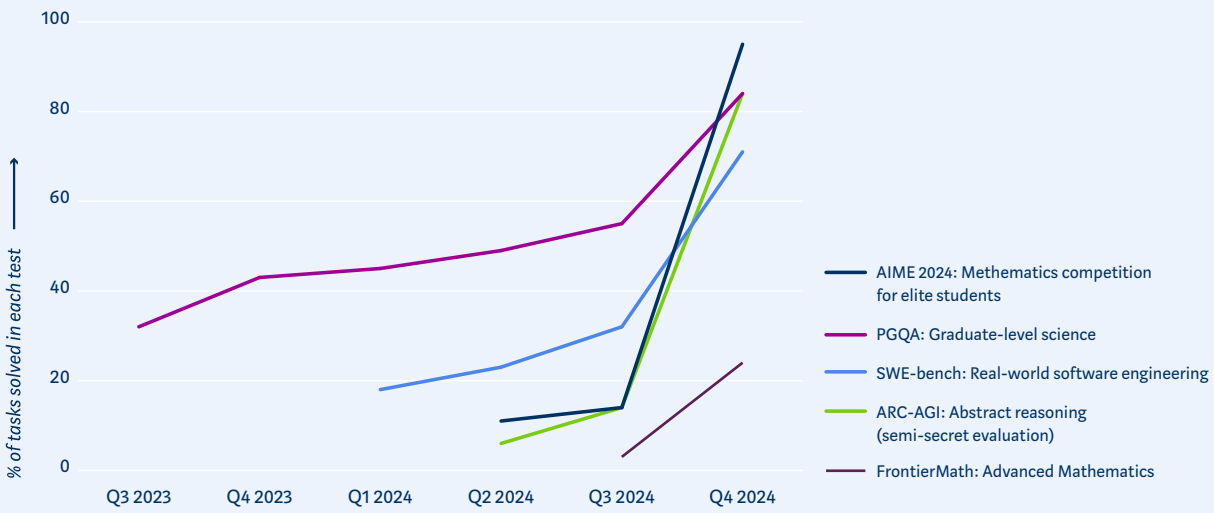
Nieuwe modellen maken nieuwe toepassingen mogelijk, waardoor data ontstaat die terug kan vloeien naar modeltraining. Het afgelopen jaar is die cyclus waar te nemen in Visual Language Models (VLMs), modellen op basis van tekst en beeld. Populaire chatapps konden op basis van de eerste multimodale modellen een plaatje genereren, die de gebruiker met enkele prompts aanscherpte. De laatste generatie modellen maakt realistische video's en wordt geïntegreerd in tekstuele, visuele en zelfs sociale interactie.

6 Mistral AI (December 2023). "Mixtral of experts", Meta (December 2024). "Large Concept Models: Language Modeling in a Sentence Representation Space", Deepseek (Januari 2025¬). "DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning".

7 Cohere. "Retrieval Augmented Generation (RAG)" OpenAI (September 2024). "Learning to reason with LLMs".

8 Anthropic (December 2024). "Building effective agents".

FIGUUR 4 | Belangrijke benchmarks door de tijd



Capaciteiten van generatieve AI

Als gevolg van de bovengenoemde ontwikkelingen is de kwaliteit van generatieve AI-output snel toegenomen. Om dit te meten worden modellen en systemen onderworpen aan allerlei tests en benchmarks, bijvoorbeeld met multiple-choice vragen over biologie en natuurkunde.[9] In 2024 leek de vooruitgang in generatieve AI-technologie te vertragen. Een steeds groter deel van de informatie op het internet is door AI gegenereerde content. Dit belemmert het verzamelen van hoge kwaliteit trainingsdata. Eind 2024 was er juist weer een versnelling in de toename van kwaliteit. Het toepassen van *reinforcement learning* is hier een belangrijk onderdeel van. Hierbij krijgt het model de vrijheid om antwoordstrategieën te ontwikkelen, en een prikkel om voort te bouwen op strategieën die goede antwoorden opleveren. Een model kan op deze manier bijvoorbeeld leren door terug te kijken op eigen antwoorden en die mogelijk ook te verbeteren.[10]

Het is de vraag in welke mate deze ontwikkelingen zich de komende jaren doorvertalen in generatieve AI-modellen die tot nog veel meer taken in staat zijn. Er bestaat nog grote onzekerheid over het verdere verloop van de kwaliteiten van deze technologie.[11] We kunnen wel aannemen dat, ook al zou er geen vooruitgang meer zijn in de modellaag, er op basis van de *bestaande* modellen veel nieuwe toepassingen bijkomen de komende jaren. Verdere vooruitgang in de modellaag zal dit dus alleen versterken.

[9] Rein, D., Hou, B. L., Stickland, A. C., Petty, J., Pang, R. Y., Dirani, J., ... & Bowman, S. R. (2024). Gpqa: A graduate-level google-proof q&a benchmark. In First Conference on Language Modeling.

[10] Guo, D., Yang, D., Zhang, H., Song, J., Zhang, R., Xu, R., ... & He, Y. (2025). Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *arXiv preprint arXiv:2501.12948*.

[11] Bengio, Y., Mindermann, S., Privitera, D., Besiroglu, T., Bommasani, R., Casper, S., ... & Zeng, Y. (2025). International AI Safety Report. arXiv preprint arXiv:2501.17805.

3. Inzet van generatieve AI: markt en toepassingen

Generatieve AI vindt razendsnel ingang in zowel de private als publieke sector. Ook voor persoonlijke doeleinden maken Nederlanders steeds vaker gebruik van toepassingen die gebaseerd zijn op generatieve AI. De belofte is groot en de verwachtingen zijn hooggespannen, waardoor organisaties over de gehele linie zich voorbereiden op een verdere gebruiks-toename in de nabij toekomst. Bij dit gebruik komt een grote centralisatie van gevoelige gegevens kijken.

De markt voor generatieve AI heeft sinds eind 2022 een explosieve groei doorgemaakt. Met de lancering van ChatGPT door OpenAI is een periode van ongekend snelle maatschappelijke adoptie ingezet waarvan het einde nog niet in zicht is. Dit heeft ook geleid tot een stevige concurrentiestrijd tussen technologiebedrijven wereldwijd, waardoor de ontwikkeling en inzet van generatieve AI inmiddels nauw verweven is geraakt met geopolitieke ontwikkelingen.

Veel sectoren werken en experimenteren inmiddels met generatieve AI. Het is belangrijk om dit gebruik in kaart te brengen om een goede inschatting te kunnen maken van de risico's en veiligheid van generatieve AI. Modellen worden doorgaans uitvoerig getest voordat ze op de markt worden gebracht, maar het gebruik in de praktijk (en de toepassingen die op de modellen voortbouwen) kan nooit volledig worden nagebootst – terwijl de risico's en incidenten zich juist daar manifesteren. Daarnaast gaat het gebruik gepaard met de verwerking van grote hoeveelheden data die de gebruiker afhankelijk en kwetsbaar kunnen maken. Het is daarom van belang dat we hier goed zicht op krijgen, zodat we kunnen leren wat er nodig is voor verantwoord gebruik en hoe we de voornaamste risico's kunnen mitigeren.

Private sector: inzet in allerlei bedrijfsprocessen

Binnen de private sector wordt generatieve AI ingezet in allerlei bedrijfsprocessen. De private sector loopt voorop als het gaat om de inzet van generatieve AI. In *The State of AI 2025* van McKinsey geeft 71% van de ondervraagde organisaties aan regelmatig gebruik te maken van generatieve AI: een verdubbeling ten opzichte van hetzelfde onderzoek een jaar eerder.[12] Capgemini schetst een vergelijkbaar beeld in hun onderzoek naar het gebruik van generatieve AI in verschillende sectoren. Dit is in 2 jaar tijd in alle domeinen sterk gegroeid.[13] Organisaties zetten generatieve AI voor een grote verscheidenheid aan bedrijfs-processen in, vaak met als doel om de efficiëntie en productiviteit te vergroten – van IT tot logistiek en van HR tot marketing. Het illustreert de veelzijdigheid van generatieve AI en laat ook zien dat mensen in heel verschillende functies in hun werk te maken zullen krijgen met een vorm van generatieve AI.

[12] McKinsey (maart 2025). "The state of AI".
[13] Capgemini Research Institute (2025). "Generative AI in organizations in 2025".

Publieke sector: experimenten in de kinderschoenen

De publieke sector experimenteert met generatieve AI, maar dit staat vaak nog in de kinderschoenen.
De quickscan van TNO uit juli 2024 laat zien dat in de publieke dienstverlening in allerlei domeinen steeds meer gebruik wordt gemaakt van AI.[14] Het aandeel generatieve AI was hierin ten tijde van het onderzoek nog laag, maar dat kunnen we deels verklaren door het feit dat veel experimenten nog in de kinderschoenen stonden.
Ook was het individuele gebruik van ambtenaren hier niet in meegenomen. TNO wijst hierbij ook op het risico van 'schaduw-IT': technologie of software die door individuele medewerkers binnen een organisatie wordt gebruikt zonder formele goedkeuring van de IT-afdeling. Het overheidsbrede standpunt (april 2025) over het gebruik van generatieve AI binnen overheidsorganisaties is in deze context ook van belang. Dat standpunt stelt dat het gebruik van generatieve AI wordt aangemoedigd, mits hierover eigen afspraken worden gemaakt met leveranciers en dit niet onder consumentenvoorwaarden valt.[15] Op basis van dit standpunt is het waarschijnlijk dat het gebruik van generatieve AI in de overheidssector hierdoor de komende tijd toeneemt.

Persoonlijk gebruik: zichtbare generatiekloof

Wat betreft het persoonlijk gebruik van generatieve AI is een generatiekloof zichtbaar. In 2024 maakte ruim 1 op de 3 Nederlanders regelmatig gebruik van generatieve AI.[16] De verschillen tussen leeftijdsgroepen zijn groot: zo geeft meer dan 50% van de ondervraagden tussen de 18 en 34 jaar aan dat zij generatieve AI gebruiken, terwijl dit in de leeftijdscategorie 65 tot 75 jaar slechts 10% is. Dit generatieverschil is niet vreemd: jongeren van nu zijn opgegroeid met het internet en hebben daardoor vaak meer vertrouwen om te experimenteren met nieuwe technologieën in hun dagelijks leven.[17] Mensen gebruiken generatieve AI-tools vooral voor het zoeken en verzamelen van informatie en het genereren van ideeën. Een ruime meerderheid van de Nederlanders gelooft dat generatieve AI hun werk de komende 2 jaar gemakkelijker (77%) en leuker (77%) zal maken.[18] Er zijn ook chatbotapps die zich specifiek richten op de mentale gezondheid van gebruikers en claimen deze te verbeteren. In een eerdere rapportage besteedde de AP hier aandacht aan, waaruit bleek dat verkeerde inzet van chatbots serieuze risico's heeft voor mensen die op zoek zijn naar hulp bij mentale problemen.[19]

Praktijkvoorbeelden

In de praktijk heeft de inzet van generatieve AI allerlei verschillende uitingsvormen. Deze fictieve voorbeelden hieronder zijn gebaseerd op bestaande toepassingen en bedoeld om een indruk te geven van wat er op dit moment mogelijk is met generatieve AI in verschillende contexten.

Accountancy

Financieel advies op maat

Accountantskantoor *Loukili & Laheij* heeft flinke stappen gezet sinds de aanschaf van fiscale software op basis van generatieve AI. Standaardrapportages en financiële prognoses kunnen nu in een paar minuten worden gegenereerd. De accountants van Loukili & Laheij hoeven deze alleen nog te valideren, waardoor zij meer tijd kunnen besteden aan persoonlijk contact en financieel advies op maat. De software ondersteunt hen hier ook bij door op basis van de historische klantdata en de klantvraag suggesties te doen voor oplossingsrichtingen. Voor klanten is dit van grote toegevoegde waarde.

De software blijkt ook effectief bij foutdetectie. Tegelijkertijd schuilt hierin ook een risico voor Loukili & Laheij. Het gevaar is dat medewerkers te veel gaan vertrouwen op de software en output onvoldoende kritisch beoordelen, terwijl dit juist bij jaarlijks wijzigende wetgeving en belastingsystemen van belang is.

14 TNO (Juni 2024). "Steeds meer kunstmatige intelligentie ingezet door overheid".
15 Rijksoverheid (April 2025). "Het overheidsbrede standpunt voor de inzet van generatieve AI".
16 KPMG (202). "Algoritme Vertrouwensmonitor 2024".
17 Algosoc (2024), "AI Opinion Monitor".
18 Deloitte (Oktober 2024). "Vertrouwen in generatieve AI: een Europees en Nederlands perspectief".
19 AP (Februari 2025). RAN4,

Onderwijs

Digitale leescoach

De 8-jarige Noor krijgt op basisschool *De Ekster* elke week les van een gepersonaliseerde digitale leescoach. Noor heeft dyslexie en komt daardoor moeilijk mee in de reguliere lessen over begrijpend lezen. Dat werkte frustrerend en demotiverend. Op de computer kan ze zich beter concentreren en krijgt ze teksten te lezen die speciaal op haar zijn afgestemd. Bijvoorbeeld een tekst over Disney, waar ze groot fan van is. De teksten die haar digitale leescoach genereert sluiten niet alleen aan bij haar interesses, maar bevatten ook specifieke woorden en spellingspatronen die op haar niveau zijn afgestemd. Voor Noor is het lezen zo niet alleen leuker geworden, maar ze gaat ook sneller vooruit.

Huisarts

Chatbot als doktersassistent

Huisartsenpraktijk *ZorgModern* heeft vanaf de start gekozen voor een chatbot als dokters-assistent: niet alleen om geld te besparen, maar ook omdat goed personeel lastig te vinden was. De chatbot is nu het eerste contactpunt voor patiënten van de praktijk. Via een app omschrijven patiënten hun klachten. De chatbot stelt een paar vervolgvragen en maakt een samenvatting die naar de huisarts wordt gestuurd. De huisarts beoordeelt het bericht en stuurt de patiënt een persoonlijk antwoord: een advies of een afspraakverzoek.

Patiënten zijn over het algemeen tevreden met het korte lijntje tussen hen en de arts. Oudere patiënten voelen zich echter minder thuis bij deze praktijk: omdat zij onvoldoende digitaal vaardig zijn, voelt de app als een obstakel voor goede zorg.

Privégebruik

Persoonlijke vakantieadviseur

Spanje, Duitsland, Franrijk, of toch een verre reis naar Thailand of Zuid-Afrika? Is het daar wel het goede seizoen voor? En zijn 3 weken voldoende om genoeg te kunnen zien? Vragen waarvoor je eerst een paar avonden het internet afstruinde op zoek naar antwoorden, kun je nu in een paar minuten laten beantwoorden door een taalmodel. En met gebruik van de juiste prompts genereer je in een handomdraai een gedetailleerd reisplan dat aan al jouw wensen en eisen voldoet: van aanbevelingen voor hotels en restaurants tot tips voor natuurparken en culturele hoogtepunten. Of informatie over lokale eetgewoonten en noodnummers. En ook tijdens je reis staat de digitale persoonlijke vakantieadviseur altijd voor je klaar.

Box 1

Afhankelijkheid door gebruik

Door het gebruik van generatieve AI vindt er momenteel een ongekend snelle centralisatie van gevoelige data plaats. Op ChatGPT zijn ondertussen, los van zakelijk gebruik, zo'n 800 miljoen mensen wekelijks actief, grofweg 10% van de wereldbevolking. In september 2025 beschreef een studie deze gebruikers in samenwerking met OpenAI. Ongeveer de helft van de gebruikers is onder de 26 jaar. Het aandeel vrouwen is de afgelopen periode gegroeid, en nu grofweg gelijk aan het aandeel mannen. De hoeveelheid gebruikersdata is dus enorm, maar de gevoeligheid van deze data is ook opvallend.[20] Zo'n 11% van de berichten wordt geclassificeerd als 'expressing', vaak het uiten van gedachten of gevoelens zonder dat daar een directe actie of vraag aan gekoppeld is. In interviews met de NOS gaven jongeren aan dat ze sommige problemen alleen met AI bespreken en geen professionele hulp zochten.[21] Zelf beschrijft OpenAI de gebruikersaccounts als 'mogelijk een van de meest gevoelige accounts die mensen ooit zullen hebben'.[22]

Het gebruik van generatieve AI is gecentraliseerd bij een klein aantal aanbieders. Dit zorgt voor afhankelijkheid van de maatschappij. OpenAI is een van de enkele aanbieders waarvan generatieve AI op deze schaal gebruikt wordt. Gebruikers bouwen afhankelijkheid op, bijvoorbeeld mensen die steeds meer rekenen op een gepersonaliseerde service, of een bedrijf dat de klantenservice heeft vervangen door AI-assistenten. Omdat het gaat om gebruik door een groot deel van de burgers en organisaties in de samenleving ontstaat hier een maatschappijbrede afhankelijkheid.

Technologische afhankelijkheid wordt de laatste jaren ingezet als drukmiddel en maakt burgers en de EU kwetsbaar. Als de afhankelijkheid te groot wordt heeft de maatschappij niet langer het laatste woord over hoe centrale processen eruitzien, zoals de informatievoorziening tijdens verkiezingen.[23] In de afgelopen jaren is technologische afhankelijkheid verschillende keren zichtbaar geworden bij internationale conflicten. De communicatie via Starlink werd stilgelegd tijdens een offensief in Oekraïne[24], en Microsoft zette het e-mailaccount van een medewerker van het Internationaal Strafhof in Den Haag uit na een uitvoerend bevel van president Trump.[25] De Amerikaanse regering stelde op de AI-summit in Parijs dat de VS geen strengere regulering van Amerikaanse AI-bedrijven door buitenlandse overheden zal accepteren.[26] Onder de CLOUD-Act kan de Amerikaanse overheid onder voorwaarden gebruikersdata bij Amerikaanse bedrijven opvragen.[27]

Bij een te grote afhankelijkheid van buitenaf kan de EU onvoldoende sturen op verantwoord gebruik. Deze afhankelijkheid kan worden verminderd door het vergroten van het Europese aanbod of beheer van generatieve AI, en een lage drempel voor gebruikers om over te stappen tussen services. Modellen op basis van openbare toegang hebben hun eigen risico's, maar kunnen Europese organisaties wel een vliegende start geven bij de ontwikkeling van generatieve AI-modellen en toepassingen. Een voorbeeld van Europese regelgeving die bijdraagt aan gezonde mate van afhankelijkheid is de Dataverordening.[28] Deze wet is in september 2025 in werking getreden en moet de interoperabiliteit tussen systemen vergroten.

---

20 https://www.nber.org/system/files/working_papers/w34255/w34255.pdf

21 https://nos.nl/artikel/2584070-chatgpt-wijst-mensen-te-snel-door-merkt-zelfdodinghulplijn-113

22 https://openai.com/nl-NL/index/teen-safety-freedom-and-privacy/

23 https://www.autoriteitpersoonsgegevens.nl/actueel/ap-onstuimige-opkomst-ai-vraagt-om-waakzaamheid-van-iedereen, https://www.tno.nl/nl/newsroom/2023/11/ai-taalmodellen-inconsistent-links/, https://www.rathenau.nl/nl/digitalisering/naar-een-nieuwe-verhouding-tot-technologiebedrijven/digitale-afhankelijkheid-beeld

24 https://www.reuters.com/investigations/musk-ordered-shutdown-starlink-satellite-service-ukraine-retook-territory-russia-2025-07-25/

25 https://www.nytimes.com/2025/06/20/technology/us-tech-europe-microsoft-trump-icc.html

26 https://www.euronews.com/2025/02/11/jd-vance-challenges-europes-excessive-regulation-of-ai-at-paris-summit

27 https://www.ncsc.nl/actueel/weblog/weblog/2022/de-werking-van-de-cloud-act-bij-dataopslag-in-europa

28 https://digital-strategy.ec.europa.eu/en/factpages/data-act-explained

4. Trends in generatieve AI: opkomende toepassingen

Generatieve AI heeft de potentie om te zorgen voor een fundamentele verandering in de manier waarop we werken, communiceren en informatie tot ons nemen. Van autonome uitvoerders (AI-agents) tot zoekmachines en nieuwe virtuele sociale actoren waarmee we in contact staan. De technologie biedt op al deze terreinen nieuwe mogelijkheden, maar stelt ons ook voor uitdagingen en kan direct en indirect impact hebben op grondrechten en fundamentele waarden.

De impact van generatieve AI zien we in de praktijk en hangt samen met de vorm waarin het wordt ingezet door organisaties en individuen. Daarbij wordt nog volop geëxperimenteerd met nieuwe toepassingsmogelijkheden. In het hoofdstuk 'Beeld van de technologie' gaven we een omschrijving van de technologische ontwikkelingen op het gebied van generatieve AI. Deze ontwikkelingen liggen ten grondslag aan de toepassingsmogelijkheden in de praktijk. Maar op welke manier brengt generatieve AI nu verandering te weeg? Dat hangt af van de wijze waarop de technologie wordt ingezet. We behandelen hier een aantal trends in toepassingsvormen die we zien met betrekking tot generatieve AI.

Generatieve AI als autonome uitvoerder (AI Agents): het idee om AI-modellen zelfstandig acties te laten uitvoeren om taken te automatiseren is niet nieuw, maar generatieve AI schept hiervoor nieuwe mogelijkheden. Ten eerste kunnen opdrachten die in menselijke taal zijn omschreven beter worden geïnterpreteerd door grote taalmodellen. Doelen worden daarbij in brokjes opgedeeld en stapsgewijs uitgevoerd. Daarnaast kunnen deze modellen instructies genereren voor allerlei verschillende instrumenten, bijvoorbeeld om de muis en het toetsenbord van computers virtueel te bedienen.[29] Deze verbeteringen samen zorgen ervoor dat het veld dichter bij autonome uitvoerders (AI agents) komt. Een enkele AI-ontwikkelaar stelt dat er eind 2027 modellen zijn die bijna alles kunnen wat de meeste mensen kunnen.[30] Deze autonome uitvoerders kunnen altijd aanstaan en op de achtergrond werken, waardoor ze zo goed als onzichtbaar zijn. Voor bedrijven kunnen autonome uitvoerders een uitkomst bieden, maar de inzet heeft mogelijk onder meer ook invloed op de werkgelegenheid.

Generatieve AI als 24/7 persoonlijke assistent: de integratie van generatieve AI in bestaande software en de combinatie van efficiëntere modellen met sterkere mobiele hardware en grotere internetbandbreedte zal het mogelijk maken om een (lokale) generatieve AI als assistent te gebruiken die continu aanstaat. Als mensen doorlopend aantekeningen willen laten maken gedurende de dag zou de AI-assistent continu mee moeten luisteren. Hierdoor zou het veel persoonlijke details mee kunnen krijgen, ook van gesprekspartners die geen besef hebben dat de assistent aanstaat, of daar geen controle over kunnen uitoefenen.

Generatieve AI als sociale actor: de output die een generatieve AI genereert lijkt heel erg op iets dat door mensen is gemaakt. Je kunt dus het gevoel hebben dat je menselijk contact hebt als je interacteert met een virtuele (chat)bot. Dit geeft AI de kans om op een nieuw vlak te opereren. Sommigen ervaren het als virtuele vriend of geliefde in zogenaamde *companion apps*, of als digitale therapeut die helpt bij mentale problemen. Ook kan een

[29] Anthropic. "Computer use (beta)".
[30] ArsTechnica (Januari 2025). "Anthropic chief says AI could surpass "almost all humans at almost everything" shortly after 2027".

generatieve AI als leraar een lesprogramma personaliseren voor een leerling, of helpen om een sollicitatie te oefenen, door een sociale setting te stimuleren. Tenslotte worden chatbots steeds vaker gebruikt om front office-achtige taken over te nemen, zoals interactie met klanten en het maken van afspraken. De AP en de Autoriteit Consument & Markt (ACM) riepen organisaties onlangs op om hun verantwoordelijkheid te nemen als zij ervoor kiezen om chatbots in te zetten.[31]

Generatieve AI als hulpmiddel voor optimalisatie en efficiëntie in organisaties: een van de beloftes van generatieve AI is dat deze routinematige cognitieve taken van werknemers over kan nemen. Bijvoorbeeld door (online) vergaderingen te transcriberen en samen te vatten, of door klantvragen over producten te beantwoorden. Dit zou potentieel de productiviteit kunnen verhogen, maar het is nog maar de vraag of dit daadwerkelijk zo is. Het uitvoeren van routinematige taken door de medewerker kan juist verlichting bieden, waardoor cognitief zware taken, die nog niet geautomatiseerd (kunnen) worden, weer uitgevoerd kunnen worden. De routinematige taken dragen dus bij aan de totale productiviteit: als deze wegvallen, is het niet evident dat meer tijd voor zware taken ook zorgt voor meer of betere uitvoering daarvan.[32]

Generatieve AI als onderzoeker: generatieve AI draagt bij aan het bedenken, controleren en uitwerken van ideeën voor onderzoek. Hierdoor ontstaat mogelijk een versnelling in onderzoeksbevindingen in tal van richtingen, bijvoorbeeld in de natuurkunde, wiskunde en biologie. Generatieve AI is heel efficiënt in het vinden van patronen en verbanden en kan tot op zekere hoogte eigen suggesties nakijken en herzien. Dit maakt het uitermate geschikt voor academische disciplines die verder bouwen op voorgaande bronnen.

Generatieve AI als basis voor beeld en geluid: steeds meer online content wordt gegenereerd door generatieve AI-modellen en -toepassingen. Aangezien deze steeds beter worden, is deze content vaak moeilijk van echt te onderscheiden.[33] Een logisch gevolg hiervan is dat mensen steeds minder vertrouwen hebben in de echtheid van digitale informatie. Ook maakt het openbaar aanbieden van modellen het relatief eenvoudig om misbruik te maken van de technologie en inhoud te genereren die misleidend is, zonder dat er direct bewijs is dat het om gegenereerde inhoud gaat. Want alle "nieuwe" inhoud die wordt gegenereerd, reproduceert de voorbeelddata waar het AI-model mee getraind is. Dat wil zeggen dat bestaande vooroordelen die in de voorbeelddata zitten ook terug zullen komen in de inhoud die gegenereerd wordt door de AI. Deze vooroordelen worden daardoor verder bestendigd in de maatschappij.

Generatieve AI als zoekmachine: generatieve AI wordt vaak ingezet als kennisbron. Dit vervangt de meer traditionele zoekmachine, die op basis van zoektermen een lijst teruggeeft van relevante bestaande bronnen. Een belangrijk verschil is dat generatieve AI ook de inhoud van openbare bronnen verwerkt, of heeft verwerkt tijdens de training. Op basis daarvan wordt nieuwe informatie gegenereerd om de zoekopdracht te beantwoorden. Dit maakt het voor de gebruiker makkelijk om informatie te krijgen, maar in dat gemak schuilt ook een risico. Er is geen garantie dat die gegenereerde informatie correct is, en door de simpele presentatie is dit voor mensen minder intuïtief in te schatten of te achterhalen. Mede om deze reden zien we recent een ontwikkeling waarbij generatieve AI-zoekmachines meer expliciet verwijzen naar bestaande bronnen. Tegelijkertijd bewegen de traditionele zoekmachines ook richting generatieve AI, waarbij bovenaan de resultaten een gegenereerde samenvatting verschijnt van gevonden informatie.

31 AP en ACM: chatbot mag mens niet volledig vervangen bij klantenservice | Autoriteit Persoonsgegevens

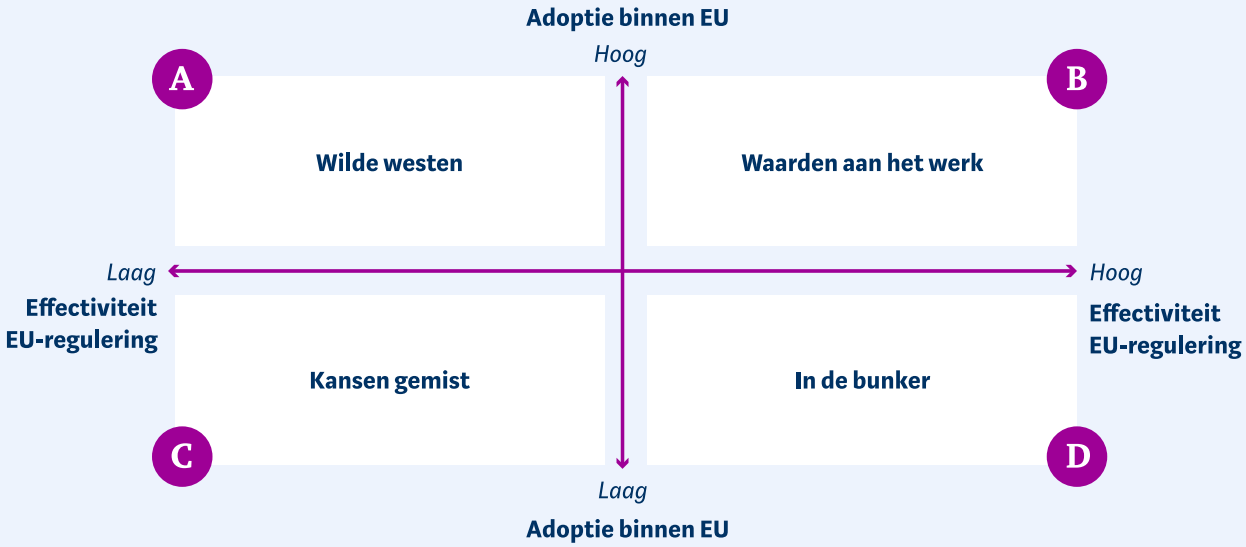
32 TNO (Januari 2025). "Technologieradar gezond en veilig werken".

33 Miller, E. J., Steward, B. A., Witkower, Z., Sutherland, C. A., Krumhuber, E. G., & Dawel, A. (2023). AI hyperrealism: Why AI faces are perceived as more real than human ones. Psychological Science, 34(12), 1390-1403.

5. Toekomstscenario's: generatieve AI in 2030

Met een technologie die zich zo snel ontwikkelt, is het moeilijk om te voorspellen hoe de toekomst eruit zal zien, zelfs op relatief korte termijn. Dit hoofdstuk verkent 4 toekomstscenario's voor de wijze waarop generatieve AI in 2030 ingebed kan zijn in onze samenleving. De scenario's zijn opgesteld langs 2 assen van kernonzekerheden op het gebied van regulering en adoptie van de technologie.

FIGUUR 5 | Toekomstscenario's Generatieve AI



De eerste kernonzekerheid gaat over de mate van adoptie van generatieve AI binnen de EU. In hoeverre speelt generatieve AI een rol in het dagelijks leven van Europese burgers en is er vertrouwen in (verantwoorde) oplossingen? In hoeverre wordt generatieve AI wijdverspreid ingezet in verschillende domeinen als onderwijs, zorg en in het bedrijfsleven? Worden verantwoorde initiatieven gestimuleerd om de risico's die gepaard gaan met deze technologie in te perken?

De tweede kernonzekerheid betreft de effectiviteit van regulering op de Europese markt. De komende jaren moet blijken of de recent aangenomen Europese wetten om (generatieve) AI in goede banen te leiden in de praktijk doeltreffend zijn, ook in combinatie met bestaande wetgeving. Zijn regels achterhaald, vaag en complex? Of zijn er guidance, standaarden, expertise en middelen voor effectieve handhaving en toepassing van wetgeving? Is er goede samenwerking tussen lidstaten en stimuleert regelgeving ontwikkeling en gebruik van verantwoorde oplossingen?

Door de 2 assen te combineren komen we tot 4 scenario's die ieder een ander toekomstbeeld voorstellen. Hieronder schetsen we op hoofdlijnen het toekomstbeeld per scenario met het jaar **2030** als horizon:

A

Scenario A:

Wilde westen

Lage effectiviteit EU-regulering, hoge adoptie generatieve AI.

Er zijn zowel verantwoorde als onverantwoorde generatieve AI-oplossingen op de markt. Organisaties trekken zich weinig aan van regels: regelgeving is niet realistisch. Veel ontwikkelaars en organisaties die generatieve AI gebruiken hebben geen oog voor grondrechtenrisico's. Onverantwoorde oplossingen hebben de overhand omdat die technologie sneller op gang komt dan regelgeving en handhaving kan bijbenen. Bovendien hebben toezichthouders onvoldoende middelen om adequaat te handhaven. Tegelijkertijd zien organisaties en burgers de meerwaarde van generatieve AI, maar is de AI- en algoritmische geletterdheid laag. Er zijn veel incidenten: burgers kunnen niet adequaat beschermd worden tegen risico's voor grondrechten. Persoonsgegevens belanden op straat, er wordt veel gebruik gemaakt van manipulatieve AI en mensen worden steeds meer afhankelijk: menselijk contact neemt af, technische afhankelijkheid neemt toe. Dit uit zich in actieve beïnvloeding van Europese burgers met grote invloed op verkiezingen. Politieke partijen proberen relevant te blijven met AI-gegenereerde deepfakes te publiceren. Medisch advies komt niet meer van de dokter maar van een chatbot. Klantenservices zijn volledig uitbesteed aan chatbots en alleen online toegankelijk. Veel van de overgebleven kantoorbanen richten zich op het bijsturen van AI-systemen naar gewenste resultaten. De verspreiding van desinformatie is wijdverbreid. Het lobbyen van grote techbedrijven heeft vruchten afgeworpen: kleine verantwoorde ideeën maken geen kans tegen de commerciële techbedrijven. Start-ups die zich wel aan de regels houden en verantwoorde oplossingen ontwikkelen, verliezen hun concurrentiepositie. Het vertrouwen in de overheid is tot een dieptepunt gedaald.

B

Scenario B:

Waarden aan het werk

Hoge effectiviteit EU-regulering, hoge adoptie generatieve AI.

Bedrijven, organisaties en consumenten maken volop gebruik van waardengedreven generatieve AI-toepassingen. De Europese regelgeving die bij de introductie enkele jaren geleden nog vele wenkbrauwen deed fronsen, heeft vruchten afgeworpen: mede door forse Europese investeringen is er een bloeiend AI-ecosysteem ontstaan. Kunstmatige intelligentie is een vast onderdeel van het curriculum van Europese lagere en middelbare scholen. Ook heeft interdisciplinair wetenschappelijk onderzoek naar AI een *boost* gekregen, waardoor het aantal patentaanvragen op AI-gedreven innovaties de afgelopen jaren sterk is gestegen. Generatieve AI-applicaties met persoonsgegevens kunnen AVG-*compliant* gebruikt worden, omdat deze gegevens weer verwijderd kunnen worden uit modellen. Toezichthouders spelen in dit alles een belangrijke faciliterende rol: onder meer door intensieve samenwerking op nationaal en Europees niveau is er een internationaal concurrerend *level playing field* gecreëerd, waarbij veel aandacht is voor harmonisatie en constante uitwisseling plaatsvindt tussen toezichthouders en de praktijk. Verantwoorde generatieve AI-toepassingen zijn hierdoor de norm – en best beschikbare producten – geworden. Europese AI-ontwikkelaars krijgen steun bij de ontwikkeling van verantwoorde initiatieven bij het voldoen aan regelgeving. Heldere standaarden en frameworks maken ook dat risico's tijdig worden opgespoord en gemitigeerd. Er heerst een sterk lerende cultuur: incidenten worden systematisch benut om systemen en applicaties te verbeteren, waardoor het vertrouwen in generatieve AI in de samenleving groot is en burgers hun rechten weten uit te oefenen. Dit scenario met hoge adoptie heeft verschillende aandachtspunten, waaronder afhankelijkheid van de technologie. Beslissingen in het dagelijks leven draaien volledig op AI-assistenten: werkplannen en medische adviezen vertrouwt men toe aan chatbots. Een kritische houding blijft van belang. Wetgeving zal dusdanig toekomstbestendig moeten zijn dat nieuwe ontwikkelingen in de technologie of de gevolgen tijdig kunnen worden opgevangen. De vaardigheden schrijven, begrijpend lezen en

vormgeven nemen mogelijk af. Er komt een shift in scholing: administratieve banen verdwijnen, terwijl er nieuwe banen opkomen die niemand ziet aankomen. Generatieve AI brengt veel, maar verandert ook fundamenteel de maatschappij.

C

Scenario C:

Kansen gemist

Lage adoptie generatieve AI,
lage effectiviteit EU-regulering.

Er is groeiend wantrouwen naar buitenlandse AI-systemen, en regelgeving slaat alle Europese initiatieven dood. De wet- en regelgeving is complex en onduidelijk; bedrijven weten niet precies wat wel of niet mag. Waar regelgeving wel meer duidelijkheid biedt, is er een 'check the box'-cultuur: regelgeving biedt daar weinig bijdrage aan daadwerkelijke bescherming van grondrechten. Beloftevolle Europese initiatieven stranden in de ontwikkelfase door complexe processen. De AI-bedrijven die er nog zijn, richten zich alleen op de meest basale generatieve AI-systemen. Buitenlandse alternatieven leiden tot impactvolle incidenten waarbij grondrechten in het geding komen. Het lukt toezichthouders niet om burgers hiertegen te beschermen: het toezicht is versnipperd en de kenniskloof tussen beleidsmakers, toezichthouders en AI-ontwikkelaars is enorm. Bij gebrek aan succesvolle Europese initiatieven en het vertrouwen in de technologie en de bescherming van grondrechten ten opzichte daarvan, is het gebruik van generatieve AI laag. Doordat er weinig gebruik en ontwikkeling van generatieve AI in de EU is, is de kennisopbouw ook zeer beperkt. De EU is niet langer relevant bij de verdere ontwikkeling van features en beheersmaatregelen. De kennisachterstand in de EU versterkt het wantrouwen naar systemen van buiten de EU. Ongelijkheid ten opzichte van andere werelddelen neemt toe, waar generatieve AI wel breed wordt omarmd in de maatschappij. Start-ups vertrekken naar buiten de EU.

D

Scenario D:

In de bunker

Lage adoptie generatieve AI,
hoge effectiviteit EU-regulering.

Door allerlei factoren, waaronder strenge regels, komt het Europese ecosysteem van generatieve AI niet op gang. De regulering van generatieve AI is effectief in de zin dat grondrechten en publieke waarden goed beschermd zijn: er zijn heldere standaarden, burgers worden op een duidelijke manier geïnformeerd over de risico's. Bij de inzet van generatieve AI is er, mede door robuuste samenwerking met toezichthouders, goed zicht op risico's. De gevolgen van incidenten blijven voor burgers relatief beperkt. Tegelijkertijd is er weinig contact tussen toezicht-houders en marktpartijen: wet- en regelgeving is steeds meer losgezongen van de wereldwijde technologische en economische ontwikkelingen. De bestaande toepassingen zijn veilig, maar het zijn er weinig. De strikte regelgeving zorgt ervoor dat maar weinig producten op de Europese markt kunnen worden aangeboden. Gecombineerd met de extra investeringskosten die nodig zijn om aan alle vereisten te voldoen, richten grote marktpartijen zich in eerste instantie op andere markten voor ontwikkeling en innovatie. Bedrijven en organisaties zijn daarbij terughoudend met het inzetten van generatieve AI, uit angst voor boetes bij incidenten. Deepfakes, misinformatie en illegale markten worden snel aangepakt, maar er worden ook kansen gemist. Sectorale markten, waar Europa oorspronkelijk marktleider was, verliezen hun internationale positie doordat hun innovatie-mogelijkheden beperkt worden. Het gat tussen de EU en andere landen groeit, tot frustratie van ondernemers en burgers die zien wat er in het buitenland wél mogelijk is. De belofte van digitale strategische autonomie blijkt moeilijk te worden ingelost. Dit heeft niet alleen economische consequenties voor de positie van Europa op het wereldtoneel, maar dit maakt ook dat onze verworven vrijheden en grondrechten steeds meer onder druk komen te staan. Er ontstaat een Europese kennisachterstand in het onderwijs en in de zorgsector. Ook missen burgers de kansen die andere landen genieten, te danken aan productiviteitsvoordelen. De kloof met landen waar de kansen van generatieve AI worden benut, groeit.

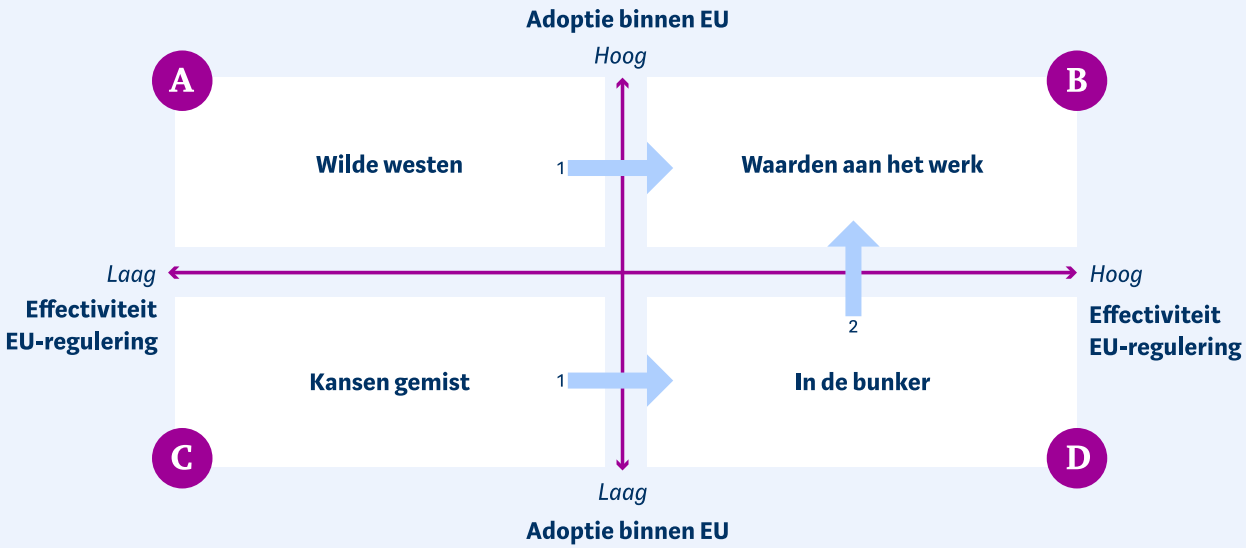
Handelingsperspectief AP

Van de omschreven toekomstscenario's ziet de AP-scenario B 'Waarden aan het werk' als het meest gewenste scenario. Hierin is regulering effectief en zorgen we samen dat waardengedreven innovatie tot stand komt. De geschetste scenario's zijn *mogelijke* toekomstscenario's die beïnvloed worden door veel externe factoren. Sommige ontwikkelingen kunnen we maar beperkt beïnvloeden, op andere ontwikkelingen hebben we meer directe invloed. Hieronder schetsen we per as wat de AP – vanuit eigen handelingsperspectief en invloedssfeer – aan acties wil ondernemen om richting scenario B te bewegen.

1. Richting effectieve regulering
Om de effectiviteit van regelgeving te verhogen, kan de AP guidance bieden, handhaving aanscherpen en invloed uitoefenen op nationale en Europese politiek en verscherpen van Europese en nationale samenwerking. Zo kan de AP een rol spelen in de verduidelijking van onduidelijke wetgeving door middel van normuitleg en verduidelijking van wet- en regelgeving. De AP kan daarnaast de handhaving versterken door middel van onderzoeken, waarschuwingen, aanbevelingen en het opleggen van sancties. Dit vergt opbouw van middelen, expertise en samenwerking. Daarnaast kan een toezichthouder bijvoorbeeld ook beleidsmakers adviseren over nieuwe regelgeving.

2. Richting adoptie van verantwoorde Europese initiatieven
Wanneer de adoptie van generatieve AI laag is, maar Europese regelgeving effectief, kan de AP een bijdrage leveren om verantwoorde Europese initiatieven te ondersteunen. Dit kan bijvoorbeeld door het opzetten van een *sandbox*-omgeving waar organisaties vragen kunnen stellen over de vertaling van regelgeving naar de praktijk,

FIGUUR 6 |Richting het gewenste toekomstperspectief



zoals de *regulatory sandbox*-omgeving voor de AI-verordening.[34] Toezichthouders kunnen bijdragen aan pilots met publieke en private instanties om verantwoorde oplossingen te stimuleren binnen veilige kaders.

Daarnaast kan de AP regionale initiatieven voor datacenters ondersteunen. De AP kan bijdragen aan kennisoverdracht over de kansen en risico's van AI om de algoritmische en AI-geletterdheid te bevorderen.

[34] De AI-verordening verplicht lidstaten tot het oprichten van ten minste één AI regulatory sandbox vanaf augustus 2026. Binnen een regulatory sandbox bieden toezichthouders ondersteuning aan aanbieders van AI-systemen die tijdens de ontwikkeling vragen hebben over hoe zij aan de AI-verordening kunnen voldoen, zodat zij hun product conform de regels op de markt kunnen brengen.

6. Juridisch kader en beheersmaatregelen

Dit juridisch kader bespreekt de rol van grondrechten en relevante Europese wet- en regelgeving voor de verantwoorde ontwikkeling en inzet van generatieve AI. Als toezichthouder op grondrechten is dit juridische kader voor de AP een waardevol vertrekpunt voor verantwoorde generatieve AI. Wet- en regelgeving en juridische instrumenten moeten ervoor zorgen dat ontwikkelaars en gebruikers van deze technologie maatregelen nemen om risico's voor grondrechten te beperken. Generatieve AI is een technologische ontwikkeling die dwars door alle sectoren heengaat en daarom niet slechts onder één wet te scharen valt. Verschillende wettelijke kaders zijn van toepassing op de ontwikkeling en het gebruik van generatieve AI-modellen en toepassingen. Generatieve AI valt dus niet in een juridisch vacuüm, maar dient zich te conformeren aan zowel technologie-specifieke regelgeving als domeinspecifieke wetgeving. Het beheersingskader verantwoorde generatieve AI is een schematische, niet-limitatieve weergave van beheersmaatregelen uit Europese wetgeving. Dit helpt ontwikkelaars en gebruikers in verschillende fases van de technologie grondrechten zo goed mogelijk te beschermen.

Grondrechten van burgers

De ontwikkeling, de inzet en het gebruik van generatieve AI heeft invloed op de grondrechten van burgers.[35] Grondrechten gaan enerzijds uit van een verticale werking. Dat wil zeggen dat zij burgers beschermen tegen overheidsoptreden. Anderzijds hebben grondrechten ook een horizontale werking. Dat houdt in dat burgers een beroep kunnen doen op bepaalde grondrechten ten opzichte van andere burgers (en bedrijven). Generatieve AI kan een bijdrage leveren aan de bevordering van grondrechten, maar kan ook risico's met zich meebrengen voor de bescherming van grondrechten.

Generatieve AI-modellen en -toepassingen kunnen een inbreuk maken op grondrechten van burgers als ontwikkelaars en gebruikers deze grondrechten onvoldoende in acht nemen. Denk bijvoorbeeld aan het trainen van modellen aan de hand van onvolledige datasets waarin bepaalde groepen ondervertegenwoordigd zijn waardoor het model hun kenmerken, taal of context niet goed leert. Dit kan leiden tot een inbreuk op het recht op non-discriminatie. Ook kan het ontwikkelen en gebruiken van generatieve AI-modellen en -toepassingen invloed hebben op het recht op bescherming van het intellectuele eigendomsrecht, bijvoorbeeld bij het genereren van afbeeldingen of video's via generatieve AI. Er zijn ook chatbotapps die zich specifiek richten op de mentale gezondheid van gebruikers en claimen deze te verbeteren. Verkeerde inzet van chatbots kan door afhankelijk en onbetrouwbaarheid serieuze impact hebben op mensen die op zoek zijn naar hulp bij mentale problemen.

Generatieve AI kan bepaalde grondrechten juist ook versterken. Als generatieve AI-toepassingen baanbrekende veranderingen teweegbrengen binnen de gezondheidszorg of het onderwijs, kan dat mogelijk leiden tot een betere toegankelijkheid voor burgers binnen die sectoren. Op die manier zorgt generatieve AI indirect voor een betere inkleuring van het recht op onderwijs en het recht op gezondheidszorg.

[35] Handvest van de Grondrechten van de Europese Unie.

Europese regelgeving

Afhankelijk van de sector waarin het generatieve AI-model wordt ontwikkeld of ingezet kan verschillende Europese wet- en regelgeving van toepassing zijn. Denk bijvoorbeeld aan financiële wetgeving, consumentenrecht, mediawetgeving, wetgeving voor de gezondheidszorg, bestuursrecht en auteursrechtenbescherming.

Naast deze bestaande wet- en regelgeving is er recentelijk veel nieuwe Europese regelgeving ontwikkeld om de inzet en het gebruik van nieuwe technologieën reguleren. Dat betreft in de eerste plaats de inwerkingtreding van de Algemene verordening gegevensbescherming (AVG, 2016) ter bescherming van de persoonsgegevens van burgers. Daarnaast zijn er de afgelopen jaren diverse nieuwe wetten aangenomen om onder andere:

- Digitale platformen te reguleren (Digital Services Act, DSA[36]).
- Marktdominantie door Big Tech bedrijven tegen te gaan (Digital Markets Act, DMA[37]).
- Cybersecurity te waarborgen (Network and Information Security directive, NIS2[38] en Cyber Resilience Act, CRA[39]).
- Eerlijke toegang tot data en de bescherming daarvan te waarborgen (Data Act, DA[40] en Digital Governance Act, DGA[41]).
- Te zorgen dat AI op een verantwoorde manier wordt ontwikkeld en ingezet (AI-verordening).

Ook staan er nog verschillende andere wetgevings-initiatieven op de Europese agenda die mogelijk ook invloed gaan hebben op generatieve AI, zoals de Digital Fairness Act.

Voor de verantwoorde ontwikkeling en inzet van (generatieve) AI is de in 2024 aangenomen AI-verordening relevant. Deze risicogebaseerde wetgeving omvat regels voor aanbieders en gebruikers van AI-systemen om betrouwbare AI in Europa te bevorderen. AI-systemen die een onaanvaardbaar risico met zich mee brengen, zijn verboden. Daarnaast gelden er regels voor hoogrisico-toepassingen en zijn er transparantieverplichtingen voor AI-systemen met bepaalde doeleinden. Voor aanbieders van AI-modellen voor algemene doeleinden (”General Purpose AI” – ”GPAI”) gelden er specifieke vereisten. Deze modellen vormen doorgaans de basis van generatieve AI. De AI-verordening is daarom specifiek relevant voor deze technologie en wordt verder besproken in hoofdstuk 7.

Bij het gebruik van generatieve AI komen vaak persoonsgegevens kijken, en is de AVG van kracht. Door bijvoorbeeld beeldmateriaal te gebruiken zijn personen al snel te identificeren. Wanneer het gaat om bijvoorbeeld medisch advies kan er ook sprake zijn van bijzondere persoonsgegevens. De AP ziet het aantal datalekken gerelateerd aan generatieve AI toepassingen stijgen, en waarschuwt voor risico’s gerelateerd aan de invoer van persoonsgegevens.

Daarnaast bevat de data waarop generatieve AI-modellen worden getraind in veel gevallen persoonsgegevens. Hierdoor is de AVG van toepassing vanaf de eerste stap in de generatieve AI-keten: de dataverzameling. Ook in alle daaropvolgende stappen blijft de AVG relevant voor deze persoonsgegevens, tenzij een model aantoonbaar anoniem wordt bevonden. De AP publiceert in 2026 een handreiking op dit onderwerp.

36 https://www.rijksoverheid.nl/actueel/nieuws/2024/02/16/dsa-extra-verantwoordelijkheden-gelden-nu-voor-alle-digitale-diensten
37 https://ondernemersplein.overheid.nl/product-dienst-en-innovatie/digitaal-ondernemen/dma/
38 https://www.digitaleoverheid.nl/overzicht-van-alle-onderwerpen/cyberbeveiligingswet/
39 https://www.rdi.nl/onderwerpen/handel-en-apparatuur/cra
40 https://www.digitaleoverheid.nl/nieuws/europese-data-act-van-toepassing-geworden/
41 https://www.rijksoverheid.nl/actueel/nieuws/2023/09/28/data-governance-act-van-kracht-om-data-betrouwbaar-en-makkelijk-te-delen

FIGUUR 7 | Beheersingskader verantwoorde generatieve AI

	Modellaag	Systeemlaag	Toepassingslaag
Risicomanagement	• Impact-assessments (bv. DPIA, FRIA) (AVG, AIV) • Security by design/secure development lifecycle (NIS2, ISO 27001) • Neutrale data intermediairs (DGA)	• Impact-assessments (bv. DPIA, FRIA) • Gegevenskwaliteitsbeoordeling (AVG, AIV) • Conformiteitsbeoordeling (AIV)	• Impact-assessments (bv. DPIA, FRIA) • Transparantieverplichtingen (bv. deepfakes) (AIV)
Technische beheersmaatregelen	• Modelkaart en technische documentatie (AIV) • Datakwaliteits- en interoperabiliteitseisen (DA) • Pseudonimisering/anonimisering (AVG) • Fairness & bias detectie	• Fairness & bias detectie • Privacy Enhancing Technologies (bv. differential privacy) (AVG) • Interoperabiliteit (DA) • Technische documentatie (AIV)	• Fairness & bias detectie • Monitoring & logging • Explainability & interpretability
Organisatorische beheersmaatregelen	• Privacy by design, Privacy by default (AVG) • Aanwijzen EU vertegenwoordiger (AIV)	• Informatiebeveiligingsplan • AI-governance • Privacy by design, Privacy by default (AVG)	• Training en bewustwording (AI-geletterdheid) (AIV) • Menselijke tussenkomst (AVG) • Privacy by design, Privacy by default (AVG)
Juridische beheersmaatregelen	• Register van gegevensbemiddelingsdiensten (DGA) • Regulatory sandbox (AIV) • Voorafgaande raadpleging (AVG) • Bescherming van intellectueel eigendomsrecht, bedrijfsgeheimen en vertrouwelijke bedrijfsinformatie (AIV)	• Verwerkingsregister (AVG) • Algoritmische audits • Onafhankelijke audits (DSA)	• Incident reportage (bv. datalek) (AVG) • Post-market monitoring (AIV) • Meldpunten/klachtmechanismen (DSA, AVG, AIV)

7. Generatieve AI en de AI-verordening

De AI-verordening is een risico-gebaseerde wet die de gezondheid, veiligheid en grondrechten van Europese burgers beschermt bij de ontwikkeling en inzet van AI. Deze Europese verordening is in 2024 aangenomen en wordt stapsgewijs ingevoerd. De mate van het risico bepaalt aan welke eisen aanbieders en gebruikers van AI zich moeten houden. De AI-verordening bevat verplichtingen voor aanbieders en gebruikers van AI-systemen, maar ook verplichtingen voor aanbieders van de onderliggende modellen. Dit hoofdstuk gaat in op de regels voor aanbieders van General-Purpose AI-modellen en de Code of Practice die aanbieders moet helpen voldoen aan deze regels. Daarbij wordt benadrukt welke rol deze regels spelen voor aanbieders hoog risico AI-systemen uit de AI-verordening. Deze kaders dragen bij aan de ontwikkeling van verantwoorde AI-systemen en modellen.

De AI-verordening kent regels voor aanbieders en gebruikers van AI, afhankelijk van de mate van risico die de toepassing met zich meebrengt. Deze regels raken ook aan zowel de modellaag als specifieke toepassingen van generatieve AI. Een aantal toepassingen zijn dusdanig risicovol dat ze verboden zijn. Zo is de inzet van een AI-systeem dat leidt tot manipulatie of misleiding met schadelijke gevolgen, verboden. Daarnaast zijn er toepassingen van AI-systemen die een hoog risico met zich meebrengen. Het merendeel van de vereisten uit de AI-verordening richt zich op deze systemen. Dit zijn bijvoorbeeld AI-systemen die worden ingezet in het onderwijs of op de werkplek. Wanneer generatieve AI wordt ingezet in deze context, moeten aanbieders en gebruikers zich aan bepaalde eisen houden. Ook zijn er transparantieverplichtingen. Aanbieders van bepaalde generatieve AI-systemen moeten bijvoorbeeld duidelijk maken dat gebruikers met AI-gegenereerde inhoud te maken hebben.

Daarnaast zijn er ook specifieke regels voor aanbieders van General-Purpose AI (GPAI)-modellen. Deze zijn sinds 2 augustus 2025 al van kracht. Dit gaat om AI-modellen voor algemene doeleinden, welke doorgaans de modellaag vormen van generatieve AI. Voor de aanbieders van GPAI-modellen gelden verplichtingen op het gebied van transparantie, auteursrechten en veiligheids- en beveiligingsmaatregelen (Safety & Security). De toepassing van de regels kan verschillen tussen GPAI-modellen. Voor aanbieders van GPAI-modellen gelden specifieke transparantieverplichtingen, maar deze vereisten zijn voor aanbieders van open source-modellen grotendeels niet van toepassing. Voor aanbieders van GPAI-modellen met een systeemrisico geldt daarnaast een strenger regime bovenop de algemene regels. Een GPAI-model kan een systeemrisico hebben vanwege het bereik en de voorzienbare negatieve gevolgen voor de gezondheid, de veiligheid, de openbare veiligheid, de grondrechten of de samenleving als geheel. Deze GPAI-modellen worden geclassificeerd als een model met systeemrisico wanneer het aan bepaalde voorwaarden voldoet, vooral met betrekking tot de technische capaciteit die gebruikt is voor de training. Deze regels gelden ook voor open source-modellen met een systeemrisico.

GPAI Code of Practice

Met de Code of Practice is een concreet kader uitgewerkt die de betrouwbare ontwikkeling van GPAI-modellen moet stimuleren. Op 10 juli 2025 heeft de AI Office de *'Code of Practice for General Purpose AI-models'* (hierna: de Code of Practice) gepubliceerd. De Code of Practice is vervolgens ook toereikend verklaard door de Europese Commissie (EC) en de lidstaten.[42] De Code of Practice dient als vrijwillige compliance tool om aanbieders van GPAI-modellen te helpen voldoen aan de vereisten uit de AI-verordening. De Nederlandse toezichthouders – met coördinatie vanuit de AP en de RDI – hebben samen met de ministeries via de AI Board input geleverd op verschillende conceptversies van de Code of Practice.

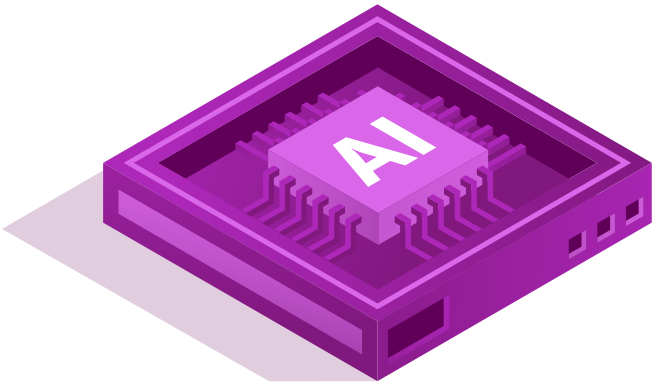
[42] https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai

De Code of Practice bestaat uit 3 delen. De eerste 2 onderdelen bevatten de algemene transparantie- en auteursrechtelijke verplichtingen. Aanbieders van GPAI-modellen moeten met een gestandaardiseerd formulier technische documentatie opstellen, zoals de trainings-gegevensbronnen van het model, licentie-informatie en de *computing power* die voor de training is gebruikt. Verder moeten aanbieders een voldoende gedetailleerde samenvatting opstellen van de content die gebruikt is bij de training, en deze openbaar maken. Het derde onderdeel van de praktijkcode geldt enkel voor aanbieders van GPAI-modellen met een systeemrisico.[43] Aanbieders van modellen met een systeemrisico moeten een uitgebreid kader voor risicobeheer implementeren om systeemrisico's te identificeren, te beheersen en te melden bij de AI Office. Dit risicomanagementproces vereist dat aanbieders het model alleen op de markt brengen als de risico's en de mitigatiemaatregelen om deze te beperken, zijn beoordeeld als acceptabel en aanvaardbaar.

Hoewel de regels voor GPAI-modellen zich in eerste instantie op de aanbieder richten, zijn deze ook van belang voor ontwikkelaars en gebruikers die gebruik-maken van deze modellen voor hun eigen systemen. GPAI-modellen vormen namelijk vaak de basis voor het (door)ontwikkelen van verschillende toepassingen van AI-systemen. De eisen uit de Code of Practice zijn daarom ook van belang voor downstreamaanbieders die toepassingen ontwikkelen waarbij deze modellen als basis dienen. Zo moeten downstreamaanbieders informatie kunnen opvragen bij aanbieders van GPAI-modellen. Bovendien moet een downstream-partij die het model finetunet of wijzigt — voor dat deel — ook voldoen aan de transparantie- en auteursrechtvereisten van de Code of Practice.

Het moet in de praktijk duidelijk worden of de transparantievereisten van de Code of Practice voor modelaanbieders specifiek genoeg zijn voor downstreamaanbieders van hoogrisico systemen. Wanneer een AI-*toepassing* als hoog risico is geclassificeerd in de AI-verordening, gelden er voor aanbieders en gebruikers van deze systemen specifieke verplichtingen. Net als voor aanbieders van GPAI-*modellen*, moeten aanbieders en gebruikers van hoogrisico-toepassingen voldoen aan eisen voor transparantie over o.a. datatraining, (bias)testen en validatie. Wanneer een organisatie een GPAI-model gebruikt in een hoogrisico-toepassing zal deze gebruiker veel kennis nodig hebben over het model, en dus mede afhankelijk zijn van de modelinformatie die de GPAI-modelaanbieder verstrekt. De regels voor hoog risico-toepassingen stellen gedetailleerde eisen aan bijvoorbeeld bias toetsing en loggingsvereisten.

Het moet in de praktijk duidelijk worden of deze transparantievereisten toereikend zijn voor de downstream hoog risico-AI-systeem aanbieders om te voldoen aan de hoogrisico verplichtingen. Als concreet voorbeeld: de Code of Practice schrijft aanbieders van GPAI-modellen voor om te informeren over het type data dat voor training is gebruikt, de dataherkomst en een algemene beschrijving van de datacuratie. Dit is waarschijnlijk in veel gevallen onvoldoende inhoudelijk voor een gebruiker van dat model om te controleren of aan de hoogrisico vereisten op het gebied van representativiteit, datakwaliteit en bias wordt voldaan voor een specifieke toepassing.



[43] AI-modellen met een systeemrisico zijn modellen die beschikken over capaciteiten met grote impact (getraind met een cumulatieve rekenimpact van 10^25 FLOPS of meer).

Het toezicht op GPAI-modellen ligt in beginsel bij de AI Office.[44] De AI Office heeft al aangegeven dat in het eerste jaar een overgangsperiode (*grace period*) geldt voor ondertekenaars van de Code of Practice die te goeder trouw handelen.[45] Voor niet-ondertekenaars, geldt dit voordeel niet. Als aanvulling op de Code of Practice heeft de AI Office op 18 juli richtsnoeren gepubliceerd over de reikwijdte van verplichtingen van GPAI-aanbieders[46], en op 24 juli een *template* voor GPAI-aanbieders om training-*content* voor deze modellen samen te vatten. De toepassing van deze grote AI-modellen – en de naleving van de regels daarvoor – werkt echter door in AI-systemen die onder nationaal toezicht vallen. In Nederland moet het nationaal toezicht nog worden uitgewerkt in implementatiewetgeving.

Op Europees niveau vinden momenteel ontwikkelingen plaats om de digitale wetgeving te vereenvoudigen, waaronder de AVG en AI-verordening. De Europese Commissie heeft hiervoor een voorstel gedaan op 19 november 2025 (*Digital Omnibus Directive*). Deze vereenvoudiging is gepresenteerd als een manier om de naleving te vereenvoudigen en de bureaucratie voor kleine en middelgrote ondernemingen te verminderen. Op deze manier hoopt Europa haar concurrentiepositie in de wereldwijde AI-competitie te versterken. Het voorstel omvat ook veranderingen voor de vereisten voor en het toezicht op GPAI-modellen, generatieve AI-systemen, en de verwerking van persoonsgegevens bij het trainen van AI. Het voorstel moet nu ter goedkeuring worden voorgelegd aan de EU-lidstaten en het Europees Parlement, voordat (een deel van) de veranderingen daadwerkelijk worden ingevoerd.

[44] De AI Office is het Europese toezichtsorgaan op de AI-verordening dat met name toezicht houdt op de naleving van de vereisten voor aanbieders van GPAI-modellen.

[45] https://digital-strategy.ec.europa.eu/en/faqs/questions-and-answers-code-practice-general-purpose-ai

[46] https://digital-strategy.ec.europa.eu/en/policies/guidelines-gpai-providers

8. Gewenst toekomstbeeld

De AP spant zich in voor een toekomst waarin de samenleving de vruchten kan plukken van verantwoorde vormen van generatieve AI op basis van effectieve regulering en vertrouwen in de technologie. In hoeverre de adoptie van generatieve AI hoog of laag zal zijn, is een onzekerheid waar de AP minder directe invloed op heeft, en zal moeten blijken. Op basis van de kennis van nu schetst de AP een gewenst toekomstbeeld; toekomstscenario B ("Waarden aan het werk") dient hierbij als inspiratie. In het gewenste toekomstbeeld zijn fundamentele waarden en grondrechten beschermd en staat de technologie in dienst van de samenleving en het maatschappelijke belang.

Dit toekomstbeeld kenmerkt zich enerzijds door de volgende overkoepelende criteria voor het ecosysteem van generatieve AI (maatschappijbrede kenmerken): Europese digitale autonomie, maatschappelijke weerbaarheid, democratische sturing, een functionerende markt voor verantwoorde oplossingen en het vermogen om door de AI-keten heen te corrigeren. Anderzijds voldoen individuele AI-toepassingen in dit toekomstbeeld aan de volgende kenmerken (op toepassingsniveau): inzichtelijk en transparant, de risico's in kaart en afgewogen, een heldere doelomschrijving, systemen en data in gecontroleerd beheer, rechtmatigheid.

FIGUUR 8 | Waarden aan het werk: uitgangspunt voor verantwoorde inzet

Kenmerken van een situatie waarin fundamentele waarden en grondrechten zijn beschermd en de technologie voor generatieve AI ten dienste staat van de samenleving en het maatschappelijke belang.

Maatschappijbrede kenmerken

Waardengedreven generatieve AI is een motor van innovatie door...

- Europese digitale autonomie
- Maatschappelijke weerbaarheid
- Democratische sturing
- Goed functionerende markt voor verantwoorde oplossingen
- Vermogen om te corrigeren door de AI-keten heen

Toepassingskenmerken

Verantwoorde inzet van generatieve AI is mogelijk door...

- Inzichtelijkheid en transparantie
- Risico's in kaart, afgewogen en gemitigeerd
- Heldere doelomschrijving
- Systemen en data in gecontroleerd beheer
- Rechtmatigheid

In de voorgaande hoofdstukken is vanuit een helikopterview gekeken naar generatieve AI. We bespraken wat de technologie is (hoofdstuk 2), hoe deze wordt gebruikt (hoofdstuk 3) en welke trends op dit moment spelen (hoofdstuk 4). Daarnaast zijn op basis van scenario's mogelijke toekomstbeelden geschetst (hoofdstuk 5).

We beschrijven de vormgeving van het gewenste toekomstbeeld aan de hand van kenmerken. Deze kenmerken maken sturing mogelijk. Want, door deze te realiseren, creëren we de maatschappelijke basis om verantwoord vooruit te kunnen met generatieve AI en het welvaarts- en welzijnsverhogende potentieel te benutten. We kijken daarbij eerst naar maatschappijbrede kenmerken en vervolgens naar kenmerken op toepassingsniveau. De visie sluit af met acties die de AP neemt om aan de gewenste koers bij te dragen.

Maatschappijbrede kenmerken

In het toekomstbeeld dat de AP wenselijk acht, vormt waardengedreven generatieve AI de motor voor innovatie. Een aantal maatschappijbrede kenmerken draagt hieraan bij:

Maatschappijbreed kenmerk 1:

Europese digitale autonomie

Het nastreven van dit kenmerk zorgt ervoor dat de EU in staat is te sturen en in te grijpen op de manier waarop de technologie de samenleving raakt. Hierbij wordt meegenomen dat een grote afhankelijkheid van een organisatie de mogelijkheid op correctie moeilijker kan maken. Binnen de EU worden activiteiten getoetst aan Europese regels, onafhankelijk van druk op de regulering van buiten de unie. In het Draghi-rapport wordt een mate van autonomie en concurrentievermogen als fundament beschreven om grondrechten te kunnen blijven beschermen in de toekomst.[47] Ook het DOSA-rapport van Economische Zaken wees in 2023 al op het belang van strategische onafhankelijkheid van de EU van (ondersteunende) technologie van AI. Afhankelijkheid kan zorgen voor de belemmering van de ontwikkeling van verantwoorde Europese initiatieven die rekening houden met Europese normen en waarden.[48]

Mogelijke instrument: stimuleren van opkomst Europese aanbieders generatieve AI.

Maatschappijbreed kenmerk 2:

Kennis en weerbaarheid

Het nastreven van dit kenmerk zorgt ervoor dat mensen in de gehele samenleving weten in de basis hoe generatieve AI werkt en hebben daardoor realistische verwachtingen van de technologie. Generatieve AI wordt gezien als ondersteunend middel, in plaats van de hoofdverdienste in het vervullen van taken. Ook kunnen mensen kritisch kijken naar de ontwikkeling en inzet ervan.[49] Dit stelt individuen in staat om gebruiksrisico's in te schatten, maar ook te doorzien hoe aspecten in het dagelijks leven beïnvloed worden door generatieve AI, zoals onze informatievoorziening. Afhankelijk van de impact worden modellen en systemen in meer of mindere mate gecontroleerd door gespecialiseerde instituten, onderzoekers, auditors, mensenrechtenorganisaties en diverse geïnteresseerden.

Mogelijke instrumenten: werken aan hoge AI-geletterdheid door maatschappij; periodieke externe controle van generatieve AI-systemen.

47 Europese Commissie (September 2024). "The Draghi report: In-depth analysis and recommendations".

48 Ministerie van Economische Zaken (Oktober 2023). "Agenda Digitale Open Strategische Autonomie"

49 AP (Januari 2025). "Aan de slag met AI-geletterdheid",

Maatschappijbreed kenmerk 3:

Democratische sturing

Het nastreven van dit kenmerk waarborgt dat via het democratisch proces de wensen en zorgen van het brede publiek effectief opgenomen worden in beleid.[50] Brede maatschappelijke discussie onderbouwt hierbij de richting, bijvoorbeeld op de vraag of open-weight-modellen de voorkeur hebben boven closed-weight-modellen. Publieksbijdragen kunnen ook bestaan uit testen of benchmarks voor generatieve AI om vanuit diverse perspectieven bij te dragen aan beheerste impact. Politieke verantwoordelijken houden zich bezig met de wenselijkheid van de toepassing van algoritmes en AI in verschillende contexten.

Mogelijke instrumenten: werken aan adequate expertise over AI-systemen binnen volks-vertegenwoordiging, laagdrempelige platforms voor specialistische benchmarks, aanjagen van maatschappelijke dialoog over generatieve AI.

Maatschappijbreed kenmerk 4:

Een goed functionerende markt voor verantwoorde oplossingen

Er is geen afhankelijkheid van een of enkele aanbieders. In plaats daarvan bestaat er een Europese markt met meerdere opties door de hele keten, van het trainen tot het toepassen van generatieve AI. De regulering creëert een *level playing field* waarin verantwoorde oplossingen niet afdoen aan kwaliteit. Daarnaast is er sprake van *interoperabiliteit* tussen toepassingen: het is voor gebruikers eenvoudig om tussen vergelijkbare toepassingen te wisselen. Dit betekent ook dat de gebruikersdata of het opgebouwde profiel van gebruikers op zo'n manier beschikbaar is dat een 'lock-in'-positie voorkomen wordt.

Mogelijke instrumenten: het nastreven van een gevarieerd aanbod van beschikbare modellen; stimuleren van interoperabiliteit tussen systemen; aanjagen van vergelijkingsplatforms.

Maatschappijbreed kenmerk 5:

Het vermogen om te corrigeren door AI-keten heen

Het nastreven van dit kenmerk zorgt ervoor dat de inbedding van generatieve AI-modellen in digitale systemen zo is opgezet dat kwetsbaarheden of onrechtmatigheden achteraf gecorrigeerd kunnen worden. Als een model persoonsgegevens bevat of ongewenste uitkomsten produceert, heeft iedere partij in de keten een systeem in werking gesteld om die persoonsgegevens of ongewenste uitkomsten te corrigeren.

Mogelijke instrumenten: correctiemethoden zoals machine unlearning; registratie van modelversies binnen hoogrisico-toepassingen.

[50] AP (Maart 2025). "Position paper AP democratische beheersing algoritmes en AI".

Kenmerken op toepassingsniveau

Er zijn verschillende kenmerken van een verantwoorde inzet van generatieve AI op toepassingsniveau. Waar de vorige kenmerken maatschappijbreed relevant zijn, kijken we nu naar gewenste kenmerken van individuele toepassingen. Dit gaat dus over hoe, in het na te streven toekomstbeeld, een organisatie de inzet benadert van individuele systemen die gebruikmaken van generatieve AI. Deze kenmerken komen terug in de aandachtspunten in figuur 1.

Toepassingskenmerk 1:

Inzichtelijk en transparant

Het nastreven van dit kenmerk waarborgt dat generatieve AI-toepassingen duidelijk herkenbaar zijn en zich lenen voor verdere analyse waar gewenst, bijvoorbeeld door op aanvraag (uitkomsten van) diverse beoordelingscriteria te delen. Verregaande mate van transparantie vormt de basis voor vertrouwen en stelt de gebruiker in staat zich weerbaar op te stellen en diens rechten uit te oefenen.

Mogelijke instrumenten: het waarborgen dat gebruikers zien dat ze te maken hebben met AI; het beschikbaar stellen van publiekelijke modelkaarten.[51]

Toepassingskenmerk 2:

Risico's in kaart, afgewogen en gemitigeerd

Het nastreven van dit kenmerk waarborgt dat een gedegen risicoafweging voorafgaat aan de inzet van generatieve AI-toepassingen. Hierin worden juridische en ethische afwegingen expliciet gemaakt en gedocumenteerd. De toepassing bevat geen onacceptabele risico's en is transparant over verhoogde risico's. Diversiteit van perspectieven is in deze analyse en afweging essentieel. Om een gedegen afweging te maken van de impact op mensenrechten kan de FRIA uit de AI-verordening (artikel 27) ter inspiratie dienen.[52]

Mogelijke instrumenten: het inrichten van risicomanagement-raamwerken voor AI-modellen; monitoring van de risico's van AI-modellen in de praktijk via steekproeven.

Toepassingskenmerk 3:

Heldere doelomschrijving

Het nastreven van dit kenmerk waarborgt dat het doel en gebruik van generatieve zo precies mogelijk wordt geformuleerd en gedocumenteerd. Dit is van belang, omdat generatieve AI inherent voor vele doelen kan worden ingezet. De toepassing wordt ingericht om het gebruik binnen de geplande kaders te houden. Na ingebruikname wordt dit gemonitord. De werknemers die betrokken zijn bij de inzet van het systeem, zijn goed op de hoogte van de werking van het systeem en de mogelijke risico's en weten hoe ze deze in de praktijk kunnen herkennen.

Toepassingskenmerk 4:

Systemen en data in gecontroleerd beheer

Het nastreven van dit kenmerk waarborgt dat er controle is over de omgeving waar het generatieve AI-systeem draait en de data die verwerkt wordt. Hierbij is de afhankelijkheid van derde partijen beperkt voor belangrijke processen. Daarnaast is het helder en transparant waar de gebruikersdata zich bevindt en voor wie dit beschikbaar is.

Mogelijke instrumenten: gebruik van open-weight modellen op eigen hardware; gebruik van private cloudoplossingen.

Toepassingskenmerk 5:

Rechtmatigheid

Het nastreven van dit kenmerk waarborgt dat wordt voldaan aan de wettelijke vereisten van de relevante wetgeving, waaronder de AVG en de AI-verordening en eventuele sectorale wetgeving. Dit geldt voor het hele proces, van het verzamelen van data tot aan het inzetten van de toepassing. De functionaris gegevensbescherming (FG) is nauw betrokken bij de implementatie. Modellen van derde partijen hebben een acceptabel niveau van anonimiteit, zoals beschreven door de EDPB.[53]

[51] Hugging Face. "Model Cards".

[52] https://artificialintelligenceact.eu/article/27/

[53] European Data Protection Board (December 2024). "Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models".

FIGUUR 9 | Vergelijking risico's van open-weight en closed-weight modellen'

Box 2

Impact van openbare modeltoegang op risico's

Voor de toekomstige ontwikkeling en inzet van generatieve AI is het al dan niet open-source beschikbaar komen van geavanceerde modellen een belangrijke en bepalende factor. Meer specifiek gaat het dan over of de zogenaamde model-parameters beschikbaar komen, oftewel open-weight-modellen. Een open-weight-model is vrij te downloaden op het internet, en met complete vrijheid in te zetten of aan te passen. Er zijn al honderdduizenden open-weight generatieve AI-modellen beschikbaar. Het volgende overzicht benoemt een aantal relevante implicaties voor risico's als gevolg van het wel of niet openbaar maken van de modellen. Dit is een versimpelde weergave. Het veilig zelf draaien van open-weight-modellen vergt expertise en is mogelijk een uitdaging voor kleinere organisaties. Licenties en benchmarks en ondersteunende tools ontwikkelen zich echter snel en kunnen hieraan bijdragen.

Uit deze vergelijking blijkt dat het open-weight beschikbaar maken van modellen andere risico's met zich meebrengt dan closed-weight-modellen. Er komen risico's bij, terwijl andere risico's gemitigeerd worden. Op dit moment zijn beide varianten in omloop. Als maatschappij zullen we dan ook rekening moeten houden met de risico's van beide varianten.

Volgens de AP is dit het moment in de tijd om risico's te beïnvloeden door het openbaar maken van modellen aan te moedigen of juist tegen te gaan. Het bepalen van de gewenste richting hierin vraagt om brede maatschappelijke discussie.

Risicocategorie →	*Open-weight modellen*		*Closed-weight modellen*
Privacy en gegevensbescherming	Eventuele persoonsgegevens in modellen raken wijdverspreid ⚠	↔	Aanbieder kan modellen terugroepen/corrigeren ✓
	Lokaal draaien mogelijk, minder verkeer van gegevens ✓	↔	Veel input en output verkeer naar aanbieder ⚠
Cybersecurity	GenAI cyber verdediging en aanvallen voor iedereen mogelijk	↔	GenAI cyber verdediging en aanvallen op basis van toegang
	Risico op backdoors, malware of verborgen acties in modellen ⚠	↔	Aanbieder heeft controle over de veiligheid van het model ✓
	Hele community kan bijdragen aan modelcontrole en testen ✓	↔	Communicatie tussen systemen introduceert kwetsbaarheden ⚠
Bias, stereotypering & discriminatie	Hele community kan bijdragen aan model alignment ✓	↔	Aanbieder heeft grote invloed op gebruikte normen en waarden ⚠
	Er ontstaan veel modellen met bias zonder aanspreekpunt ⚠	↔	Aanbieder kan aangesproken worden op bias ✓
Gebruik met slechte intenties	Het wordt makkelijker om misinformatie te genereren ⚠	↔	Aanbieder kan toegang ontzeggen bij misbruik ✓
	Krachtige AI kan gebruikt worden veel schade aan te richten ⚠		
Autonomie en marktmacht	Onafhankelijkheid groter, makkelijker toetreden markt ✓	↔	Afhankelijkheid van aanbieder die groeit door gebruik ⚠
	Macht en controle op applicatieniveau ✓	↔	Aanbieders krijgen machtspositie over applicaties ⚠
	Iedereen kan bijdragen aan ontwikkelingen; Meer divers en minder gestuurd	↔	Aanbieders bepalen richting van ontwikkelingen, mogelijk niet gewenste richting maar wel een aanspreekpunt

⚠ Hoog risico
✓ Laag risico

Inzet vanuit de AP

De AP gaat zich komende periode inzetten om een bijdrage te leveren aan een toekomstbeeld waar fundamentele waarden en grondrechten worden beschermd en waar de technologie ten dienste staat van de samenleving en het maatschappelijk belang. Dit doen we enerzijds door extra aandacht te hebben voor generatieve AI binnen onze bestaande werkzaamheden en anderzijds door een aantal nieuwe activiteiten te starten die een positieve bijdrage leveren aan de verantwoorde inzet van generatieve AI in onze samenleving.

Vanuit het reguliere werk levert de AP een bijdrage aan heldere en realistische werkmethoden, onder andere door actief mee te schrijven aan standaarden en opinies. Bijvoorbeeld de EDPB-opinie uit december en de standaarden voor hoogrisico systemen uit de AI-verordening. Daarnaast werkt de AP aan digitale weerbaarheid, onder andere door te investeren in het kennisniveau van de samenleving. Bijvoorbeeld door het bieden van guidance rondom AI-geletterdheid, begrijpelijke informatie op onze website en het organiseren van seminars. Ook risico-signalering vormt een integraal onderdeel van de werkzaamheden van de AP. Bijvoorbeeld via het overzicht aan AVG-risico's voor generatieve AI en de halfjaarlijkse risicorapportages over algoritmes en AI. Als laatste is de AP beschikbaar voor voorafgaande raadpleging en richt het een sandbox op. Hierin kunnen toezichthouders ondersteuning bieden aan aanbieders van AI-systemen die tijdens de ontwikkeling vragen hebben over hoe zij aan de AI-verordening kunnen voldoen.

De AP zet ook een aantal extra stappen om verantwoord vooruit te gaan met generatieve AI. Organisaties kunnen ons bereiken op een loket voor vragen over generatieve AI. We gaan door met het in kaart brengen van welke vragen en uitdagingen er liggen rondom de verantwoorde ontwikkeling en inzet van generatieve AI. Hiervoor organiseren we een aantal bijeenkomsten en start de AP een online loket voor vragen en ideeën over generatieve AI. De informatie die we hiermee ophalen, geeft inzicht in bijvoorbeeld de compliance-vragen en issues die 'top of mind' zijn binnen organisaties. Dit vormt de basis voor verdere periodieke dialoog over de verantwoorde ontwikkeling en inzet van generatieve AI. En het helpt om de focus te leggen op die vraagstukken die het meest urgent zijn.

Ook gaan we aan de slag met concrete instrumenten en handvatten die bijdragen aan het beschermen van fundamentele waarden bij de ontwikkeling en inzet van generatieve AI. Denk bijvoorbeeld aan Europese richtlijnen, het stimuleren van methodes voor het anonimiseren of verwijderen van persoonsgegevens uit modellen, handreikingen voor AI-geletterdheid en uitgangspunten voor transparantie. En samen met andere toezichthouders in het digitale domein zetten we ons daarnaast in voor gezamenlijke normuitleg, zodat we zoveel mogelijk duidelijkheid en daarmee rechtszekerheid creëren voor organisaties die aan de slag willen met generatieve AI. Positieve voorbeelden en *use cases* geven we een podium om daarmee te laten zien en te inspireren hoe verantwoorde generatieve AI er in de praktijk uit kan zien. En natuurlijk blijft de AP toezichthouden op onrechtmatigheden die zich voordoen in het speelveld van generatieve AI.

Zo maken we samen de verantwoorde inzet van generatieve AI mogelijk.

Bijlage 1. Begrippenlijst

Hieronder volgt een korte (niet-juridische) omschrijving van de gehanteerde termen en begrippen:

AI
(artificial intelligence, artificiële intelligentie, kunstmatige intelligentie) Het nabootsen van menselijke vaardigheden met een computersysteem, zoals leren, plannen, redeneren, anticiperen en zelfstandig beslissen zonder tussenkomst van menselijke intelligentie.

AI-model
Een softwarecomponent waarin kennis ligt besloten om op basis van invoer tot uitkomsten (zoals voorspellingen of classificaties) te komen. Modellen leren patronen uit voorbeeld data (voorbeeld invoer met een correcte uitkomst) en verwerken nieuwe invoer op basis hiervan tot een zo goed mogelijk uitkomst.

AI-systeem
Een systeem dat is ontworpen om met verschillende niveaus van autonomie te werken en dat gebruik maakt van een AI-model.

Generatieve AI
Een vorm van AI die in staat is om nieuwe data te genereren. De meest populaire toepassingen maken teksten en afbeeldingen waar niet meer aan af is te zien dat deze door AI gegenereerd zijn. Een invoer van een gebruiker geeft vaak aan wat te genereren.

Machine learning
Methodes om patronen uit data op te nemen in een AI-model.

Neuraal netwerk
Een type AI-model waarin een abstractie van informatie kan worden opgeslagen in de vorm van parameters. Deze parameters zijn getallen die worden gekozen door te trainen op voorbeelddata. Een neuraal netwerk bepaalt op basis van een input en de parameters een output.

Deep learning
Een methode binnen machine learning. Deze methode maakt gebruikt van neurale netwerken met veel parameters.

Open-source model
Een AI-model dat openbaar beschikbaar is. In dat geval is het bijvoorbeeld mogelijk de besliscriteria van een beslisboom, of de parameters van een neuraal netwerk, te downloaden.

Closed-source model
Een AI-model dat alleen beschikbaar wordt gesteld op een manier waarbij je output terugkrijgt op basis van de input. De exacte berekening van de beslissing is dus niet openbaar.

Foundation model
(ook wel: General Purpose model) Een AI-model dat op basis van veel data patronen heeft opgeslagen in grote neurale netwerken. Het model kan hierdoor worden ingezet in veel verschillende toepassingen. Daarnaast kan het model als 'fundament' dienen voor modellen voor specifieke doeleinden, waarbij het foundation model gespecialiseerd wordt.

Large Language Model (LLM)
Een AI-model dat op basis van veel tekstdata patronen heeft opgeslagen in grote neurale netwerken. Het model kan tekst genereren en kan bijvoorbeeld worden toegepast in vraag-antwoord-situaties.

Autoriteit Persoonsgegevens
Postbus 93374
2509 AJ Den Haag

autoriteitpersoonsgegevens.nl