

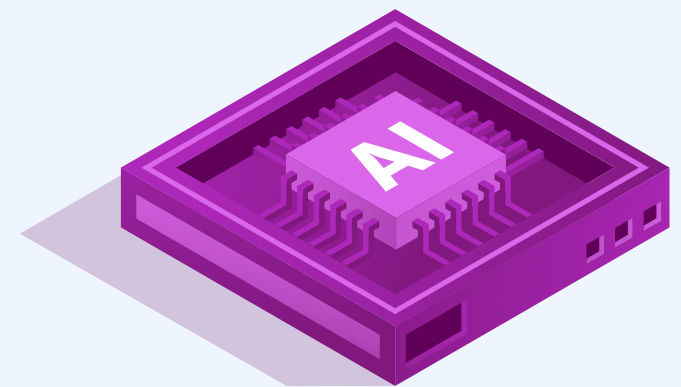
An aerial photograph of a public square paved with grey tiles. Several people are walking across the square. In the bottom right corner, there is a landscaped area with green grass, yellow flowers, and purple flowers. A large purple rectangular box is overlaid on the left side of the image, containing the text 'Verantwoord vooruit' in white.

Verantwoord vooruit

AVG-Randvoorwaarden voor generatieve AI

Inhoud

Consultatie	3
Managementsamenvatting	4
1. Inleiding	5
2. Basisbegrip van generatieve AI	6
Definitie generatieve AI	6
Training van een generatief AI-model	6
Schematische weergave rechtmatigheid	7
AVG-Randvoorwaarden generatieve AI	9
3. Afsluitend: de inzet vanuit de AP	17



The background of the entire page is a repeating pattern of blue, isometric AI chip icons. Each icon is a square with rounded corners, featuring a central square with the letters 'AI' in white, and several small rectangular protrusions along the bottom edge, resembling a microchip. The icons are arranged in a staggered grid across the entire page.

Consultatie

De Autoriteit Persoonsgegevens (AP) nodigt u uit om te reageren op dit document via een e-mail naar **genai-loket@autoriteitpersoonsgegevens.nl**. Dat kan tot en met 27 juni 2025. De reacties die we ontvangen, vatten we samen in één document. We noemen daarbij geen namen, organisaties of contactgegevens. De samenvatting publiceren we op de website van de AP en gebruiken we om de versie te verbeteren. De definitieve visie volgt later in 2025.

Managementsamenvatting

1. **Waar generatieve AI tot op heden tekort schiet in rechtmatigheid, ziet de AP dat het mogelijk is om toe te werken naar een rechtmatige ontwikkeling en inzet van generatieve AI.** De AP heeft enkele onrechtmatigheden en onzekerheden blootgesteld in een juridische analyse. Vervolgens kan de ontwikkeling en inzet van deze technologie worden ingericht om onrechtmatigheden en onzekerheden te voorkomen. We bespreken in dit document de voorwaarden en mogelijkheden in het licht van de AVG.
2. **Het is volgens de AP aannemelijk dat tijdens de ontwikkeling van foundation modellen onrechtmatigheden hebben plaatsgevonden.** De overkoepelende inschatting is dat op basis van de huidige praktijk het overgrote gedeelte van alle generatieve AI-modellen op dit moment tekortschiet qua rechtmatigheid. Generatieve AI-applicaties bouwen voort op zogeheten ‘foundation modellen’. Om deze modellen te trainen, is vrijwel alle publiekstoegankelijke data op het internet gebruikt (gescrapet). Op het internet zijn bijzondere persoonsgegevens geplaatst, die niet door een betrokkene zelf openbaar zijn gemaakt. Het is daarom aannemelijk dat een foundation model onrechtmatig verzamelde bijzondere persoonsgegevens bevat. De AP merkt op dat die bijzondere persoonsgegevens een klein deel van de gehele dataverzameling vormen voor het trainen van foundation modellen. Maar dat neemt niet weg dat deze bijzondere persoonsgegevens onrechtmatig verkregen zijn. Toch is de verdere inzet van deze foundation modellen door Nederlandse en Europese partijen daardoor niet bij voorbaat onrechtmatig, zoals volgt uit een analyse van de EDPB.
3. **Op toepassingsniveau identificeert de AP een aantal randvoorwaarden, uitdagingen en mogelijkheden voor verantwoorde inzet.** Deze toepassingen hebben een helder doel en worden voor dat doel gebruikt na een gedegen risico-afweging met passende waarborgen. De AP ziet verdere mogelijkheden en uitdagingen op het gebied van generatieve AI, die los staan van onderstaande randvoorwaarden voor gegevensbescherming. Bijvoorbeeld het creëren van bewustwording bij gebruikers over het delen van gevoelige data in generatieve AI-toepassingen. Over deze mogelijkheden en uitdagingen weidt het visiestuk ‘Verantwoord vooruit: visie op generatieve AI’ verder uit.
4. **Randvoorwaarden voor verdere verwerking van reeds verzamelde persoonsgegevens zullen later worden uitgewerkt.** Er kunnen zich situaties voordoen waarin een partij eigen, reeds verzamelde persoonsgegevens verder wil verwerken in een generatief AI-model. Doorgaans zal die eigen data op zichzelf niet voldoende zijn voor het trainen van een foundation model. In die situatie wijkt de verdere verwerking af van het oorspronkelijk doel van de verwerking. Er zal dan moeten worden beoordeeld of die verdere verwerking verenigbaar is met het oorspronkelijke doel (onder artikel 6, vierde lid, AVG). Deze analyse valt buiten de reikwijdte van dit stuk en wordt mogelijk in een volgende iteratie van dit document uitgewerkt.

1. Inleiding

In het consulterende visiestuk 'Verantwoord vooruit: visie op generatieve AI' beschrijft de AP wat zij verstaat onder generatieve AI, welke kansen deze technologie de samenleving biedt en welke risico's aan deze technologie kleven. Daarnaast beschrijft dit visiestuk wat er moet gebeuren om generatieve AI verantwoord te omarmen. Daarvoor zullen ontwikkelaars en gebruikers van generatieve AI de verplichtingen uit de AVG moeten naleven bij het verwerken van persoonsgegevens in deze modellen en/of toepassingen. Dit stuk is met name geschikt voor professionals die generatieve AI ontwikkelen of willen inzetten in hun eigen bedrijfsvoering.

Dit document kan op zichzelf staan of als aanvulling op het visiestuk 'Verantwoord vooruit: visie op generatieve AI' gelezen worden. In dit document geven we allereerst een korte definitie van generatieve AI. Daarna zetten we de generatieve AI-keten uiteen en vervolgens bespreken we enkele niet-limitatieve AVG-randvoorwaarden voor generatieve AI. Per randvoorwaarde geven we aan voor welke partijen deze randvoorwaarde van toepassing is en waar mogelijkheden voor verantwoorde ontwikkeling en inzet liggen.

De AP is er zich bewust van dat deze technologie onderhevig is aan snelle veranderingen en momenteel veel in bespreking is in de samenleving. Deze AVG-randvoorwaarden schetsen daarom het beeld van de AP op het moment van publicatie (mei 2025). Het European Data Protection Board (hierna: EDPB) is momenteel bezig met het opstellen van richtlijnen voor generatieve AI, waaraan de AP een actieve bijdrage levert. Hierdoor kunnen onderstaande randvoorwaarden onderhevig zijn aan latere aanpassingen. Ook benadrukt de AP dat er naast onderstaande randvoorwaarden nog andere AVG-verplichtingen van toepassing kunnen zijn op het ontwikkelen of gebruiken van generatieve AI-modellen en -toepassingen.

2. Basisbegrip van generatieve AI

Definitie generatieve AI

Generatieve artificiële intelligentie (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 guidance gaat over AI-modellen die in staat zijn realistische data te genereren en 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.

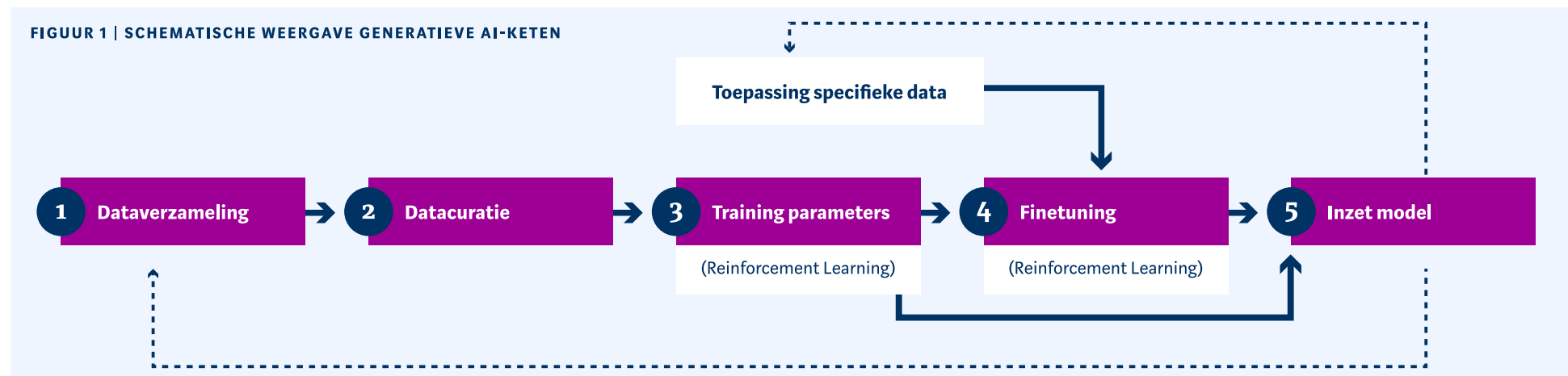
Buiten de reikwijdte van dit stuk, liggen tal van andere vormen van AI. AI-technologie is al lange tijd in ontwikkeling en wordt bijvoorbeeld veel gebruikt voor dataclassificatie. Een dergelijk model kan niet zomaar data genereren, dit is het belangrijkste verschil met generatieve AI.

Training van een generatief AI-model

Versimpeld kan het trainen van een generatief AI-model worden gezien als een set opeenvolgende stappen:

1. **Dataverzameling:** voorbeelddata worden verzameld, vaak door middel van scraping.
2. **Datacuratie:** ongewenste voorbeelden (zoals persoonsgegevens, of haatdragende inhoud) worden verwijderd in een curatiestap.
3. **Training parameters:** de parameters worden getraind op basis van patronen in voorbeelddata en mogelijk met aanscherping door reinforcement learning.
4. **Finetuning:** door finetuning wordt het model afgestemd op een specifieke toepassing of begrenzing.

FIGUUR 1 | SCHEMATISCHE WEERGAVE GENERATIEVE AI-KETEN



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

Gezien de gegeneerde data in veel gevallen weer wordt verzameld voor training, is er sprake van een feedback loop.

Schematische weergave rechtmatigheid

In onderstaand schema geven we verschillende situaties weer voor de ontwikkeling en inzet van generatieve AI-modellen. Het antwoord op de vraag of de ontwikkeling en/of inzet van een generatief AI-model of -applicatie rechtmatig heeft plaatsgevonden, is niet eenduidig. Veel hangt af van de contextuele factoren bij de ontwikkeling of inzet van dergelijke modellen en applicaties, zoals blijkt uit de randvoorwaarden.

In onderstaande situatie ontwikkelt verwerkingsverantwoordelijke A een foundation model. Dit foundation model wordt vervolgens verder gefinetuned en ingezet door verwerkingsverantwoordelijke B.

Verwerkingsverantwoordelijke A heeft in dit scenario een foundation model ontwikkeld dat persoonsgegevens bevat die onrechtmatig zijn verzameld door scraping. De persoonsgegevens zijn vervolgens gecureerd. Die onrechtmatige dataverzameling kan volgen uit het ontbreken van een grondslag onder artikel 6 AVG. Of uit het ontbreken van de mogelijkheid om een beroep te doen op een uitzonderingsgrond onder artikel 9 AVG. Vervolgens stelt verwerkingsverantwoordelijke A dit model beschikbaar aan verwerkingsverantwoordelijke B.

Laatstgenoemde gebruikt het foundation model om verder te finetunen en/of een generatieve AI-applicatie te bouwen. Er zijn dan twee scenario's te onderscheiden:

Scenario 1: Verwerkingsverantwoordelijke A heeft alle persoonsgegevens in het foundation model aantoonbaar geanonimiseerd

Volgens de EDPB en het Hof van Justitie van de Europese Unie (hierna: HvJEU) moet de aantoonbare anonimiteit van een generatief AI-model onder andere worden beoordeeld op de mogelijkheid van identificatie van de betrokkenen.¹ Indien verwerkingsverantwoordelijke B geen persoonsgegevens verwerkt in het geanonimiseerde foundation model tijdens het finetunen of de inzet van zijn generatieve AI-applicatie, is de AVG niet meer van toepassing zoals vastgesteld door de EDPB.² In dat geval zijn er namelijk geen persoonsgegevens meegetraind in het foundation model en verwerkt verwerkingsverantwoordelijke B geen persoonsgegevens.

Wanneer verwerkingsverantwoordelijke B het geanonimiseerde model finetunet met een dataset die persoonsgegevens bevat of bij het gebruik daarvan persoonsgegevens verwerkt, is het finetunen en het gebruik niet bij voorbaat onrechtmatig. De onrechtmatigheid van het ontwikkelen van het foundation model werkt in dat geval niet door naar

1. Zie voor een uitgebreidere uitleg van de anonimiteit van AI modellen: EDPB Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models, hoofdstuk 3.2.

Zie ook HvJEU *Breyer v. Bundesrepublik Deutschland*, ECLI:EU:C:2016:779, par. 42.

2. EDPB Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models, paragraaf 134.

verwerkingsverantwoordelijke B.³ De EDPB heeft in dat kader geoordeeld dat verwerkingsverantwoordelijke B zelf een grondslag nodig heeft voor de verwerking van persoonsgegevens.⁴ Dit laat echter onverlet dat verwerkingsverantwoordelijke A door de bevoegde autoriteit verantwoordelijk kan worden gehouden voor de onrechtmatigheden tijdens de dataverzameling, datacuratie en ontwikkeling van het foundation model.

Scenario 2: Verwerkingsverantwoordelijke A heeft het foundation model niet (aantoonbaar) geanonimiseerd

Wanneer verwerkingsverantwoordelijke A het foundation model niet (aantoonbaar) heeft geanonimiseerd, blijft de AVG van toepassing. De onrechtmatigheid van de verwerking door verwerkingsverantwoordelijke A betekent echter niet dat de uitgevoerde verwerkingen door verwerkingsverantwoordelijke B bij voorbaat ook onrechtmatig zijn. Verwerkingsverantwoordelijke B zal vanuit zijn verantwoordingsplicht een adequate beoordeling moeten uitvoeren voor de ontwikkeling van het model met onrechtmatig verkregen persoonsgegevens door verwerkingsverantwoordelijke A.⁵ De diepgang van die beoordeling hangt af van verschillende factoren. Bijvoorbeeld welke risico's de inzet van het foundation model teweegbrengt voor betrokkenen.⁶ Als uit deze beoordeling volgt dat er bijzondere persoonsgegevens in het foundation model zijn meegetraind, dan kan verwerkingsverantwoordelijke B dit model niet gebruiken, vanwege het

3. Ibid, paragraaf 135.

4. Ibid, paragraaf, 126.

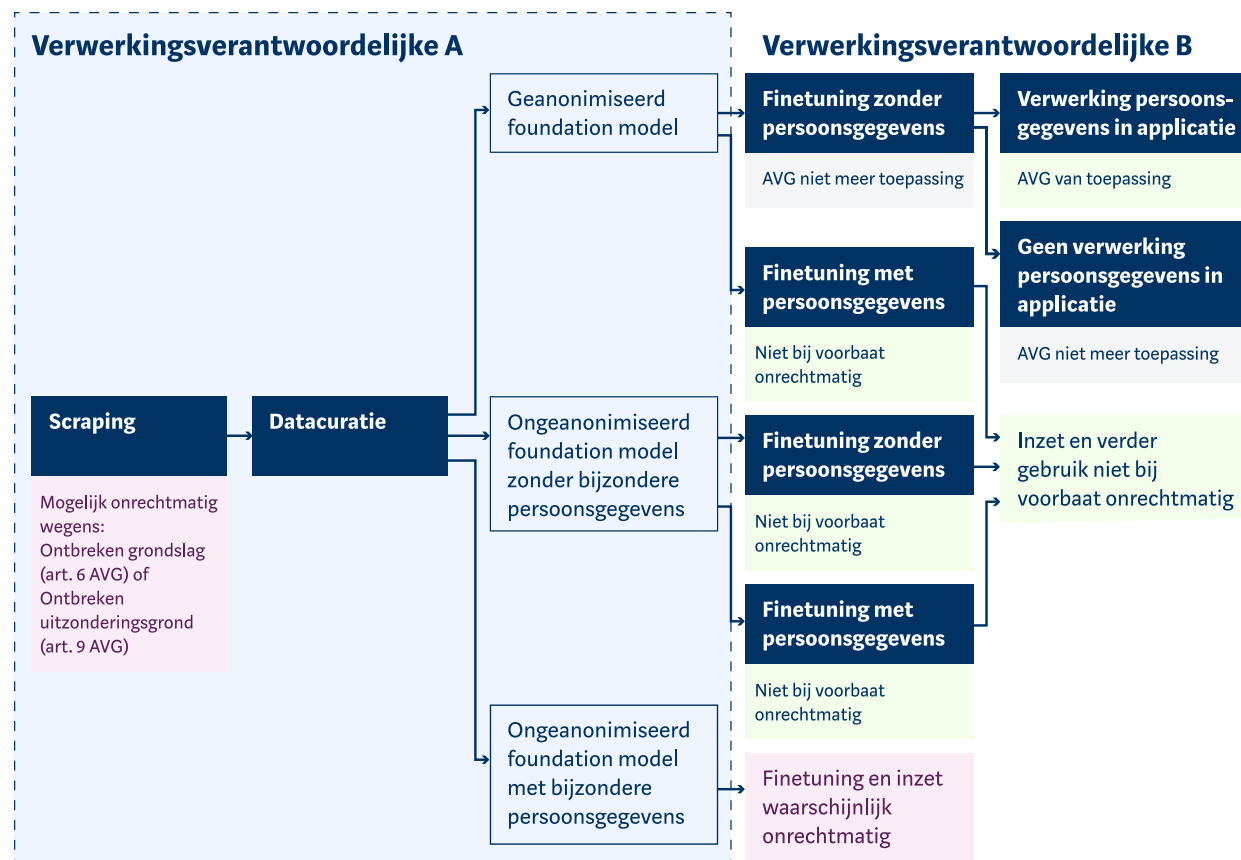
5. Ibid, paragraaf 129.

6. Ibid, paragraaf 130.

verwerkingsverbod in artikel 9 AVG.⁷ Het zal immers voor verwerkingsverantwoordelijke B ook niet mogelijk zijn om zich op een uitzonderingsgrond te beroepen.

Indien uit de beoordeling volgt dat er geen bijzondere persoonsgegevens zijn meegetraind in het foundation model, is het finetunen en de verdere inzet van dat foundation model niet bij voorbaat onrechtmatig. Verwerkingsverantwoordelijke B zal wel moeten beschikken over een eigen grondslag voor de finetuning en inzet van zijn generatieve AI-modellen en -applicaties. Net als in het eerste scenario blijft verwerkingsverantwoordelijke A verantwoordelijk voor de onrechtmatigheden die zich hebben voorgedaan tijdens de dataverzameling, datacuratie en ontwikkeling van het foundation model.

FIGUUR 2 | SCHEMATISCHE WEERGAVE RECHTMATIGHEID GENERATIEVE AI KETEN



7. Zie voor een uitgebreidere uitleg van 'bijzondere persoonsgegevens':

[Wat zijn persoonsgegevens? | Autoriteit Persoonsgegevens](#)

AVG-Randvoorwaarden generatieve AI

Zoals dit schema illustreert, is de AVG in verschillende stappen van de generatieve AI-keten van toepassing, zodra er persoonsgegevens worden verwerkt. Tijdens het verzamelen worden persoonsgegevens verwerkt en ook na het cureren van datasets kunnen deze datasets nog persoonsgegevens bevatten. Zo kunnen datasets, waarop generatieve AI-modellen worden getraind of gefinetuned, persoonsgegevens bevatten. Verder kunnen persoonsgegevens worden ingevoerd tijdens het gebruik van een generatieve AI-toepassing. Denk aan een generatieve AI-toepassing die wordt ingezet als ondersteuning in de zorgsector. Een generatieve AI-toepassing kan bijvoorbeeld conceptantwoorden opstellen voor patiëntvragen, die zorgverleners vervolgens beoordelen en versturen.⁸ Maar ook inmiddels steeds populairder wordende 'AI-agents' kunnen persoonsgegevens verwerken.⁹ Denk aan een AI-agent die assisteert bij het bestellen van producten via internet. Doorgaans zal de AI-agent daarvoor onder meer locatiegegevens en financiële gegevens raadplegen.

De AVG stelt niet alleen verplichtingen aan ontwikkelaars van generatieve AI-modellen ('aanbieders'¹⁰), maar ook aan partijen die deze modellen verder finetunen en/of inzetten voor eigen gebruik ('gebruiksverantwoordelijken'¹¹).

8. Zie bijvoorbeeld: TNO (April 2024). "Generatieve AI in de Nederlandse zorg", <<https://publications.tno.nl/publication/34643183/U5tb8oyL/TNO-2024-R10662.pdf>>.

9. Zie bijvoorbeeld: OpenAI (Januari 2025). "Introducing Operator", <<https://openai.com/index/introducing-operator/>>.

10. Voor een definitie van aanbieders, zie art. 4, derde lid, AI-verordening.

11. Voor een definitie van gebruiksverantwoordelijken, zie art. 4, vierde lid, AI-verordening.

FIGUUR 3 | AVG-RANDVOORWAARDEN GENERATIEVE AI



Hieronder zetten we niet-limitatief enkele randvoorwaarden uiteen voor de verschillende stappen in de generatieve AI-keten: dataverzameling, datacuratie, training en finetunen van het foundation model en inzet van generatieve AI-toepassingen. Per randvoorwaarde wordt de relevantie voor bepaalde partijen gesuggereerd, maar het is aan aanbieders en gebruiksverantwoordelijken zelf om te bepalen welke randvoorwaarden voor hen relevant zijn. Onderstaande randvoorwaarden gaan er dan ook vanuit dat in de verschillende stappen van de generatieve AI-keten sprake is van de verwerking van persoonsgegevens. De randvoorwaarden zijn in de figuur hieronder schematisch weergegeven.

Randvoorwaarden voor dataverzameling

1. Trainingsdata voor het trainen en finetunen van een model moet rechtmatig zijn verkregen

Relevant voor: ontwikkelaars en finetuners

Om generatieve AI-modellen te trainen, zijn grote hoeveelheden data nodig. Momenteel lijkt het erop dat de hiervoor benodigde omvangrijke datasets voor foundation modellen alleen tot stand kunnen komen door in elk geval ook gebruik te maken van (ongerichte) scraping. Daardoor komen er vrijwel zeker persoonsgegevens in dergelijke datasets terecht. In het geval dat eigen reeds verzamelde datasets met persoonsgegevens worden gebruikt voor de training van een foundation model, zal die data afzonderlijk doorgaans niet toereikend zijn om een foundation model te trainen. Voor het verzamelen van die persoonsgegevens heeft de verwerkingsverantwoordelijke een grondslag nodig, zoals beschreven in artikel 6 AVG. Doorgaans bestaat er geen directe relatie tussen betrokkenen van wie data van het internet worden gescrapet en de generatieve AI-ontwikkelaars die de data verzamelen. Daarom lijkt het

erop dat de grondslag 'gerechtvaardigd belang'¹² op dit moment als enige mogelijke grondslag kan worden gebruikt voor het verzamelen van datasets om die foundation modellen te trainen. Of er een beroep op de grondslag gerechtvaardigd belang kan worden gedaan, hangt af van de [EDPB-richtlijnen over gerechtvaardigd belang](#).¹³ Voor een geslaagd beroep op de grondslag gerechtvaardigd belang, moet een gerechtvaardigd belang bestaan. Daarnaast moet de verwerkingsverantwoordelijke de noodzakelijkheidstoets doorstaan en een belangenafweging uitvoeren. Zoals eerder vastgesteld in de [Handreiking scraping van de AP](#), is het niet vanzelfsprekend dat een beroep op de grondslag gerechtvaardigd belang slaagt in het kader van scraping. De voorwaarden voor een geslaagd beroep op de grondslag gerechtvaardigd belang moeten altijd per geval beoordeeld worden.

Voor het ontwikkelen van generatieve AI-modellen en het inzetten van generatieve AI-applicaties is een aantal voorwaarden uit de EDPB-richtlijnen van belang. In het algemeen geldt onder de noodzakelijkheidstoets dat striktere verzamelcriteria voor de dataverzameling sneller zullen leiden tot het slagen van die toets. Gerichtere verzameling draagt uiteraard bij aan de proportionaliteit van de verwerking van persoonsgegevens. Verwerkingsverantwoordelijken zullen dus moeten nagaan of, en zo ja hoeveel, persoonsgegevens moeten worden gescrapet tijdens de dataverzameling. Wanneer persoonsgegevens

worden geanonimiseerd voor zij worden meegetraind in een generatief AI-model, levert dit een aanzienlijke bijdrage aan het slagen van de toets voor gerechtvaardigd belang.

Bij de belangenafweging kunnen striktere verzamelcriteria ook een rol spelen. Strengere verzamelcriteria zullen over het algemeen leiden tot een geringere impact op de rechten van betrokkenen, omdat de omvang van de verzamelde persoonsgegevens hierdoor vermindert. Daar komt bij dat het verwerken van persoonsgegevens doorgaans een bijkomend gevolg en dus een klein onderdeel is van het trainen van een generatief AI-model. Tot slot weegt anonimisering van persoonsgegevens, voordat deze worden meegetraind in een generatief AI-model, zwaar in het voordeel van de verwerkingsverantwoordelijke.

Na het trainen van een foundation model kan het model verder worden gefinetuned op kleinere datasets. De ontwikkelaar van het foundation model kan het model finetunen. Maar een partij die een gefinetuned model uiteindelijk wil integreren in een generatieve AI-toepassing kan dit ook doen. Doorgaans zijn deze datasets voor finetuning een stuk kleiner en niet altijd verkregen door scraping. Als de datasets voor het finetunen persoonsgegevens bevatten, dan geldt dat de verwerkingsverantwoordelijke over een grondslag moet beschikken, zoals beschreven in artikel 6 AVG. De geringere omvang van deze datasets maakt het echter mogelijk dat bijvoorbeeld ook toestemming van betrokkenen als grondslag kan gelden, of dat er sprake kan zijn van een verenigbare verdere verwerking van eigen data onder artikel 6, vierde lid, AVG. Want er kan bij deze kleinere datasets een directe relatie bestaan tussen de betrokkenen en de verwerkingsverantwoordelijke. Verwerkingsverantwoordelijken zullen

12. Artikel 6, eerste lid en onder f, van de AVG.

13. European Data Protection Board (Oktober 2024). "Guidelines 1/2024 on processing of personal data based on article 6(1)(f) GDPR", <https://www.edpb.europa.eu/system/files/2024-10/edpb_guidelines_202401_legitimateinterest_en.pdf>.

echter per geval, en voorafgaand aan de initiële verzameling van persoonsgegevens, moeten beoordelen welke grondslag van toepassing is.

Strengere randvoorwaarden voor de verzameling van bijzondere persoonsgegevens

Als er tijdens de dataverzameling bijzondere persoonsgegevens worden verzameld, moet de verwerkingsverantwoordelijke zich kunnen beroepen op een uitzondering op het verwerkingsverbod voor bijzondere persoonsgegevens.¹⁴ Die beoordeling zal moeten plaatsvinden voordat de bijzondere persoonsgegevens worden meegescrapet. Finetuners van foundation modellen kunnen zich mogelijk op de uitzonderingsgrond toestemming beroepen, vanwege voornoemde directe relatie.¹⁵ Bij (ongericht) scrapen voor het trainen van foundation modellen bestaat die directe relatie doorgaans niet en dus is een beroep op de uitzonderingsgrond toestemming vrijwel onmogelijk. Als de betrokkene de bijzondere persoonsgegevens kennelijk openbaar heeft gemaakt, kan dit als uitzonderingsgrond op het verwerkingsverbod gelden bij ongerichte scraping.¹⁶ Dit zal door de verwerkingsverantwoordelijke moeten worden beoordeeld, voorafgaand aan de dataverzameling. Als de verwerkingsverantwoordelijke zich niet op een uitzonderingsgrond kan beroepen, kan de verwerking van die bijzondere persoonsgegevens niet als rechtmatig worden aangemerkt.

De situatie waarin bijzondere persoonsgegevens worden meegescrapet, zonder dat er sprake is van kennelijke openbaarmaking van die persoonsgegevens door betrokkenen zelf, vormt waarschijnlijk een klein deel van de gehele dataverzameling. Maar dat maakt de verzameling daarvan niet minder onrechtmatig. Voor het verzamelen van bijzondere persoonsgegevens heeft het HvJEU geoordeeld dat het verwerkingsverbod niet mag worden toegepast op een exploitant van een zoekmachine alsof die exploitant zelf de bijzondere persoonsgegevens heeft geplaatst.¹⁷ In een dergelijk geval heeft het HvJEU gesteld dat een exploitant van een zoekmachine voorafgaand aan de dataverzameling niet hoeft te toetsen of er een uitzonderingsgrond van toepassing is.¹⁸ Die beoordeling hoeft door de exploitant pas plaats te vinden nadat er een verwijderingsverzoek is ingediend door een betrokkene. Daarnaast heeft het HvJEU geoordeeld dat er bij zoekmachines een belang speelt voor het grote publiek om toegang tot informatie te krijgen.¹⁹ Dat kan bijdragen aan het recht op informatie dat is beschreven in artikel 11 van het Handvest van de Grondrechten van de Europese Unie.

Het verzamelen van data voor het opbouwen van een trainingsset voor generatieve AI toont gelijkenissen met het verzamelen van data om een zoekdienst aan te bieden. In beide gevallen worden er op een ongerichte manier zoveel mogelijk gegevens binnengehaald. Ook zal in beide gevallen een stukje analyse plaatsvinden op de gegevens.

Daarentegen zijn er ook belangrijke verschillen:

- **Opslaan van gegevens.** Voor een zoekmachine is het van belang om gegevens van een webpagina te analyseren zodat deze goed wordt geïndexeerd. Dit betekent dat de zoekmachine de brongegevens na deze verwerking niet volledig hoeft op te slaan, terwijl dit voor het opbouwen van een trainingsdataset bij generatieve AI juist wel het doel is. De verschillen tussen zoekmachines en gegevensverzameling voor generatieve AI zit in de manier waarop gegevens worden opgeslagen en verder worden gebruikt. Een zoekmachineaanbieder heeft alle gegevens gestructureerd en doorzoekbaar gemaakt. Hierdoor kunnen gegevens makkelijk worden ingezien en verwijderd. Bij generatieve AI worden datasets gecureerd, waardoor bepaalde ongewenste data en persoonsgegevens verwijderd kunnen worden. Toch kan het gebeuren dat er tijdens het trainen van het model achtergebleven stukken trainingsdata in het model terechtkomen, die woordelijk gereproduceerd kunnen worden. Echter kunnen op dit moment specifieke stukjes data nog niet uit een getraind model 'ontleerd' worden. In de toekomst biedt de techniek 'Machine Unlearning' mogelijk een uitkomst, maar op moment van schrijven biedt deze techniek nog geen theoretische garantie op verwijdering van de bijzondere persoonsgegevens uit een getraind model.

14. Artikel 9 AVG.

15. Artikel 9, tweede lid en onder a, AVG.

16. Artikel 9, tweede lid en onder e, AVG. Voor een uitgebreidere uitleg over de kennelijke openbaarmaking bij scraping, zie de [Handreiking scraping door particulieren en private organisaties van de AP](#).

17. Zie de conclusie van de AG in de zaak GC & Others (ECLI:EU:C:2019:14), overweging 55.

18. Zie de conclusie van de AG in de zaak GC & Others (ECLI:EU:C:2019:14), overweging 55 en 56, alsook overweging 48 van onderhavig arrest (ECLI:EU:C:2019:773).

19. Zie ECLI:EU:C:2019:773, overweging 53.

- **Doel en frequentie.** Een zoekmachineaanbieder zal alles scrapen wat technisch toegankelijk is (en niet volgens Robots.txt wordt afgewezen). Zoekmachines proberen per definitie zo goed en actueel mogelijke zoekresultaten aan te bieden. Dit betekent dat ze doorlopend het internet doorzoeken en in principe alle beschikbare informatie raadplegen. Voor scraping wordt op een bepaald moment data van het internet gescrapet en een trainingsset vastgesteld waar een model vervolgens weken of maanden op wordt getraind. Dit proces is wel iteratief, maar is desalniettemin gebaseerd op een momentopname van de gescrapete data. Ook kunnen er regels worden meegegeven om bepaalde websites direct uit te sluiten in de verzamelcriteria van de dataverzameling. Bij scraping is het namelijk geen doel op zich om alle websites terug te geven in zoekresultaten. Lage kwaliteit of ongewenste data (denk aan extremistische fora, of gezondheidsfora waar veel bijzondere persoonsgegevens te vinden zijn) zal juist uitgesloten moeten worden om een goed werkend model te maken.

De AP constateert daarom dat zoekmachines en scraping voor generatieve AI aanzienlijke overeenkomsten vertonen. De momentopname van de verwerking van bijzondere persoonsgegevens bij scraping heeft mogelijk een lagere impact op betrokkenen dan de continue verwerking daarvan door zoekmachines. Echter vormt de mogelijkheid om te voldoen aan een verwijderverzoek bij zoekmachines een belangrijk verschil met het voldoen aan verwijderverzoeken bij getrainde generatieve AI-modellen. De techniek 'Machine Unlearning' biedt mogelijk in de toekomst een oplossing voor reeds getrainde modellen. De AP ziet gelijkenissen en verschillen tussen scraping voor Genera-

tieve AI en zoekmachines als het gaat om het recht op informatie. Sommige generatieve AI-applicaties zouden namelijk bij kunnen dragen aan de informatievoorziening voor het grotere publiek. Maar door de brede inzetbaarheid van foundation modellen kan ook worden gesteld dat de functie van die modellen niet alleen blijft bij informatieverschaffing aan een breder publiek.

De AP sluit daarom niet uit dat de beoordeling van het HvJEU of er sprake is van een uitzonderingsgrond voor het verwerken van bijzondere persoonsgegevens in de toekomst ook op getrainde generatieve AI-modellen van toepassing kan zijn. Echter vormt het aanzienlijke karakter dat foundation modellen vertonen een belangrijk verschil in de informatieverschaffing aan een breder publiek. Voornoemde uitspraak kan daarom niet een-op-een toegepast worden op scraping voor het trainen van generatieve AI-modellen. Dat betekent dat verwerkingsverantwoordelijken, voorafgaand aan het scrapen van data voor generatieve AI, zullen moeten beoordelen of zij zich op een uitzonderingsgrond kunnen beroepen voor het verwerken van bijzondere persoonsgegevens.

Randvoorwaarden voor datacuratie

2. Trainingsdata moet rechtmatig en zorgvuldig gecureerd worden en zo goed als mogelijk worden ontdaan van (ongewenste) persoonsgegevens

Relevant voor: ontwikkelaars en finetuners

Doorgaans worden datasets opgeschoond voordat zij gebruikt worden om een generatief AI-model te trainen of te finetunen. Dit wordt ook wel 'datacuratie' genoemd. Tijdens deze datacuratie worden ongewenste data, zoals beledigende taal of (ongewenste) persoonsgegevens, uit datasets gehaald. In lijn met het vorige vereiste, is het van

belang dat datasets die worden gecureerd door een verwerkingsverantwoordelijke, rechtmatig zijn verzameld. Deze verplichting geldt zowel voor ontwikkelaars van generatieve AI-modellen als voor finetuners van die modellen, die hun modellen trainen of finetunen op een gecureerde dataset. De datacuratie is voor de bescherming van persoonsgegevens een belangrijke stap, omdat het de laatste stap is waarbij expliciet gecontroleerd kan worden welke persoonsgegevens worden meegetraind in het model. Zodra de persoonsgegevens impliciet in het model worden meegetraind, is de controle en het uitvoeren van bijkomende rechten van complexere aard. Het cureren van datasets kan onder de vereisten voor de grondslag gerechtvaardigd belang zelfs als noodzakelijk worden geacht. Dit geldt wanneer bepaalde persoonsgegevens niet noodzakelijk zijn voor de ontwikkeling van een generatief AI-model. Het cureren van datasets op bepaalde (ongewenste) persoonsgegevens draagt immers bij aan het beginsel van dataminimalisatie.²⁰

Omdat datacuratie als verwerking onder de AVG kan worden aangemerkt²¹, moet de verwerkingsverantwoordelijke over een rechtmatige grondslag beschikken, zoals beschreven in artikel 6 van de AVG. Als dezelfde verwerkingsverantwoordelijke zowel de dataverzameling als de datacuratie uitvoert, kunnen beide verwerkingen mogelijk onder dezelfde grondslag plaatsvinden,

20. Het gaat hier om de tweede stap van de toets die verwerkingsverantwoordelijken moeten uitvoeren wanneer zij zich op de grondslag gerechtvaardigd belang beroepen, ook wel de 'noodzakelijkheidstoets' genoemd. Zie Guidelines 1/2024 on processing of personal data based on Article 6(1)(f) GDPR, p. 12.

21. Het bewerken of wissen van gegevens wordt aangemerkt als verwerking onder artikel 4, aanhef en onder 2, AVG.

mits voldaan aan de transparantievereisten uit de AVG.²² De verwerkingsverantwoordelijken zullen dit echter per geval en voorafgaand aan de verwerking moeten beoordelen. Ook deze verplichting geldt zowel voor ontwikkelaars van generatieve AI-modellen als voor finetuners van deze modellen, in het geval zij hun modellen (verder) trainen op een gecureerde dataset.

Tot slot vindt datacuratie doorgaans geautomatiseerd plaats, waardoor de vraag rijst of die curatie 100% garantie biedt dat ongewenste persoonsgegevens niet meer terugkomen in de opgeschoonde dataset. Er bestaan echter ook uitzonderlijke gevallen waarin generatieve AI-modellen persoonsgegevens hebben gereproduceerd, ondanks het cureren van datasets.²³ Dit wordt ook wel ‘ophoesting’ genoemd. Een dergelijke ophoesting is dan het gevolg van het meetrainen van persoonsgegevens, omdat de automatische curatie van die persoonsgegevens niet (volledig) is geslaagd. Het is daarom van belang dat verwerkingsverantwoordelijken state-of-the-art technieken gebruiken om dergelijke ophoesting te voorkomen.

Ophoesting kan mogelijk worden onderdrukt, door na de training op ruwe data een extra training te doen op specifieke voorbeelden, waarvan mensen hebben aangegeven

dat dit ongewenste informatie is.²⁴ Door het model te leren dat het niet gepast is om persoonsgegevens te produceren, zal het minder geneigd zijn om die persoonsgegevens op te hoesten. Als een model verder gefinetuned kan worden, is het wellicht wel mogelijk om dit ongedaan te maken. Naast onderdrukking in het model, is het mogelijk om de output te checken (*filteren*) op persoonsgegevens. Als deze worden gedetecteerd, zou een standaardantwoord kunnen worden teruggegeven of een nieuw antwoord kunnen worden gegenereerd in plaats van de gegenereerde output. Dit is een oplossing die op systeemniveau werkt, maar niet het model direct aanpast. Als gevolg daarvan zitten de persoonsgegevens wel nog steeds in het model. Dit werkt feitelijk dus wel, maar komt niet in de plaats van de uitoefening van rechten door betrokkenen.

Als een generatief AI-model **aantoonbaar** is geanonimiseerd en toch persoonsgegevens produceert, kan worden aangenomen dat een dergelijke reproductie van persoonsgegevens een ‘hallucinatie’ vormt. Een generatief AI-model produceert aannemelijke inhoud die gebaseerd is op veelvoorkomende patronen in de data. Het model heeft echter geen daadwerkelijke kennis van de wereld en kan dus aannemelijk klinkende onwaarheden produceren. Dit worden wel hallucinaties genoemd. Die hallucinatie is dan geen gevolg van het meetrainen van een persoonsgegeven, maar een inherent falen van de technologie. Het verschil tussen hallucinaties (geen persoonsgegevens) en ophoesting (wel persoonsgegevens) door een generatief AI-model is van belang voor het inwilligen van rechten van

betrokkenen. Persoonsgegevens die niet in de dataset staan of meegetraind zijn in het model, kunnen niet worden gerectificeerd of verwijderd. De volgende randvoorwaarde gaat verder in op rechten van betrokkenen, onder andere bij ophoesting.

Randvoorwaarden voor training van een foundation model & finetuning van een model

3. Verwerkingsverantwoordelijken hebben een systeem ingesteld om rechten van betrokkenen te faciliteren
Relevant voor: ontwikkelaars, finetuners en gebruiksverantwoordelijken

Op basis van het voorgaande is bij een geanonimiseerd model geen sprake van een verwerking van persoonsgegevens. Deze randvoorwaarde gaat dus over de situatie waarin een model niet aantoonbaar geanonimiseerd is. Betrokkenen hebben verschillende rechten wat betreft hun persoonsgegevens in de AVG. Bijvoorbeeld het recht op inzage, het recht op rectificatie en het recht op gegevenswissing.²⁵ De verplichting om die rechten te faciliteren, ligt bij de verwerkingsverantwoordelijke die het verzoek ontvangt. Betrokkenen kunnen hun rechten uitoefenen bij verschillende stappen in de generatieve AI-keten. Aanbieders en gebruiksverantwoordelijken van generatieve AI-modellen moeten een systeem inrichten om te voldoen aan de rechten van betrokkenen, op het moment dat zij dergelijke verzoeken binnenkrijgen.

De geleerde patronen van generatieve AI-modellen zitten besloten in getallen, ook wel ‘gewichten’ genoemd. De informatie die hierin ligt opgeslagen, is niet langer expliciet

22. In lijn met het beginsel van doelbinding en de transparantieplichtingen richting betrokkenen moeten betrokkenen voorafgaande aan de dataverzameling en datacuratie op de hoogte zijn van de doeleinden voor die verwerking. Zie voor meer informatie artikel 5, eerste lid en onder b, AVG en artikel 12-14 AVG.

23. Zie bijvoorbeeld: Carlini et al. (2021). “Extracting Training Data from Large Language Models”, <<https://www.usenix.org/conference/usenixsecurity21/presentation/carlini-extracting>>. Zie ook: Lukas et al. (2023). “Analyzing Leakage of Personally Identifiable Information in Language Models”, <<https://arxiv.org/abs/2302.00539>>.

24. Dit wordt ook wel “Reinforcement Learning from Human Feedback” genoemd.

25. Zie onder andere artikelen 15, 16 en 17 van de AVG.

gerepresenteerd, maar is impliciet onderdeel van de verzameling gewichten. Eventuele persoonsgegevens daarin zijn dus niet direct aanwijsbaar. Echter vormen de parameters een statistische relatie tussen de traininggegevens, waardoor het mogelijk is om persoonsgegevens te extraheren uit het model. Bijvoorbeeld door het model te bevragen of door rechtstreeks relaties tussen de gegevens in het model te extraheren. De informatie in het model kan dus betrekking hebben op een natuurlijke persoon, ondanks dat het model technisch is georganiseerd of gecodeerd. Daardoor is een relatie met een natuurlijke persoon niet onmiddellijk duidelijk.²⁶ Dat betekent dat generatieve AI-modellen wel degelijk persoonsgegevens kunnen bevatten.

De AP is zich bewust van de moeilijkheden bij het faciliteren van rechten van betrokkenen in reeds getrainde modellen die niet aantoonbaar zijn geanonimiseerd. Om te voldoen aan het recht van inzage zou een verwerkingsverantwoordelijke (met behulp van een zoekfunctie) inzage kunnen geven in de gebruikte dataset voor het trainen of finetunen van een model. Ook kunnen betrokkenen hun recht op gegevenswissing en rectificatie van deze datasets uitoefenen. Door oude modellen snel uit te faseren, kunnen verwerkingsverantwoordelijken voldoen aan het recht op gegevenswissing en rectificatie van een reeds getraind of gefinetuned model. Het nieuwe model moet dan worden getraind of gefinetuned op een dataset zonder de persoonsgegevens. In de toekomst kunnen bepaalde gegevens mogelijk 'ontleerd' worden uit deze modellen met de techniek 'Machine Unlearning'.

26. EDPB Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models, paragraaf 37.

Echter biedt deze techniek op het moment van schrijven nog geen garantie dat de persoonsgegevens worden ontleerd uit het model. Hertraining van het model is daarom nu de enige oplossing.

In de praktijk stellen aanbieders foundation modellen vaak zonder trainingssets beschikbaar aan finetuners of gebruiksverantwoordelijken. Eventuele meegetrainde persoonsgegevens in dat foundation model zijn niet expliciet gerepresenteerd. Zonder beschikking over de trainingsset, hebben die partijen vrijwel geen middelen om betrokkenen te identificeren.²⁷ Met andere woorden, de persoonsgegevens zijn wel verwerkt in het generatieve AI-model van de aanbieder, maar identificatie door de finetuner of gebruiksverantwoordelijke is daarbij geen doel op zich. In dat geval hoeft de finetuner of gebruiksverantwoordelijke niet te voldoen aan de rechten van betrokkenen, volgens artikel 11, tweede lid, AVG. Pas op het moment dat een betrokkene zelf aan de finetuner of gebruiksverantwoordelijke demonstreert dat zijn persoonsgegevens in het model zitten, komt de verplichting om rechten van betrokkenen te faciliteren weer tot leven.²⁸

Inzage in de trainingsdata zal in dat geval niet mogelijk zijn, omdat finetuners en gebruiksverantwoordelijken niet over die trainingsdata beschikken. Voor verwijdering en rectificatie lijkt hertraining van het model op dit moment de enige oplossing. Maar omdat het foundation model niet door hen zelf is getraind, biedt dit geen oplossing. In de toekomst kan 'Machine Unlearning' voor dit probleem

27. Zie HvJEU *Breyer v. Bundesrepublik Deutschland*, ECLI:EU:C:2016:779, par. 42.

28. Zie ook overweging 57 van de AVG.

mogelijk wel een uitkomst bieden. Deze moeilijkheden bij het inwilligen van rechten van betrokkenen kunnen worden voorkomen. Aanbieders en finetuners kunnen contractueel afspreken dat aanbieders inzage verschaffen in recente trainingssets aan finetuners of gebruiksverantwoordelijken, op het moment dat zij daarvoor een verzoek ontvangen. Ook kan contractueel worden afgesproken dat aanbieders een nieuw model aanleveren zonder de betreffende persoonsgegevens, als er een finetuner of gebruiksverantwoordelijke een rectificatie of verwijderingsverzoek ontvangt. Verder zouden aanbieders bij het beschikbaar stellen van foundation modellen de meest recent gebruikte trainingsset kunnen aanleveren aan finetuners of gebruiksverantwoordelijken. Dan kan namelijk het recht op inzage worden gefaciliteerd door inzage te geven in die trainingsdata. Als deze contractuele afspraken niet plaatsvinden, lijkt het erop dat een finetuner of gebruiksverantwoordelijke geen gevolg kan geven aan de rechten van betrokkenen. Daardoor kan het beschikbaar stellen van foundation modellen niet als rechtmatig worden aangemerkt.

4. Doeleinden voor het meetrainen van persoonsgegevens in generatieve AI-modellen en voor het verwerken van persoonsgegevens in generatieve AI-toepassingen moeten voorafgaand worden bepaald en gecommuniceerd aan betrokkenen

Relevant voor: ontwikkelaars, finetuners en gebruikersverantwoordelijken

Volgens het beginsel van doelbinding mogen persoonsgegevens alleen worden verwerkt als verwerkingsverantwoordelijken voorafgaand aan die verwerking welbepaalde verwerkingsdoeleinden opstellen en communiceren aan

betrokkenen.²⁹ Deze doeleinden moeten rechtmatig zijn en specifiek en expliciet genoeg worden geformuleerd, zodat betrokkenen weten waarvoor hun persoonsgegevens worden verwerkt. Als het meetrainen van persoonsgegevens in een model noodzakelijk is, moeten verwerkingsverantwoordelijken hier dus duidelijk over communiceren aan betrokkenen.

Generatieve AI-modellen kunnen voor een breed scala aan taken worden ingezet, waarbij finetuning vaak in een later stadium plaatsvindt. Ontwikkelaars en finetuners van deze generatieve AI-modellen moeten verwerkingsdoeleinden specifiek genoeg formuleren voordat zij persoonsgegevens meetrainen in een foundation model. Daarbij is het van belang dat in de verwerkingsdoeleinden in ieder geval enige context over de inzet van een foundation model gegeven wordt. Bijvoorbeeld met een beschrijving van de functionaliteiten van het model, of door te beschrijven of het model is ontwikkeld voor interne doeleinden of bedoeld is voor latere distributie of verkoop.³⁰ Finetuners kunnen doeleinden doorgaans specifiek formuleren bij het finetunen van foundation modellen dan ontwikkelaars bij het trainen van foundation modellen.

Wanneer ontwikkelaars of finetuners van generatieve AI-modellen persoonsgegevens verzamelen met scraping, is mogelijk de uitzondering van artikel 14, vijfde lid onder b, AVG van toepassing. Daaruit volgt dat wanneer informatieverstrekking aan betrokkenen onmogelijk blijkt

of onevenredig veel inspanning zou vergen, die informatieverstrekking achterwege kan blijven. In dat geval moet de verwerkingsverantwoordelijke nog wel informatie verschaffen over de verwerkingsdoeleinden op zijn website of in zijn toepassing.

Ook gebruiksverantwoordelijken van generatieve AI-toepassingen moeten voldoen aan het beginsel van doelbinding. Als de gebruiksverantwoordelijke van een generatieve AI-toepassing persoonsgegevens verwerkt, moet hij deze doeleinden voorafgaand aan de verwerking bepalen en communiceren aan betrokkenen. De doeleinden hangen af van de soort generatieve AI-toepassing die wordt ingezet. Gebruiksverantwoordelijken van generatieve AI-toepassingen zullen doorgaans een directe relatie hebben met betrokkenen, waardoor de uitzondering van artikel 14, vijfde lid AVG niet van toepassing is.

Randvoorwaarden voor de inzet van generatieve AI-toepassingen

5. Generatieve AI-toepassingen genereren zo weinig mogelijk foutieve of ongewenste persoonsgegevens

Relevant voor: gebruiksverantwoordelijken

Getrainde generatieve AI-modellen werken niet met een expliciet kennismodel, maar produceren goed lopende tekst, plaatjes of video's op basis van kansberekening. Persoonsgegevens kunnen daardoor foutief worden geproduceerd, omdat het model een onterechte associatie maakt. Daarnaast kunnen generatieve AI-modellen per ongeluk persoonsgegevens uit de trainingsset ophoesten (ook wel 'regurgitation' genoemd). Daardoor kunnen uitkomsten in generatieve AI-toepassingen ongewenst zijn.

Onder de AVG moeten persoonsgegevens juist zijn met inachtneming van de doeleinden waarvoor zij worden verwerkt.³¹ Indien persoonsgegevens onjuist zijn, moeten deze, met inachtneming van alle redelijke mogelijkheden, worden gewist of gerectificeerd.³² Het beginsel van juistheid moet worden gezien in relatie tot de risico's en gevolgen voor betrokkenen bij het gebruik van hun persoonsgegevens.³³

Voor gebruiksverantwoordelijken van generatieve AI-toepassingen betekent dit dat zij alle redelijke maatregelen moeten nemen om reproductie van foutieve of ongewenste persoonsgegevens te voorkomen. Nieuwe technologieën zoals Retrieval Augmented Generation (RAG) en Chains of Thought (CoT) kunnen daarbij uitkomst bieden om reproductie van foutieve en ongewenste persoonsgegevens te verminderen.³⁴ Door Retrieval Augmented Generation genereert het systeem inhoud op basis van opgezochte bronnen. Bij Chains of Thought gebeurt de besluitvorming aan de hand van meerdere stappen of aaneenschakelingen.

Verder is het van belang dat de verwerkingsverantwoordelijke de betrokkenen op een duidelijke en begrijpelijke manier op de hoogte stelt van de mogelijkheid op foutieve en ongewenste uitkomsten in generatieve AI-toepassingen. Hierdoor draagt de verwerkingsverantwoordelijke bij aan

31. EDPB Richtsnoeren 4/2019 inzake artikel 25 Gegevensbescherming door ontwerp en door standaardinstellingen, paragraaf 77.

32. Artikel 5, eerste lid en onder d, AVG.

33. EDPB Richtsnoeren 4/2019 inzake artikel 25 Gegevensbescherming door ontwerp en door standaardinstellingen, paragraaf 78.

34. Wei et al. (2022). "Chain-of-thought prompting elicits reasoning in large language models", <<https://arxiv.org/abs/2201.11903>>.

29. Zie artikel 5, eerste lid en onder b, AVG.

30. Zie voor meer informatie: EDPB Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models, paragraaf 64.

de toename van 'AI-geletterdheid' in de samenleving. Burgers begrijpen daardoor dat er fouten kunnen voorkomen in de (output van) generatieve AI-toepassingen. Als uit een inzageverzoek blijkt dat het persoonsgegeven niet in de trainingsset zat, maar er wel een antwoord is geproduceerd met dit persoonsgegeven, dan kan worden aangenomen dat dit persoonsgegeven foutief is geproduceerd (gehallucineerd). De verwerkingsverantwoordelijke dient zo veel mogelijk mitigerende maatregelen te nemen om die foutieve of ongewenste output te vermijden. Als het model toch een persoonsgegeven hallucineert, moet de verwerkingsverantwoordelijke kunnen aantonen dat die hallucinatie geen gevolg is van het meetraineren van dat persoonsgegeven.

3. Afsluitend: de inzet vanuit de AP

De AP gaat zich komende periode inzetten om een bijdrage te leveren aan de in het visiestuk beschreven en gewenste toekomstbeeld: "Waarden aan het werk". 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 gaat de AP bijdragen aan heldere en realistische werkmethoden, onder andere door actief mee te schrijven aan standaarden en opinies. Bijvoorbeeld de [EDPB-opinie uit december 2024](#) over het trainen van AI-modellen. Daarnaast werkt de AP aan digitale weerbaarheid, onder andere door te investeren in het kennisniveau van de samenleving. Bijvoorbeeld met [guidance](#) voor AI-geletterdheid, begrijpelijke informatie op onze [website](#) en seminars. Ook risicosignalering vormt een integraal onderdeel van de werkzaamheden van de AP. Bijvoorbeeld via de halfjaarlijkse [risicorapportages](#) over algoritmes en AI. Als laatste is de AP beschikbaar voor [voorafgaande raadpleging](#) en richt het een sandboxtraject in voor hoogrisicotoepassingen volgens de AI-verordening.

De AP gaat ook een aantal extra stappen zetten om verantwoord vooruit te gaan met generatieve AI. We beginnen in te kaart brengen 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, vormt de basis voor verdere periodieke dialoog over de verantwoorde ontwikkeling en inzet van generatieve AI.

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 toezicht-houders in het digitale domein zetten we ons daarnaast in voor gezamenlijke normuitleg, zodat we zo veel 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 hoe verantwoorde generatieve AI er in de praktijk uit kan zien en gebruikers te inspireren. En natuurlijk blijft de AP toezichthouden op onrechtmatigheden die zich voordoen in het speelveld van generatieve AI.

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