Kenniscentrum
Data & Maatschappij

ARTIFICIËLE INTELLIGENTIE EN GEGEVENS-BESCHERMING: EEN VERKENNENDE GIDS

Kenniscentrum
Data & Maatschappij

JUNI 2020

Deze gids citeren als:
Gils, T., Wauters, E., Bénichou, B., De Bruyne, J. & Valcke, P. (mei 2020). Artificiële Intelligentie en gegevensbescherming: een verkennende gids. Kenniscentrum Data & Maatschappij, Brussel, België.

www.data-en-maatschappij.ai

Inhoudstafel

1. INLEIDING EN OPBOUW GIDS 7

2.WAT IS ARTIFICIËLE INTELLIGENTIE? 10

3.OP WELKE ACTIVITEITEN IS DE AVG VAN TOEPASSING EN WELKE ROLLEN KAN EEN ORGANISATIE VERVULLEN ONDER DE AVG? 16

3.1. WAT ZIJN PERSOONSGEGEVENS? 17

3.2. WAT ZIJN BIJZONDERE CATEGORIEËN VAN PERSOONSGEGEVENS? 19

3.3. WAT IS HET BELANG VAN HET ONDERSCHEID TUSSEN ANONIMISERING EN PSEUDONIMISERING? 20

3.4. WAT MET NIET-PERSOONSGEBONDEN GEGEVENS EN/OF GEMENGDE GEGEVENSSETS? 23

3.5. WAT ZIJN DE VERSCHILLENDE ROLLEN DIE EEN ORGANISATIE KAN VERVULLEN ONDER DE GDPR IN EEN AI-CONTEXT? 24

4. HOE KAN GEGEVENSBESCHERMING WORDEN VERZEKERD BIJ HET ONTWERP (DESIGN) EN DE ONTWIKKELING (DEVELOPMENT) VAN AI-SYSTEMEN? 29

4.1. WAT BETEKENT GEGEVENSBESCHERMING DOOR ONTWERP EN HOE KAN DIT GEÏMPLEMENTEERD WORDEN IN AI-SYSTEMEN? 29

4.1.A. Gegevensbescherming inbouwen in applicaties en processen 31

4.1.B. Risico-gebaseerde benadering 32

4.1.C. Gegevensbescherming door ontwerp heeft betrekking op het naleven van alle bepalingen in de AVG 33

4.1.D. Risicoanalyse 34

4.1.E. Documenteer de gemaakte evaluaties en de genomen maatregelen 35

4.1.F. Vermijd '*technical debt*' door gegevensbescherming door ontwerp toe te passen 36

4.1.G. Gegevensbescherming door standaardinstellingen 36

4.1.H. Certificering 37

4.2. WAT HOUDT DE VEREISTE VAN MINIMALE GEGEVENSVERWERKING VAN PERSOONSGEGEVENS IN VOOR AI-SYSTEMEN? 37

4.2.A. Algemeen 38

4.2.B. Risico-gebaseerde benadering 39

4.2.C. Minimale gegevensbescherming en de doeltreffendheid van AI-systemen 40

4.2.D. Verzamel en gebruik minder persoonsgegevens (trainingsfase en gebruiksfase) 40

4.2.E. Een performant en afgedwongen databeleid: beperk de toegang tot persoonsgegevens 49

4.3. WAT ZIJN AANDACHTSPUNTEN VOOR DE BEVEILIGING VAN DE VERWERKING VAN

PERSOONSGEGEVENS DOOR AI-SYSTEMEN? 49
4.3.A. Technische en organisatorische beveiliging van de gehele omgeving waarin persoonsgegevens verwerkt worden 50
4.3.B. Risico-gebaseerde benadering 51
4.3.C. Gegevensveiligheidsbeleid 51
4.3.D. Databeheer en datamapping 52
4.3.E. Technische maatregelen 52
4.3.F. Bewustmaking en opleiding van al wie toegang heeft tot persoonsgegevens 55
4.3.G. Afspraken met leveranciers en verwerkers 56
4.3.H. Documenteer 57
4.4. WANNEER MOET EEN DPIA OF GEB WORDEN UITGEVOERD VOOR EEN VERWERKING VAN PERSOONSGEGEVENS DOOR AI-SYSTEMEN? 57
4.4.A. Algemeen 58
4.4.B. Bepalen van het risico 58
4.4.C. Lijst van de GBA 61
4.4.D. Tijdstip van het uitvoeren van de DPIA 63
4.4.E. Wie moet een DPIA uitvoeren? 63
4.4.F. Stappenplan 64
4.5. WELKE VERPLICHTINGEN GELDEN ER WANNEER PERSOONSGEGEVENS WORDEN VERWERKT MET HET OOG OP WETENSCHAPPELIJK ONDERZOEK OF STATISTISCHE DOELEINDEN? 65
4.5.A. Wetenschappelijke doeleinden of statistische doeleinden 67
4.5.B. Toepasselijke AVG-principes 68
4.5.C. Toepasselijke Belgische wetgeving 69

5.HOE KAN GEGEVENSBESCHERMING WORDEN VERZEKERD TIJDENS DE GEBRUIKSFASE (DEPLOYMENT) VAN AI-SYSTEMEN? 75

5.1. WELKE TRANSPARANTIEVERPLICHTINGEN LEGT DE AVG OP EN WAT ZIJN DE SPECIFIEKE AANDACHTSPUNTEN IN EEN AI-CONTEXT? 75
5.1.A. Externe en interne transparantie 76
5.1.B. Transparantieverplichtingen onder de AVG 77
5.1.C. Kort toegelicht: explainability 83
5.2. WELKE BEPERKINGEN LEGT DE AVG OP VOOR DE BEWARING VAN PERSOONSGEGEVENS? 84
5.3. WELKE RECHTEN HEBBEN DE BETROKKENEN ALS HUN GEGEVENS DOOR AI-SYSTEMEN WORDEN VERWERKT? 86
5.3.A. Gemeenschappelijke bepalingen 87
5.3.B. Recht van inzage 89

5.3.C. Recht op rectificatie 90
5.3.D. Recht op gegevenswissing ('recht op vergetelheid') 91
5.3.E. Recht op beperking van de verwerking 92
5.3.F. Recht op overdraagbaarheid van gegevens 93
5.3.G. Recht van bezwaar 95
5.4. WAT ZEGT DE AVG OVER GEAUTOMATISEERDE INDIVIDUELE BESLUITVORMING EN PROFILERING EN WAT IS DE IMPACT HIERVAN OP AI-SYSTEMEN? 97
5.4.A. Voordelen en nadelen profilering en geautomatiseerde besluitvorming 99
5.4.B. Profilering 99
5.4.C. Geautomatiseerde besluitvorming: de mens beslist 101
5.4.D. Uitsluitend geautomatiseerde besluitvorming: het AI-systeem beslist 101
5.4.E. Bijzondere categorieën van gegevens en kinderen 104
5.4.F. Gerelateerde rechten van betrokkenen 105

6.BESLUIT 108

7. BIBLIOGRAFIE 110

Hoofdstuk 1: Inleiding

1. Inleiding en opbouw gids

Context gids – Bijna dagelijks verschijnen wel berichten over nieuwe systemen en applicaties die gebruik maken van artificiële intelligentie (AI). Deze snelle ontwikkeling van AI-systemen is een goede zaak gelet op de vele voordelen die het met zich meebrengt. Toch zijn er ook een aantal juridische, ethische en maatschappelijke uitdagingen die moeten worden aangepakt. Het is van belang dat de ontwikkeling en het gebruik van AI-systemen gebeurt binnen het bestaande regelgevende kader.

Omdat AI-systemen in regel grote hoeveelheden data gebruiken, is de Algemene Verordening Gegevensbescherming (AVG) van groot belang. De AVG beschermt de grondrechten en de fundamentele vrijheden van natuurlijke personen en met name hun recht op bescherming van persoonsgegevens. Bepalingen uit de AVG moeten dus worden nageleefd bij het ontwerp, de ontwikkeling en het gebruik van AI-systemen.

Verschillende buitenlandse gegevensbeschermingsautoriteiten publiceerden hierover de voorbije maanden studies en officiële beleidsdocumenten. Met deze verkennende gids over AI en gegevensbescherming wil ook het Vlaams Kenniscentrum Data en Maatschappij (KCDM) de toepassing van de AVG op AI-systemen verduidelijken.

Totstandkoming gids – Deze gids kwam tot stand door middel van overleg met en input door belanghebbenden en met steun van het Vlaams Departement Economie, Wetenschap & Innovatie (EWI). Na intern overleg werd een inhoudstafel opgesteld en een voorstel van te behandelen onderwerpen opgemaakt. Dit werd aan de belanghebbenden bezorgd. Op grond van hun input werden de inhoudstafel en de behandelde onderwerpen veranderd en/of verfijnd. Onderzoekers aan het KU Leuven Centre for IT & IP Law (CiTiP) zijn verantwoordelijk voor de inhoud en coördinatie van deze gids. Bijkomende feedback, vragen en input op/over deze gids kan te alle tijden aan hun worden overgemaakt. De fiches en praktische tools op grond van deze gids worden in overleg met onderzoekers aan VUB-SMIT verspreid. Tot slot wordt ook David Stevens, voorzitter van de Belgische Gegevensbeschermingsautoriteit, bedankt voor zijn waardevolle feedback.

Doelstellingen gids – Deze gids heeft twee doelstellingen. Enerzijds wil het organisaties en gebruikers informatie geven over de toepassing van de AVG bij het ontwerp, de ontwikkeling en het gebruik van AI-systemen. Anderzijds vormt deze gids het raamwerk waar andere praktische instrumenten uit zullen volgen. Zo werden bijvoorbeeld al de eerste praktische fiches door het KCDM verspreid, getiteld 'Wanneer is de AVG wel/niet van toepassing?' en 'Welke rol heb ik onder de AVG?'. In de loop van de komende maanden zullen gelijkaardige praktische tools worden verspreid die berusten op de in deze gids behandelde onderwerpen.

Opbouw gids – Deze gids is verder opgebouwd uit vier onderdelen:

- Hoofdstuk 2 bespreekt het begrip artificiële intelligentie en enkele andere fundamentele concepten.
- Hoofdstuk 3 onderzoekt het toepassingsgebied van de AVG en past dit waar nodig toe in de context van AI.
- Hoofdstuk 4 gaat na hoe gegevensbescherming bij het ontwerp en de ontwikkeling van AI-systemen kan worden verzekerd.
- Hoofdstuk 5 bestudeert hoe gegevensbescherming kan worden verzekerd tijdens het gebruik van AI-systemen.

Waar nuttig vangt elk onderdeel aan met een overzichtskader, waarin zowel de essentie van het besproken onderdeel als een aantal concrete acties worden toegelicht. Vervolgens worden de toepasselijke bepalingen uit de AVG uitvoerig besproken. Deze *meerlagige aanpak* zorgt ervoor dat deze gids naast

een uitgebreide (juridische) analyse ook een praktisch instrument tracht te zijn. Concreet komt het er dus op neer om eerst na te gaan of er een fiche door het KCDM werd gepubliceerd aangaande een bepaald onderwerp. Is dit (nog) niet het geval, dan kunnen de praktische stappen in dit rapport worden bekeken. Bijkomende informatie kan vervolgens worden gevonden in de respectievelijke onderdelen. Er werd ook gekozen om in een algemene bibliografie per hoofdstuk aan te geven met welke officiële beleids- en overheidsdocumenten werd rekening gehouden. Voetnoten werden gebruikt om naar de relevante bepalingen in de AVG te verwijzen of naar specifieke auteurs/bronnen (andere dan de eerder algemene beleids- en overheidsdocumenten).

Beperkingen gids – Binnen deze gids is het niet mogelijk om alle thema's over gegevensbescherming en AI te behandelen (zoals de rol en taken van de functionaris van gegevensbescherming of bindende bedrijfsvoorschriften). Er werd gekozen om op die thema's in te gaan die specifiek van belang zijn voor AI waarbij voor algemene vragen in verband met de AVG naar algemene gidsen en tools gekeken kan worden die elders gepubliceerd werden. Niet-behandelde thema's kunnen nog afzonderlijk worden behandeld door fiches en/of andere praktische tools.

Over het KCDM – Het Kenniscentrum Data & Maatschappij is een samenwerking tussen drie universitaire onderzoeksgroepen: imec-SMIT-VUB, KU-Leuven CiTiP en imec-MICT-UGent. Het maakt deel uit van het Vlaams Beleidsplan Artificiële Intelligentie en krijgt steun van de Vlaamse overheid (EWI). Het KCDM is de centrale hub voor de juridische, maatschappelijke en ethische aspecten van data-gedreven applicaties en AI-toepassingen.

Hoofdstuk 2: Wat is artificiële intelligentie?

2. Wat is artificiële intelligentie?

Artificiële intelligentie verwijst volgens de Europese Commissie (EC) naar systemen die intelligent gedrag vertonen door hun omgeving te analyseren en – in zekere mate zelfstandig – actie te ondernemen om specifieke doelstellingen te verwezenlijken. Op AI-gebaseerde systemen kunnen uitsluitend uit software bestaan en actief zijn in de virtuele wereld (bijvoorbeeld stem-gestuurde assistenten, software voor beeldanalyse, zoekmachines en systemen voor spraak- en gezichtsherkenning). AI kan ook in hardware apparaten worden geïntegreerd (bijvoorbeeld autonome motorvoertuigen).

De Deskundigengroep inzake artificiële intelligentie[1] hanteert een **ruimere definitie**. Systemen op basis van AI zijn door mensen ontworpen softwaresystemen (en mogelijk ook hardware-systemen) die, met een complex doel, in de fysieke of digitale wereld in actie komen op basis van gegevens die zij in hun omgeving waarnemen, waarbij ze de verzamelde gestructureerde of ongestructureerde gegevens interpreteren, redeneren op basis van de uit deze gegevens verkregen kennis of de verkregen informatie verwerken en beslissen met welke handeling(en) het gestelde doel het best kan worden bereikt. AI-systemen kunnen gebruik maken van symbolische regels of een numeriek model leren en kunnen hun gedrag ook aanpassen door te analyseren welke invloed hun eerdere handelingen op de omgeving hebben.

Een AI-systeem kan **sterk of zwak** zijn.[2] Een sterk AI-systeem is bedoeld als een systeem dat de meeste activiteiten kan uitvoeren waartoe mensen in staat zijn. Dergelijke sterke AI-systemen bestaan vooralsnog niet. Zwakke AI-systemen daarentegen kunnen slechts één taak of een paar specifieke taken uitvoeren. Voorbeelden zijn zelfrijdende wagens of gezichtsherkenningstoepassingen.

AI als wetenschappelijke discipline kent **verschillende deelgebieden** waaronder *natural language processing*[3], experten systemen[4] en robotica.[5] Ruwweg kan een onderscheid worden gemaakt tussen een **'kennis-gebaseerde'** en een **'data-gebaseerde'** benadering van AI. De eerste benadering probeert de kennis van een menselijke expert zo goed als mogelijk in kaart te brengen door observaties en door gesprekken met de expert, en probeert die kennis vervolgens te gieten in de representaties, regels en zoekstrategieën die het gedrag van de expert benaderen. Vooral de tweede data-gebaseerde benadering of machinaal leren (ML)[6] krijgt tegenwoordig veel aandacht.[7] Hier vertrekt men van data[8] over het gedrag van mensen, over de beslissingen die ze hebben genomen of over verschijnselen die via sensoren zijn waargenomen. Vervolgens worden statistische technieken gebruikt om in die data patronen te ontdekken en die patronen worden dan weer aangewend om nieuwe problemen op te lossen.[9]

[1] Beter bekend als de High-Level Expert Group on AI of de AI HLEG.
[2] Soms worden ook de termen: generalistisch en specifieke AI-systemen of nog smalle of algemene AI-systemen.
[3] De vaardigheid om gesproken en geschreven taal te verwerken en produceren.
[4] Systemen die kennis van een bepaald gebied hebben en die kennis al redenerend op de feiten van een geval kunnen toepassen, bijvoorbeeld in een medische context.
[5] M.J. Vetzo, J.H. Gerards en R. Nehmelman, Algoritmes en grondrechten, Boom Juridisch, den Haag, 2018, p. 43.
[6] In het Engels wordt de term machine learning gehanteerd.
[7] Volgens imec verwijst AI "naar machines die dankzij een uitzonderlijk inzicht in data uit zichzelf kunnen leren, redeneren, beslissingen nemen en handelen", dus zonder dat iemand hen telkens moet vertellen wat ze moeten doen. Merk op dat in deze definitie het woord "data" opduikt. Zonder data geen AI". Zie: https://www.imec.be/nl/artikelen/wat-is-artificiele-intelligentie-en-wat-ben-ik-ermee.
[8] Gegevens(set) en data(set) worden in deze gids als synoniemen gebruikt.
[9] L. Steels, "Artificiële intelligentie. Naar een vierde industriële revolutie?", Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten, 2017, p. 14-17.

ML is een subcategorie of een type van artificiële intelligentie. ML is gebaseerd op algoritmes die in staat zijn om te leren op basis van eerdere ervaringen, zogenaamde zelflerende algoritmes. Het geeft computers de mogelijkheid om te leren zonder hiertoe expliciet geprogrammeerd te zijn.[10] Hoe meer data deze systemen of tools verwerken, hoe beter de algoritmen in deze systemen patronen zullen ontdekken in de verzamelde data. Ze doen dit zelfstandig, zonder instructies, maar met behulp van voorbeelden of suggesties. Belangrijke elementen zijn volgens het recente Witboek over AI, waarin de strategie van de EC rond AI wordt uitgestippeld, **data** en **algoritmes**. Data is elke vorm van informatie die door een computer verwerkt kan worden. Dit kan variëren van enkele minimale gegevens tot meerdere gegevensverzamelingen (datasets). Het verwerken van een enorme hoeveelheid aan data is gekend als **big data**. Een algoritme is een opeenvolging van regels en instructies die een vooraf bepaald doel bereiken. Een algoritme leest, doorzoekt en sorteert data om kennis te creëren.[11]

Dieplerende systemen (DL)[12] zijn vandaag de meest voorkomende en een geavanceerde vorm van ML. DL maakt gebruik van 'artificiële neurale netwerken'. Dat zijn netwerken van digitale neuronen, geïnspireerd door het menselijke brein. Het dieplerende algoritme voert hierbij een gelaagde analyse uit, waarbij resultaten uit de ene laag worden gebruikt als input voor de analyse van een volgende laag. Zo kunnen complexe, verborgen verbanden in grote datasets worden ontdekt.[13] Een netwerk dat bijvoorbeeld een verkeersbord moet herkennen, zal zich focussen op vormen, kleuren en formaten. Een eerste laag kan op zoek gaan naar een omgekeerde driehoek, een tweede naar fel rood, en een derde naar wit. Iedere laag zal aangeven of het haar specifiek item heeft gevonden en hoe zeker ze daarvan is. Zo kan een neuraal netwerk voor beeldherkenning naar een foto kijken en een voorrangsbord herkennen.[14]

AI-systemen gebruiken dus in regel **grote hoeveelheden data** die deze systemen de mogelijkheid geven om te leren en intelligent te worden. Het gaat daarbij niet noodzakelijk om persoonsgegevens, denk bijvoorbeeld aan meteorologische of financiële gegevens die niet gekoppeld zijn aan personen. Maar, indien het AI-systeem **persoonsgegevens** gebruikt, is zoals reeds aangehaald de AVG van toepassing. De AVG zal zowel bij het design, de ontwikkeling als de uitrol en het gebruik van AI-systemen moeten worden nageleefd.

[10] M.J. Vetzo, J.H. Gerards en R. Nehmelman, Algoritmes en grondrechten, Boom Juridisch, den Haag, 2018, p. 43.
[11] Zie voor meer informatie over deze begrippen: M.J. Vetzo, J.H. Gerards en R. Nehmelman, Algoritmes en grondrechten, Boom Juridisch, den Haag, 2018, 244 p.
[12] In het Engels wordt de term deep learning gehanteerd.
[13] M.J. Vetzo, J.H. Gerards en R. Nehmelman, "Algoritmes en grondrechten", Boom Juridisch, den Haag, 2018, p. 43.
[14] Het voorbeeld werd aangehaald op: https://www.techzine.be/blogs/trends/25516/ai-machine-learning-en-deep-learning-wat-is-het-verschil.

Toepassingen in deze gids

Er zijn talrijke toepassingen van AI-systemen waarop de in deze gids besproken bepalingen over gegevensbescherming van toepassing zijn. Het is niet mogelijk om al die AI-systemen uitvoerig te bespreken. Van belang is om volgende vragen bij het ontwerp, de ontwikkeling en het gebruik van een AI-systeem te onthouden. Antwoorden hierop zullen bepalen of en welke vereisten uit de AVG van belang zijn:

- Doel: waarvoor wordt de applicatie gebruikt?
- Kanalen: via welke kanalen wordt data verzameld en worden personen benaderd?
- Welke data: welke gegevens worden verwerkt door het AI-systeem in de verschillende fases?

In deze gids worden twee gevalstudies gebruikt om de toepassing van de AVG op AI-systemen waar nodig te verduidelijken en illustreren.

AI in e-commerce

Een eerste toepassing is het gebruik van een AI-gedreven e-commerce verkoopprogramma. Bekende voorbeelden hiervan zijn Amazon Webshop, Alibaba of Collect&Go van Colruyt. Er zijn ook heel wat online platformen die gebruik maken van *cognitive computing technology*. Cognitieve computers leren uit de data die hen aangereikt wordt, zowel uit gestructureerde bronnen (zoals documentatie, handleidingen, specificaties) als ongestructureerde bronnen (zoals blogs, recensies, sociale media). Ze proberen op basis van deze (big) data de context te begrijpen. De belangrijkste karakteristieken van cognitieve systemen zijn dat ze begrijpen, leren, en op een voor ons natuurlijke manier met de mens redeneren en interacteren.[15] Denk bijvoorbeeld aan de Expert Personal Shopper (XPS), een platform en internet bot die online gesprekken voert met mensen om te weten te komen wat ze willen kopen en om hen daar vervolgens mee te helpen.

1. Het doel van een dergelijk AI-systeem is om de klantenervaringen en de verkoop te bevorderen door personalisatie van bijvoorbeeld:
   a. het aanraden van gepaste artikels aan bezoekers van een webshop;
   b. het aanraden van gepaste artikels bij andere artikels die reeds in het winkelmandje zijn opgenomen;
   c. het geven van gerichte kortingen aan klanten;
   d. het creëren van zo genaamde leadgeneratie (*lead generation*)[16], hetzij het aantrekken van potentiële klanten door middel van voorschotelen relevante inhoud ('content').

[15] R. Nijman, "Cognitive Computing en IBM Watson – Wat is het en wat biedt het de overheid?", https://www.ibm.com/blogs/think/nl-en/2015/01/12/cognitive-computing-en-ibm-watson-wat-is-het-en-wat-biedt-het-de-overheid/. Zie ook: https://www2.cio.nl/development/85006-wat-is-cognitive-computing.

[16] Leadgeneratie is het proces om potentiële klanten te identificeren en hun interesse te wekken voor de producten of diensten van een organisatie.

2. Er kunnen verschillende kanalen worden gebruikt bij dergelijke AI-gedreven e-commerce verkoopprogramma's zoals:
   a. webshops met inbegrip van het gebruik van pop-ups, chatbots of productmeldingen;
   b. e-mail correspondentie, vooral met bestaande klanten;
   c. websites en applicaties van derden-dienstverleners zoals Google Ads of sociale media zoals Facebook, Twitter of Instagram.
3. Er kunnen verschillende persoonsgegevens worden gebruikt en verwerkt door een AI-gedreven e-commerce verkoopprogramma waaronder:
   a. eerdere aankopen van de gebruiker;
   b. gedrag van de gebruiker op website (bijvoorbeeld klikken, terugkeren, bepaalde items opnieuw bekijken,…);
   c. online te vinden informatie van gebruiker (via Google bijvoorbeeld);
   d. informatie ingegeven door de gebruiker in de webapplicatie;
   e. locatiegegevens;
   f. vergelijking met gegevens van 'gelijkaardige' gebruikers;
   g. vergelijking met gelijkaardige gecombineerde aankopen/views (bijvoorbeeld andere klanten die X kochten/bekeken, kochten/bekeken ook Y);
   h. biometrische gegevens, zoals gezichtsherkenning toegepast op een profielfoto.

AI in rekrutering

Een tweede toepassing is het gebruik van AI-gedreven systemen bij aanwervingen en rekrutering (*recruitment*). Er zijn hiervan talrijke applicaties. Skeeled is bijvoorbeeld op AI-gebaseerde recruitment software die in verschillende delen van het wervingsproces wordt gebruikt zoals bij de pre-screening, bij het maken van de ranglijst van sollicitanten of bij het geven van feedback aan de rekruteerder. AI kan ook helpen om onduidelijke functiebeschrijvingen te verduidelijken voordat ze worden gepubliceerd. De VDAB maakt sinds eind 2018 bijvoorbeeld gebruik van AI – Jobnet – om de resultaten van hun automatisch matchingsysteem te verbeteren en optimaliseren. Toch kan het ook soms misgaan door het bestaan van bias.[17] Het algoritme dat door Amazon werd ontwikkeld voor het scannen van sollicitatiebrieven bleek bijvoorbeeld vrouwen te benadelen.

1. Het doel van het gebruik van een AI-systeem is om de juiste profielen aan te werven en dus het proces efficiënter te maken door bijvoorbeeld:
   a. vacatures online te laten zien aan de geschikte personen;
   b. uit de kandidaturen een ranking en/of selectie te maken en vervolgens een geschikte kandidaat te kiezen;
   c. feedback te geven aan de mensen betrokken in het wervingsproces.

[17] Een bias verwijst naar onbewust redeneerfouten/inherente vooroordelen ingebouwd in het AI-systeem.

2. Er kunnen verschillende kanalen worden gebruikt om gegevens van een kandidaat te verzamelen te verzamelen waaronder:
    a. rechtstreekse kanalen zoals online applicatiepagina e-mails;
    b. sociale media platformen (bijvoorbeeld Facebook, Twitter, LinkedIn en Instagram).
3. Er kunnen verschillende persoonsgegevens worden gebruikt en verwerkt door AI-gedreven rekruteringssystemen waaronder:
    a. informatie meegedeeld door een kandidaat in een CV of motivatiebrief zoals de identiteit, hobby's, foto's, eerdere werkervaringen en opleiding;
    b. informatie uit sociale media profielen zoals aantal en inhoud van posts, vrienden, profiel, sociale en geografische omgeving of foto's;
    c. andere persoonlijke online informatie zoals een e-commerce profiel;
    d. andere informatie te vinden bij derde partijen.

Hoofdstuk 3: Toepassingsgebied AVG

3. Op welke activiteiten is de AVG van toepassing en welke rollen kan een organisatie vervullen onder de AVG?

Essentie

Het is uitermate belangrijk om te weten welke soort gegevens (zullen) worden gebruikt in een bepaald AI-systeem. Indien het systeem persoonsgegevens verwerkt, dan moet de AVG worden nageleefd.

Het begrip 'verwerken' is heel breed. Persoonsgegevens worden verwerkt zodra er iets met deze gegevens gedaan wordt of zelfs zodra de gegevens door een door de organisatie gecontroleerde omgeving passeren, ook al is er geen effectieve toegang en doet de organisatie verder niets met de persoonsgegevens.[18]

Persoonsgegevens zijn zowel gegevens die toelaten om een natuurlijk persoon te identificeren, als gegevens die betrekking hebben op een geïdentificeerde of identificeerbare persoon. Personen kunnen geïdentificeerd worden aan de hand van een naam of adres ('rechtstreekse identificatie'), maar het kan ook gaan om IP-adressen, cookie identificatoren of andere factoren ('onrechtstreekse identificatie'). Als een persoon niet onmiddellijk geïdentificeerd kan worden, moet nagegaan worden of identificatie mogelijk is of niet.

Pseudonieme gegevens vallen onder de toepassing van de AVG. De AVG is niet van toepassing op anonieme gegevens. Hier moet wel worden nagegaan of geen her-identificatie mogelijk is. Het probleem is dat big data de mogelijkheid van her-identificatie faciliteert door de combinatie van verschillende gegevenssets. Anonimiseren van persoonsgegevens is dus niet altijd permanent en niet elke anonimiseringsmethode is een geschikte methode voor de bescherming van gegevens. Gegevens circuleren immers op het internet, worden verhandeld, nieuwe sets van gegevens worden aangemaakt en derde partijen kunnen in het bezit zijn van informatie die het linken van gegevens toelaat waarvan de originele verwerkingsverantwoordelijke geen weet heeft. Dit betekent dat het steeds moeilijker zal worden om duidelijke grenzen te trekken tussen persoonsgegevens en niet-persoonsgebonden gegevens.

AI-systemen kunnen daarenboven gebruik maken van 'gemengde gegevenssets', die zowel persoonsgebonden als niet-persoonsgebonden gegevens bevatten. Dit betekent voor deze soort gegevenssets dat: (i) niet-persoonsgebonden gegevens onder de Verordening inzake het vrije verkeer van niet-persoonsgebonden gegevens vallen en (ii) persoonsgegevens onder de AVG vallen. Indien de beide sets van gegevens 'onlosmakelijk met elkaar verbonden zijn', zal de AVG van toepassing zijn op de volledige gegevensset, ook indien persoonsgegevens slechts een klein deel van de set uitmaken. De kans is groot dat een AI-systeem gebruik maakt van gemengde datasets die 'onlosmakelijk' met elkaar verbonden zijn en dat de AVG dus van toepassing zal zijn.

[18] Art. 4, 2) AVG.

Een van de belangrijkste aspecten onder de AVG is het bepalen van de verschillende rollen en verantwoordelijkheden met betrekking tot de verwerking van persoonsgegevens. Het is daarom uitermate belangrijk om te bepalen of een organisatie een verwerkingsverantwoordelijke, dan wel een verwerker is, aangezien dit een fundamentele impact heeft op de na te leven verplichtingen.

Actiepunten

- ◊ Maak een inventaris van de gegevens die het AI-systeem zal gebruiken.
- ◊ Ga na welke gegevens persoonsgegevens zijn en welke niet.
- ◊ Indien het om persoonsgegevens gaat, bepaal of werkelijk al deze gegevens nodig zijn voor de ontwikkeling en het functioneren van het AI-systeem.
- ◊ Ga na of de persoonsgegevens die gebruikt werden in de trainingsfase, geanonimiseerd of minstens gepseudonimiseerd kunnen worden voor de operationele fase, indien ze daar nog gebruikt worden.
- ◊ Bekijk of er een mogelijkheid is om bepaalde/alle gegevens te anonimiseren/pseudonimiseren zonder een grote (functionele of technische) impact op het AI-systeem.
- ◊ Analyseer hoe groot de kans op her-identificatie is indien er anonieme gegevens worden gebruikt.
- ◊ Ga na of gemengde datasets uit elkaar kunnen worden gehaald of 'onlosmakelijk' met elkaar verbonden zijn.

3.1. Wat zijn persoonsgegevens?

De AVG is van toepassing op elke **verwerking van persoonsgegevens**.[19] De definitie van verwerking is heel ruim opgevat[20] en heeft tot gevolg dat de AVG bij vrijwel iedere verrichting met persoonsgegevens van toepassing zal zijn.

Het begrip 'persoonsgegevens' verwijst naar alle i**nformatie over een identificeerbare levende natuurlijke persoon**. Voorbeelden van persoonsgegevens zijn:

- een naam en achternaam;
- thuis- of afleveradres;
- e-mailadres zoals naam.achternaam@onderneming.be;
- nummer van identiteitskaart;
- locatiegegevens (bijvoorbeeld de locatiegegevens op een mobiele telefoon);
- internetprotocoladres (IP-adres);
- identificatiecookie[21];

[19] Art. 2 AVG.

[20] Art. 4 2) AVG omschrijft verwerking als "een bewerking of een geheel van bewerkingen met betrekking tot persoonsgegevens of een geheel van persoonsgegevens, al dan niet uitgevoerd via geautomatiseerde procedés, zoals het verzamelen, vastleggen, ordenen, structureren, opslaan, bijwerken of wijzigen, opvragen, raadplegen, gebruiken, verstrekken door middel van doorzending, verspreiden of op andere wijze ter beschikking stellen, aligneren of combineren, afschermen, wissen of vernietigen van gegevens".

[21] Het gebruik van cookies of gelijkaardige technologieën om een persoon te traceren op verschillende websites houdt bijvoorbeeld in dat persoonsgegevens worden verwerkt als deze tracering gepaard gaat met online-identificatiemiddelen die worden gebruikt om een profiel van het individu aan te maken. Zo is een sociale media 'handle' of de gebruikersnaam van een individu, die anoniem of onzinnig lijkt, nog steeds voldoende om hen te identificeren omdat het die persoon op een unieke manier identificeert. De gebruikersnaam is een persoons gegeven als deze het ene individu van het andere onderscheidt, ongeacht of het mogelijk is om de 'online' identiteit te koppelen aan een 'echte wereld' persoon.

- advertentie-ID van telefoon;
- gegevens in het bezit van een zorg-of dienstverlener, bijvoorbeeld in de vorm van een symbool dat iemand een unieke identiteit geeft;
- een of meer elementen die kenmerkend zijn voor de fysieke, fysiologische, genetische, psychische, economische, culturele of sociale identiteit van die natuurlijke persoon;
- letterlijk elke andere informatie die aan een natuurlijke persoon gekoppeld kan worden.

Het is niet altijd eenvoudig om te bepalen of een gegeven een persoonsgegeven is. Het is in ieder geval aangeraden **voorzichtig met persoonsgegevens om te gaan** en te verzekeren dat er een duidelijke reden is om deze gegevens te verwerken.

Persoonsgegevens bestaan dus eigenlijk uit een **aantal belangrijke bouwstenen**:[22]

Alle informatie
Het gaat ten eerste over alle informatie, wat erop wijst dat persoonsgegevens een ruim concept is. Wat de **aard** van de informatie betreft, omvat het gegevens van alle soorten over een bepaalde persoon. Deze informatie kan objectief zijn zoals iemands bloedtype, of subjectief zijn zoals opinies of oordelen. Wat de **inhoud** van de informatie betreft, kan het gaan over gegevens die informatie verstrekken van welke aard dan ook, dat kunnen zelfs technische gegevens zijn. Persoonsgegevens kunnen verder betrekking hebben op het privé- en familieleven van een persoon, maar ook op de activiteiten die een persoon onderneemt zoals in zijn werk, vrije tijd of als consument. De **vorm** van het medium waarop de informatie is opgeslagen, kan in iedere vorm zijn zoals alfabetisch, numeriek of grafisch.
Informatie over een persoon
Ten tweede moet de informatie een natuurlijk persoon betreffen. Men kan er vanuit gaan dat informatie betrekking heeft op een persoon wanneer het **over die persoon gaat**. In veel situaties kan deze relatie gemakkelijk worden vastgesteld. Zo zijn bijvoorbeeld de gegevens in een individueel personeelsdossier duidelijk 'gerelateerd' aan de situatie van de persoon als werknemer. Daarbij zal elke informatie die toelaat om een persoon rechtstreeks of onrechtstreeks te identificeren in elk geval als een persoonsgegeven beschouwd moeten worden. Toch zijn er ook situaties waarin het vaststellen of gegevens betrekking hebben op een persoon **niet vanzelfsprekend** is. In sommige situaties heeft de informatie immers betrekking op objecten zoals een auto of een huis. Die objecten zijn meestal iemands eigendom, staan onder iemands beheer of oefenen invloed uit over een persoon. Dan kan enkel indirect worden aangenomen dat de informatie betrekking heeft op een bepaalde persoon, door de koppeling aan andere gegevens die identificatie toestaan, maar ook hier gaat het over persoonsgegevens.

[22] De Groep Gegevensbescherming Artikel 29 is de onafhankelijke Europese werkgroep die tot 25 mei 2018 verantwoordelijk was voor de behandeling van kwesties in verband met de bescherming van de persoonlijke levenssfeer en van persoonsgegevens. De Groep werd vervangen door het Europees Comité voor gegevensbescherming (European Data Protection Board). Zie voor meer informatie: https://edpb.europa.eu/our-work-tools/article-29-working-party_nl).

Om na te gaan of gegevens betrekking hebben op een persoon, zijn volgende elementen van belang:

- **inhoud** van de gegevens (zijn ze rechtstreeks verbonden met een persoon of zijn activiteiten?);
- **doel** waarvoor de gegevens verwerkt worden;
- **resultaten** of de **gevolgen** voor de persoon doordat de gegevens verwerkt worden.

Geïdentificeerde of identificeerbare levende persoon

Ten derde moet het gaan over een geïdentificeerde of identificeerbare levende persoon. Over het algemeen kan een persoon als geïdentificeerd worden beschouwd wanneer die **duidelijk** van andere leden van een groep **onderscheiden** kan worden. Een persoon kan geïdentificeerd worden door middel van **identificatoren**. Voorbeelden zijn uiterlijke kenmerken of een kwaliteit van een persoon die niet onmiddellijk kan worden waargenomen zoals een naam, functie of beroep. Identificeerbaar impliceert dat de mogelijkheid bestaat dat de persoon onderscheiden kan worden. Identificatie is ook op indirecte wijze mogelijk. Daarbij gaat het doorgaans om een klein of groot aantal 'unieke combinaties'.

In gevallen waarin het op het eerste zicht niet mogelijk is om met de beschikbare identificatiemiddelen één bepaalde persoon te onderscheiden, kan die persoon wellicht toch identificeerbaar zijn doordat aan de hand van de **combinatie van die informatie** met andere gegevens (die al dan niet ter beschikking zijn van de verwerkingsverantwoordelijke) de betrokkene van andere personen kan worden onderscheiden. Een typevoorbeeld is informatie met betrekking tot voorwerpen. Voorwerpen zijn namelijk doorgaans iemands eigendom, staan onder het beheer van of oefenen invloed uit op een persoon, of staan in een bepaalde fysieke of geografische nabijheidsrelatie tot personen of andere voorwerpen. De informatie kan echter leiden tot een persoon en kan in dergelijke gevallen slechts indirect geacht worden die personen te identificeren.

3.2. Wat zijn bijzondere categorieën van persoonsgegevens?

Er zijn een aantal bijzondere categorieën van **gegevens die in principe niet verwerkt mogen worden** tenzij daarvoor een **uitzonderingsgrond** is. Andere 'gewone' gegevens mogen in principe wel verwerkt worden, mits de verwerking conform de AVG gebeurt en er dus onder meer een rechtmatige verwerkingsgrond bestaat.

Deze 'gevoelige' categorieën van gegevens zijn: persoonsgegevens waaruit ras of etnische afkomst, politieke opvattingen, religieuze of levensbeschouwelijke overtuigingen, het lidmaatschap van een vakbond of seksueel gedrag of seksuele gerichtheid blijken; genetische gegevens en biometrische gegevens met het oog op de unieke identificatie van een persoon; en gegevens over gezondheid.[23]

Het verbod om dergelijke gegevens te verwerken geldt echter niet in een aantal wettelijk opgesomde gevallen.[24] Indien een organisatie deze persoonsgegevens in deze gevallen zou willen verwerken, moet ze sowieso alle andere principes en beginselen van de AVG in acht nemen, evenals de eventuele specifieke regels op de verwerking van die gegevens.

[23] Art. 9 en 10 AVG.
[24] Art. 9 AVG.

In een **AI-context** kunnen volgende **uitzonderingsgevallen** mogelijks relevant zijn:

1.	Als er **expliciete toestemming** door de betrokkene werd gegeven voor de verwerking van die gegevens voor een of meer welbepaalde doeleinden.
2.	Als de gegevens door de betrokkene **openbaar** zijn gemaakt.
3.	Als de verwerking noodzakelijk is met het oog op **wetenschappelijk onderzoek of statistische doeleinden** op basis van een wettelijke bepaling.
4.	Als de verwerking van die gegevens noodzakelijk is met oog op de uitvoering van verplichtingen en de uitoefening van rechten van de verwerkingsverantwoordelijke of de betrokkene op het gebied van het **arbeidsrecht, socialezekerheidsrecht en socialebeschermingsrecht**.
5.	Als de verwerking noodzakelijk is voor doeleinden van **preventieve of arbeidsgeneeskunde** of voor de beoordeling van de **arbeidsgeschiktheid van de werknemer**.

Gezien de verwerking van deze soort persoonsgegevens vaak een verhoogd risico met zich meebrengt, zal waarschijnlijk een **Gegevensbeschermingseffectbeoordeling** (GEB) of ***Data Protection Impact Assessment*** (DPIA) moeten worden uitgevoerd.[25]

3.3. Wat is het belang van het onderscheid tussen anonimisering en pseudonimisering?

Het onderscheid tussen anonimisering en pseudonimisering van gegevens is van belang in de context van gegevensbescherming. De AVG gebruikt de term pseudonimisering om te verwijzen naar gecodeerde gegevens die niet meer aan een specifieke natuurlijke persoon kunnen worden gekoppeld zonder aanvullende gegevens die als sleutel dienen. De aanvullende gegevens om de persoonsgegevens aan een specifieke persoon te koppelen worden apart bewaard.[26]

Toepassing
Stel dat iemand solliciteert voor een bepaalde functie. Het HR-departement heeft een AI-systeem laten ontwikkelen dat het sollicitatiedossier opsplitst in twee 'folders'. In de eerste stap wordt de eerste pagina verwijderd met onder andere de naam en contactgegevens en de rest van het document wordt in 'Folder 1' bewaard. Dit document krijgt in een tweede stap een automatisch gegenereerd nummer en wordt vervolgens doorgestuurd naar een rekruteerder. Het HR-departement bewaart de eerste pagina van de sollicitatie met naam en contactgegevens samen met het automatisch gegenereerd nummer in 'Folder 2'. Op zich laat de informatie in 'Folder 1' geen identificatie meer toe, maar gecombineerd met de informatie in 'Folder 2' kan de sollicitant geïdentificeerd worden.

[25] We hanteren in deze gids DPIA in plaats van het Nederlandse GEB omdat dit een meer gangbare term is. Zie ook "4.4. Wanneer moet een DPIA of GEB worden uitgevoerd voor een verwerking van persoonsgegevens door AI-systemen?"
[26] Overweging 29 AVG.

Pseudonimisering is dus geen methode van anonimisering, maar v**ermindert de mogelijkheid dat gegevens aan de betrokken persoon gekoppeld worden**. Gepseudonimiseerde persoonsgegevens waarvoor een sleutel bestaat om de oorspronkelijke persoonsgegevens opnieuw te verkrijgen, blijven persoonsgegevens en zijn dus aan de verplichtingen van de AVG onderworpen.

De AVG verbindt een **aantal voordelen** aan pseudonimisering die ook voor AI-systemen nuttig kunnen zijn:

1.	Er is meer ruimte om de gegevens voor een ander doel te verwerken dan waarvoor ze werden verzameld.[27]
2.	Deze techniek kan als een technische en organisatorische maatregel helpen om het principe van doelbinding en de verplichtingen van gegevensbescherming door ontwerp en gegevensbescherming door standaardinstellingen.[28]
3.	Het draagt bij om aan de vereisten van data veiligheid te voldoen.[29]
4.	Het is een belangrijke waarborg in verband met de verwerking met het oog op archivering in het algemeen belang, wetenschappelijk of historisch onderzoek of statistische doeleinden.[30]

Anonimiseren wordt door de AVG niet gedefinieerd, maar komt erop neer dat de **natuurlijke persoon op wie de gegevens betrekking hebben niet of niet langer kan worden geïdentificeerd**. De gegevens zijn pas echt geanonimiseerd als de anonimisering onomkeerbaar is. Is een organisatie in staat tot een onomkeerbare anonimisering, dan is de AVG niet (meer) van toepassing.[31]

De **identificeerbaarheid** is dus het criterium om te beoordelen of gegevens pseudoniem of anoniem zijn. Daarbij moet rekening worden gehouden met alle objectieve factoren, zoals de kosten van en de tijd benodigd voor identificatie. De beschikbare technologie op het tijdstip van verwerking en de technologische ontwikkelingen moeten daarbij in acht worden genomen.[32]

Ondanks anonimisering kan er **toch nog een kans op her-identificatie** bestaan. Dit is het proces van het opnieuw omzetten van geanonimiseerde gegevens in persoonsgegevens door het gebruik van datamatching of gelijkaardige technieken. Deze technieken maken vaak gebruik van een vorm van ML wat betekent dat er voor sommige van deze toepassingen een risico op her-identificatie bestaat. Zo hebben onderzoekers in 2019 een model ontwikkeld waarmee ze 99.98% van de Amerikanen correct kunnen re-identificeren in elke dataset, met behulp van 15 demografische kenmerken.[33]

[27] Art. 6.4 AVG.
[28] Art. 25 AVG. Zie ook "4.1. Wat betekent gegevensbescherming door ontwerp en hoe kan dit geïmplementeerd worden in AI-systemen?"
[29] Art. 32, 33 en 34 AVG. Zie ook "4.3. Wat zijn aandachtspunten voor de beveiliging van de verwerking van persoonsgegevens door AI-systemen?"
[30] Art. 89.1 AVG. Zie ook deel "4.5. Welke verplichtingen gelden er wanneer persoonsgegevens worden verwerkt met het oog op wetenschappelijk onderzoek of statistische doeleinden?"
[31] Overweging 26 AVG.
[32] Overweging 26 AVG.
[33] L. Rocher, J.M. Hendrickx en Y. de Montjoye, "Estimating the success of re-identifications in incomplete datasets using generative mod-els", Nature Communication, 2019, vol. 10, nr. 3069, https://doi.org/10.1038/s41467-019-10933-3.

In dit verband stelt de WP29 dat er **drie criteria** zijn die in overweging genomen moeten worden om te bepalen of her-identificatie kan plaatsvinden, namelijk:

- **herleidbaarheid:** de mogelijkheid om een persoon te individualiseren;
- **koppelbaarheid:** de mogelijkheid om ten minste twee gegevens met betrekking tot dezelfde persoon of groep personen met elkaar in verband te brengen;
- **deduceerbaarheid:** de mogelijkheid om persoonsgebonden informatie af te leiden.

Volgens de WP29 is een anonimiseringsoplossing die deze drie risico's uitsluit, voldoende bestand tegen het risico op her-identificatie. Ze heeft zelf echter al aangegeven dat het bereiken van deze drempel zeer moeilijk is. Elke methode heeft immers minstens een klein risico op her-identificatie. Enkel **een combinatie van verschillende technieken** zou het mogelijk maken om persoonsgegevens volledig te anonimiseren.

Wat met big data?	Een probleem is dat big data (zeker in combinatie met de rekenkracht van AI-systemen) de mogelijkheid tot her-identificatie vergroot door de mogelijke **combinatie van verschillende gegevenssets**. Anonimiseren van persoonsgegevens is dus niet altijd permanent en misschien niet langer een geschikte methode voor de bescherming van gegevens.[34] Gegevens circuleren immers op het internet, worden verhandeld, worden geïntegreerd in nieuwe sets van gegevens,… Derde partijen kunnen daarenboven in het bezit zijn van informatie die het linken van gegevens toelaat en waarvan de originele verwerkingsverantwoordelijke geen weet heeft en wat de deur openzet voor een mogelijke aansprakelijkheid. Dit betekent dat het **steeds moeilijker** zal worden om duidelijke **grenzen** te trekken tussen **persoonsgegevens en niet-persoonsgebonden gegevens**.

De vraag die zich stelt is dan ook of het in bepaalde omstandigheden niet beter kan zijn om **gegevens te pseudonimiseren** dan te streven naar een haast onmogelijke anonimisering van gegevens. **Anonimisering** heeft volgens sommigen immers potentieel **enkele nadelen**:

- verminderde mogelijkheid om gegevens aan personen terug te koppelen, waardoor het niet altijd geweten is of de gegevens in eerste instantie in overeenstemming met de AVG verzameld werden;
- verminderd zicht op de herkomst van de gegevens, transformaties en de bewegingen doorheen de tijd, waardoor het moeilijker is een verantwoord beleid rond gegevens uit te bouwen;
- Door de combinaties van verschillende gegevenssets, bestaat er een risico op her-identificatie, wat een verhoogde aansprakelijkheid met zich kan meebrengen;
- De accuraatheid van de gegevens kan verminderen afhankelijk van de toegepaste anonimiseringstechnieken. Daardoor kan de gegevensset minder bruikbaar worden.[35]

[34] Zie in dit verband: G. LaFever, "Anonymisation does not work for big data due to lack of protection for direct & indirect identifiers and easy re-identification vs pseudonymization", gdpr.report/news/2019/08/12/anonymisation-does-not-work-for-big-data-due-to-lack-of-protection-for-direct-indirect-identifiers-and-easy-re-identification-vs-pseudonymisation.

[35] Zie voor de praktische toepassing van anonimisering en de technieken die daarvoor gebruikt kunnen worden deel 4.2. over minimale gegevensverwerking.

3.4. Wat met niet-persoonsgebonden gegevens en/of gemengde gegevenssets?

Zoals reeds aangehaald gebruiken AI-systemen niet noodzakelijk persoonsgegevens. Ze kunnen ook niet-persoonsgebonden gegevens gebruiken. Niet-persoonsgebonden gegevens worden gedefinieerd als **andere gegevens dan persoonsgegevens**.[36]

Het gaat over gegevens die geen betrekking (meer) hebben op een geïdentificeerde of identificeerbare natuurlijke persoon zoals gegevens over weersomstandigheden, voor zover deze niet gekoppeld worden aan een persoon uiteraard. Het kan dus ook gaan over persoonsgegevens die werden geanonimiseerd. Bij de beoordeling of de gegevens naar behoren zijn geanonimiseerd, moet rekening worden gehouden met de specifieke en unieke omstandigheden van elk afzonderlijk geval.

In realiteit kan het soms **moeilijk** zijn om een **duidelijke scheidingslijn** te trekken tussen persoonsgegevens (en dus toepassing van de AVG) en niet-persoonsgebonden gegevens (en dus toepassing van de regels inzake niet-persoonsgebonden gegevens). AI-systemen kunnen daarenboven gebruik maken van **'gemengde gegevenssets'**.

Toepassing
Sommige webshops kunnen gebruik maken van door derden geleverde diensten in het kader van klantrelatiebeheer (*customer relationship management* – "CRM"). AI kan daarbij worden gebruikt om de output en effectiviteit van CRM tools te verbeteren.[37] De gegevens van een klant moeten beschikbaar worden gesteld in de CRM-omgeving. Tot de gegevens die nodig zijn voor de CRM-dienst behoort alle informatie die nodig is om de interactie met de klanten doeltreffend te beheren. Dit zijn bijvoorbeeld hun postadres, e-mailadres, telefoonnummer. Het kan echter ook gaan over producten en diensten die zij kopen, alsook verkoopverslagen met inbegrip van geaggregeerde gegevens, wat niet-persoonsgebonden gegevens zijn. De gegevens in de CRM-omgeving kunnen dus zowel persoonsgegevens als niet-persoonsgebonden gegevens omvatten.

De Europese Commissie heeft in dit kader enkele richtsnoeren gepubliceerd over de wisselwerking tussen gegevenssets die bestaan uit zowel persoonsgegevens als niet-persoonsgebonden gegevens.

Wanneer een gegevensset **zowel uit persoonsgegevens als niet-persoonsgebonden gegevens bestaat**, betekent dit dat:

- niet-persoonsgebonden gegevens onder de Verordening inzake het vrije verkeer van niet-persoonsgebonden gegevens vallen;
- persoonsgegevens onder de AVG vallen.

Indien de beide sets van gegevens **'onlosmakelijk met elkaar verbonden zijn'**, zal de AVG van toepassing zijn op de volledige gegevensset, ook indien persoonsgegevens slechts een klein deel van de set uitmaken.

[36] Art. 3.1 Verordening 2018/1807 van het Europees Parlement en de Raad van 14 november 2018, inzake een kader voor het vrije verkeer van niet-persoonsgebonden gegevens in de Europese Unie, PE/53/2018/REV/1, OJ L 303, 28.11.2018, p. 59-68.

[37] Zie bijvoorbeeld: Commercient, "Why Artificial Intelligence Integration in CRM is the Future for your Businesss", https://www.commercient.com/why-artificial-intelligence-integration-in-crm-is-the-future-for-your-business.

Het begrip 'onlosmakelijk met elkaar verbonden' is niet gedefinieerd. Het kan verwijzen naar een situatie waarin een gegevensset zowel persoonsgegevens als niet-persoonsgebonden gegevens bevat en het scheiden van deze gegevens ofwel:

- onmogelijk is;
- economisch inefficiënt; of
- technisch onhaalbaar wordt geacht door de verwerkingsverantwoordelijke.

De kans is groot dat **AI-systemen gebruik maken van gemengde datasets die 'onlosmakelijk' met elkaar verbonden**, waardoor de AVG dus van toepassing zal zijn.

Toepassing
Zo is bijvoorbeeld het gegeven 'behaalde Masterdiploma' op zichzelf een niet-persoonsgebonden gegeven wanneer niet meer uit de gegevensset volgt op wie dit gegeven betrekking had. In een AI-systeem voor rekruteringsdoeleinden is het echter noodzakelijk dat het 'behaalde diploma'en de 'sollicitant' correct aan elkaar gekoppeld zijn. Deze uit elkaar halen is dus niet altijd mogelijk waardoor ze 'onlosmakelijk' met elkaar verbonden zijn en de AVG dus van toepassing is.

3.5. Wat zijn de verschillende rollen die een organisatie kan vervullen onder de GDPR in een AI-context?

Een van de belangrijkste aspecten onder de AVG is het bepalen van de verschillende rollen en verantwoordelijkheden met betrekking tot de verwerking van persoonsgegevens. Het onderscheid tussen **verwerkingsverantwoordelijke** en **verwerker** is van belang omdat ze elk andere verplichtingen hebben onder de AVG. Slechts als deze hoedanigheid vaststaat, weet men dus wie of welke organisatie instaat voor bijvoorbeeld de transparantie- of verantwoordingsplichten.

In de verschillende stadia van de levenscyclus van een AI-systeem is de **verwerkingsverantwoordelijke** de natuurlijke of rechtspersoon, overheidsinstantie of andere organisatie die de beslissing neemt omtrent het doel van en de middelen voor de verwerking van persoonsgegevens.

Een verwerkingsverantwoordelijke kan verschillende taken aan derden uitbesteden, die deze taken ten behoeve van de verwerkingsverantwoordelijke en conform diens instructies uitvoeren. Deze derden worden **verwerkers** genoemd. Indien zij deze persoonsgegevens verder verwerken of bijkomende verwerkingen uitvoeren, worden zij de verwerkingsverantwoordelijken voor die verwerkingen.

Hieronder lichten we per fase/activiteit in de levensloop van een AI-systeem kort toe wie als verwerkingsverantwoordelijke en wie als verwerker moet worden beschouwd. De tabel licht enkele veelvoorkomende gevallen toe en kan als richtlijn dienen voor nieuwe situaties.

Fase/Activiteit	Wie is verwerkingsverantwoordelijke?	Wie is verwerker?
Ontwikkeling/ Training/Validatie	De organisatie die het AI-systeem (door)ontwikkelt, traint of valideert en beslist welke persoonsgegevens worden gebruikt om het systeem te trainen (en dus het doel en de middelen bepaalt). Ook indien deze organisatie een set persoonsgegevens verkrijgt, heeft deze organisatie de status van verwerkingsverantwoordelijke. Indien de ontwikkeling, training, validatie of (door)ontwikkeling aan een derde organisatie wordt uitbesteed en deze derde organisatie beslist welk type persoonsgegevens hiervoor wordt gebruikt, is deze de verwerkingsverantwoordelijke.	De organisatie waaraan de ontwikkeling, training, validatie of (door)ontwikkeling wordt uitbesteed, mits de opdrachtgever: 1. het doel van de verwerkingsactiviteit vaststelt en; 2. de significante kenmerken van de te verwerken gegevens bepaalt. Dit ongeacht of deze opdrachtgever/ verwerkingsverantwoordelijke de persoonsgegevens aan de verwerker overdraagt of de verwerker deze via zijn eigen kanalen verkrijgt en; 3. de verwerker deze gegevens uitsluitend verwerkt voor de door de verwerkingsverantwoordelijke vastgestelde doeleinden.
Lancering/ Release/ Ingebruikname	Als het AI-systeem (al dan niet deel uitmakend van een ruimer product of dienst) verkocht of gelicentieerd wordt én persoonsgegevens bevat, wisselen beide organisaties persoonsgegevens uit en zijn ze beide verwerkings-verantwoordelijke. Ook indien bijvoorbeeld een licentiegever een systeem ter beschikking stelt aan een licentienemer en enkel deze laatste de verwerkingsverantwoordelijke is (zie hiernaast), wordt de licentiegever alsnog mede-verwerkingsverantwoordelijke wanneer deze persoonsgegevens, afkomstig van de licentienemer, voor eigen doeleinden gaat verwerken (bijvoorbeeld teneinde de efficiëntie van het AI-systeem te meten).	Elke organisatie die een AI-systeem ter beschikking stelt van een verwerkings-verantwoordelijke waarbij het AI-systeem wordt geïntegreerd in diens product of dienstverlening, of elke organisatie die dit doet omdat het noodzakelijk is voor de correcte uitvoering van haar dienstverlening, maar die zelf geen persoonsgegevens, verkregen van de verwerkingsverantwoordelijke, voor eigen doeleinden verwerkt. Een organisatie (dienstverlener) die een AI-systeem ter beschikking stelt van een andere organisatie (gebruiker) is geen verwerker en ook geen verwerkingsverantwoordelijke wanneer: 1. dit systeem lokaal en op zichzelf staand bij de gebruiker geïnstalleerd wordt; 2. de dienstverlener geen toegang heeft tot de lokale installatie, bijv. voor onderhoud.

Profilering	De organisatie die besluit persoonsgegevens via een AI-systeem voor eigen doeleinden te verwerken. De uitzondering voor zuiver persoonlijke of huishoudelijke activiteiten geldt niet voor organisaties (bijvoorbeeld Amazon) die middelen (Amazon Echo/Alexa) verschaffen om persoonsgegevens te verwerken voor hun eigen doeleinden, waarbij deze middelen doorgaans in de context van dergelijke persoonlijke of huishoudelijke activiteiten worden gebruikt, zoals bijvoorbeeld spraakassistenten.	Zie hierboven.
Geautomatiseerde besluitvorming	De entiteit die de geautomatiseerde besluitvorming met betrekking tot de betrokkenen voor eigen doeleinden uitvoert.	Zie hierboven.

Daarenboven is het niet uitgesloten dat als twee organisaties samen het doel en de middelen bepalen voor de verwerking door middel van een AI-systeem zij als **gezamenlijke verwerkingsverantwoordelijken** dienen beschouwd te worden. Dit kan bijvoorbeeld het geval zijn als een organisatie samenwerkt met een andere organisatie ter ontwikkeling van een product of dienst waarvoor beide partijen persoonsgegevens aanleveren voor de training en/of validatie van de tool en waarbij zij gezamenlijk het doel van die verwerking bepalen en hun technische middelen combineren, zonder dat de ene partij louter op instructie van de andere partij persoonsgegevens verwerkt.

In principe kan enkel de verwerkingsverantwoordelijke beslissen om in het **kader van een verwerkings-activiteit** (al dan niet) voor **een technische oplossing op basis van AI** (of een andere technologie) te kiezen. De verwerkingsverantwoordelijke is dan ook verplicht met de nodige zorgvuldigheid te handelen bij de keuze van de concrete IT-tool, en met name bij het uitbesteden van de verwerking of verwerven van de tool. Niettemin kan in bepaalde omstandigheden een verwerker zelf beslissen over de aangewende technische middelen. In dat geval zal ook de verwerker een deel van deze verantwoordelijkheid dragen.

Een verwerkingsverantwoordelijke moet dus op voorhand de **eventuele kwaliteitsspecificaties van de relevante oplossing op lijsten en beoordelen** en de (benodigde) **omvang van de verwerking vaststellen**. Een verwerkingsverantwoordelijke dient immers de gevolgen van de gerelateerde beslissingen op zich te nemen. Ten aanzien van de betrokkenen zal hij/zij zich niet kunnen onttrekken van zijn/haar verantwoordelijkheid door te beweren dat hij/zij niet over de correcte informatie of technische kennis beschikte. Het is zijn/haar verantwoordelijkheid om zo nodig een audit uit te voeren en te beslissen of het beoogde systeem wel in aanmerking komt voor het beoogde doeleinde, en niet bijvoorbeeld disproportioneel veel persoonsgegevens verwerkt.

Verwerkingsverantwoordelijken en verwerkers kunnen hun **verantwoordelijkheid** in ieder geval nooit **afschuiven op het AI-systeem** zelf en kunnen zich dus bijvoorbeeld niet verschuilen achter de mogelijke

complexiteit of ondoordringbaarheid van een AI-systeem om schendingen van de AVG te rechtvaardigen.

Hoofdstuk 4: Gegevensbescherming bij het ontwerp en de ontwikkeling van AI-systemen

4. Hoe kan gegevensbescherming worden verzekerd bij het ontwerp (design) en de ontwikkeling (development) van AI-systemen?

In de volgende delen worden een aantal zaken besproken uit de AVG die van belang zijn bij het ontwerp en de ontwikkeling van AI-systemen, namelijk gegevensbescherming door ontwerp (*data protection by design*), minimale gegevensbescherming (*data minimisation*), beveiliging van verwerking (*data security*), DPIAs en de verwerking van persoonsgegevens met het oog op wetenschappelijk onderzoek of statistische doeleinden.

4.1. Wat betekent gegevensbescherming door ontwerp en hoe kan dit geïmplementeerd worden in AI-systemen?

Essentie
De ontwikkeling en het gebruik van AI-systemen en de daarop betrekking hebbende processen moeten zo ontworpen zijn dat ze van nature uit zoveel mogelijk bescherming bieden aan persoonsgegevens. Dit is het principe van gegevensbescherming door ontwerp. Bij elk proces en elke ontwikkeling moet vanaf de ontwerpfase worden nagedacht of en hoe deze invloed (kunnen) hebben op de manier waarop persoonsgegevens worden verwerkt. Op basis daarvan worden dan de nodige veiligheden in het proces of het product mee ingebouwd. Ook moeten producten en diensten zo ontworpen worden dat ze bij een standaardgebruik, de meest gegevensbeschermingsvriendelijke instellingen gebruiken, zodat de eindgebruikers er enkel van kunnen afwijken indien dat uitdrukkelijk hun bedoeling is. Dit is het principe van gegevensbescherming door standaardinstellingen (*privacy by default*), dat deel uitmaakt van gegevensbescherming door ontwerp. Door gegevensbescherming door ontwerp toe te passen, wordt een zogenaamde '*tech debt*'[38] vermeden: de kost van het naleven van de AVG wordt van meet af aan in rekening gebracht. Bovendien wordt vermeden dat systemen achteraf tegen een (vaak) grotere kost herschreven moeten worden om aan de AVG te voldoen, als dat al mogelijk is.

[38] Een zogenaamde technical debt of tech debt is een begrip in softwareontwikkeling dat de impliciete kosten van extra herbewerking weerspiegelt, veroorzaakt door het kiezen van een gemakkelijke (beperkte) oplossing tijdens de ontwikkeling in plaats van een betere aanpak die langer zou duren. Door bepaalde inspanningen niet van meet af aan mee te nemen, wordt een zekere schuld opgebouwd aangezien deze in de uiteindelijke versie alsnog rechtgezet moeten worden.

Actiepunten

Algemeen:[39]

- ◊ Zorg dat alle medewerkers zich bewust zijn van het belang van gegevensbescherming, aandacht hebben voor de risico's en zich hier mee voor verantwoordelijk voelen (opleiding en bewustmaking).
- ◊ Voorzie praktische interne documenten en richtlijnen die toegepast moeten worden door personeel wanneer met persoonsgegevens wordt gewerkt.
- ◊ Zet een performant databeleid op dat de toegang tot rauwe data beperkt en toelaat data, de toegang ertoe en het gebruik ervan te regelen, te herkennen en te traceren.[40]
- ◊ Anonimiseer en pseudonimiseer persoonsgegevens wanneer mogelijk.[41]
- ◊ Zet een performant IT-beleid op met onder andere performante technische en organisatorische veiligheden, rolbeleid, encryptie en bewustmaking van het personeel.[42]
- ◊ Maak gebruik van technologische en veiligheidstoepassingen die voldoen aan de stand van de techniek en zorg dat ze hieraan blijven voldoen.[43]
- ◊ Neem gegevensbescherming op als productvereiste bij elke nieuwe ontwikkeling en elk nieuw proces.
- ◊ Stel in elke fase en bij elke ontwikkeling de vraag of de uitvoering daarvan op dat moment of op een later ogenblik een impact kan hebben op de gegevensbescherming.
- ◊ Voorzie duidelijke richtlijnen om te bepalen wanneer een DPIA uitgevoerd moet worden.[44]
- ◊ Voorzie veiligheden en zogenaamde *nudges*[45] in systemen voor eindgebruikers, zodat ze zich ervan bewust worden gemaakt dat bepaalde acties mogelijk een risico voor de gegevensbescherming inhouden.
- ◊ Documenteer de gemaakte analyses, afwegingen en keuzes aangaande gegevensbescherming die in elke fase worden gemaakt, zodat kan worden aangetoond dat aan de verplichting van gegevensbescherming door ontwerp is voldaan. Dit maakt het ook mogelijk om terug te vinden waarom bepaalde maatregelen wel of niet werden genomen. Motiveer daarbij telkens de keuze.
- ◊ Stel de standaardinstellingen van gebruikssoftware zo in dat persoonsgegevens standaard op de meest voorzichtige manier verwerkt worden en de eindgebruiker er uitdrukkelijk van moet afwijken om persoonsgegevens op een minder voorzichtige manier te verwerken.
- ◊ Leg aan leveranciers op dat hun diensten en producten voldoen aan de AVG-vereisten, waaronder de verplichting van gegevensbescherming door ontwerp en standaardinstellingen.
- ◊ Maak gebruik van zogenaamde gegevensbeschermings-bevorderende technologieën of *Privacy Enhancing Technologies* (PET's).[46]

[39] Gegevensbescherming door ontwerp is vooral een procesverplichting die inhoudt dat andere substantiële verplichtingen mee in het ontwerp van een product, dienst of proces worden opgenomen. Deze actiepunten hebben dan ook steeds betrekking op andere uit de AVG volgende verplichtingen en principes, zoals de minimale gegevensverwerking, de verplichting om een adequaat veiligheidsbeleid op te zetten of de verplichting om te documenteren op welke wijze de AVG nageleefd wordt. Deze worden verder besproken.

[40] Zie ook "4.2. Wat houdt de vereiste van minimale gegevensverwerking van persoonsgegevens in voor AI-systemen?" en "4.3. Wat zijn aandachtspunten voor de beveiliging van de verwerking van persoonsgegevens door AI-systemen?"

[41] Zie ook "5.2. Welke beperkingen legt de AVG op voor de bewaring van persoonsgegevens?"

[42] Zie ook "4.2. Wat houdt de vereiste van minimale gegevensverwerking van persoonsgegevens in voor AI-systemen?" en "4.3. Wat zijn aandachtspunten voor de beveiliging van de verwerking van persoonsgegevens door AI-systemen?"

[43] Zie ook "4.2. Wat houdt de vereiste van minimale gegevensverwerking van persoonsgegevens in voor AI-systemen?" en "4.3. Wat zijn aandachtspunten voor de beveiliging van de verwerking van persoonsgegevens door AI-systemen?"

[44] Zie ook "4.4. Wanneer moet een DPIA of GEB worden uitgevoerd voor een verwerking van persoonsgegevens door AI-systemen?"

[45] Nudges zijn architecturale keuzes in de software die een bepaald gedrag bevorderen. Bijvoorbeeld het vooraf aanvinken van een bepaalde optie waarvan gewenst is dat de gebruiker deze meeneemt.

[46] Zie ook "4.2. Wat houdt de vereiste van minimale gegevensverwerking van persoonsgegevens in voor AI-systemen?"

Specifiek voor AI-systemen:

- ◊ Bouw AI-systemen op die manier dat ze de volgende acties toelaten:
  - ◊ het achterhalen (tot op zekere hoogte) van de ratio achter gegenereerde uitkomsten (transparantie) en van de hiervoor gebruikte persoonsgegevens;
  - ◊ het uitoefenen door betrokkenen van hun rechten;
  - ◊ getraind worden met gepseudonimiseerde en/of versleutelde gegevens;
  - ◊ getraind worden met zo weinig mogelijk, maar kwalitatieve data.[47]
- ◊ Vermijd dat her-identificatie plaatsvindt bij verwerking van anonieme gegevens of dat uit de beschikbare gegevens andere persoonsgegevens afgeleid worden, die niet verwerkt mogen worden, zoals gevoelige gegevens.[48]
- ◊ Zorg dat de effectiviteit van het AI-systeem voldoende getest is en valse positieven en negatieven (*hidden failures*) afdoende uitgesloten worden.
- ◊ Ken je data:
  - ◊ Waak over de herkomst van en de rechten op persoonsgegevens die gebruikt worden om AI-systemen te trainen. Verzeker dat de dataset gebruikt mag worden.
  - ◊ Zorg dat degenen die de gegevens verzamelen en behandelen in staat zijn om persoonsgegevens te herkennen.
- ◊ Zorg voor degelijk opgeschoonde datasets:
  - ◊ verwijder overtollige gegevens;
  - ◊ zorg voor voldoende representatieve datasets;
  - ◊ controleer dat datasets geen vooroordelen of bias in zich meedragen die eventuele sociale ongelijkheid of discriminatie versterken.
- ◊ Zorg voor de nodige documentatie in kader van de verantwoordingsplicht (accountability):
  - ◊ Documenteer de toegepaste gegevensbeschermingsanalyses, de daarbij gemaakte keuzes en afwegingen en de DPIA's uitgevoerd tijdens de ontwikkeling, het testen en het onderhoud.
  - ◊ Documenteer de eigenschappen van datasets, de manier waarop deze opgeschoond werden en de redenering daarachter. Bewaar ook minstens een representatief staal van de dataset, indien mogelijk geanonimiseerd.

4.1.A. Gegevensbescherming inbouwen in applicaties en processen

Gegevensbescherming mag geen laag vernis zijn die bovenop een AI-systeem gelegd wordt, maar moet er inherent aan zijn. Conform de verplichting van gegevensbescherming door ontwerp moeten de nodige maatregelen al **bij het bepalen van de verwerkingsmiddelen[49] worden genomen** om:

- de nodige waarborgen voor AVG-conforme verwerking van persoonsgegevens te verzekeren; en
- de nastreving van de gegevensbeschermingsbeginselen[50] in de geplande verwerking in te bouwen.

Dit betekent niet dat pas bij het gebruik van een AI-systeem gekeken mag worden hoe persoonsgegevens in dat systeem AVG-conform kunnen worden verwerkt. De **conforme verwerking moet vanaf het begin** in het systeem ingebakken zijn.

---

en "4.3. Wat zijn aandachtspunten voor de beveiliging van de verwerking van persoonsgegevens door AI-systemen?"

[47] Zie ook "4.2. Wat houdt de vereiste van minimale gegevensverwerking van persoonsgegevens in voor AI-systemen?"

[48] Zie ook "3.2. Wat zijn bijzondere categorieën van persoonsgegevens?"

[49] Met "op het ogenblik van het bepalen van de verwerkingsmiddelen" wordt gedoeld op het ogenblik waarop de te gebruiken processen, technieken en werkwijzen bepaald worden, waarmee persoonsgegevens verwerkt zullen worden. Dit gebeurt normaliter in de ontwerpfase.

[50] Deze beginselen zijn rechtmatigheid, behoorlijkheid en transparantie, doelbinding, minimale gegevensverwerking, juistheid, opslagbeperking, integriteit en vertrouwelijkheid en verantwoordingsplicht (art. 5 AVG).

Ook na de ontwerpfase moet de naleving van de gegevensbescherming door ontwerp-vereiste verder geëvalueerd worden, rekening houdend met eventuele wijzigende omstandigheden. Voldoen aan deze verplichting is dan ook een **voortdurend proces en continue oefening** die gedurende de hele levensduur van een AI-systeem loopt.

Producenten van producten, diensten en toepassingen voor andere gebruikers die deze niet zelf gaan gebruiken of uitvoeren en die dus zelf geen persoonsgegevens verwerken, zijn strikt gezien niet verplicht om deze verplichting toe te passen.[51] Software die persoonsgegevens verwerkt, moet dus strikt gezien niet aan de gegevensbescherming door ontwerp verplichtingen voldoen indien de ontwikkelaar deze software niet zelf gebruikt. De gebruikers van deze software, die door middel van deze software persoonsgegevens verwerken, moeten wel aan de regels van gegevensbescherming voldoen. Vanuit de gebruikersmarkt zal er dus wel een nood zal zijn om producten af te nemen die voldoen aan de vereiste van gegevensbescherming door ontwerp, ook al moeten de producenten deze strikt gezien niet inbouwen. Wanneer de producent via zijn product ook diensten levert en daardoor verwerker wordt, moet hij echter wel aan deze verplichting voldoen.[52]

Toepassing
In e-commerce kan software bijvoorbeeld in staat gemaakt worden om 'gevoelige' gegevens te herkennen en de gebruiker te waarschuwen wanneer vermoedelijk gevoelige gegevens verwerkt worden. Ook zouden automatisch AVG-conforme meldingen kunnen worden toegevoegd aan reclameberichten die gegenereerd worden door de toepassing waar het AI-systeem deel van uitmaakt. Het is ook aangewezen dat de persoon die gegevens van prospecten invoert, steeds via een keuzemenu moet bevestigen van welke (vooraf bepaalde) bron deze afkomstig zijn, zodat deze persoon zich bewust is van de herkomst van deze gegevens én de gegevensbron gemakkelijk terug te vinden is. Ook in rekrutering is software idealiter in staat om 'gevoelige' gegevens te herkennen en de gebruiker te waarschuwen wanneer vermoedelijk gevoelige gegevens verwerkt worden. Er kunnen ook meldingen worden ingebouwd in de gebruikersinterface die de softwaregebruiker waarschuwen dat aan bepaalde voorwaarden voldaan moet zijn, wanneer bijvoorbeeld de keuze gemaakt wordt om publiek beschikbare sociale media informatie van een kandidaat op te halen (wat niet zomaar mag). Bij de aanmaak van een vacaturemelding kan ook worden gezorgd dat informatie automatisch wordt toegevoegd die de kandidaten informeert over de verwerking van hun persoonsgegevens, zoals de verwijzing naar het toepasselijke gegevensbeschermingsbeleid (de '*privacy policy*').

4.1.B. Risico-gebaseerde benadering

Het risico-gebaseerde karakter van de AVG komt sterk naar voor bij de vereiste van gegevensbescherming door ontwerp aangezien de omvang van de bijhorende verplichtingen volledig van de context afhangt. Onder meer de stand van de techniek, de kosten, het doel van de verwerking en de risico's verbonden aan de verwerking moeten daarbij in aanmerking genomen worden.[53]

[51] Zie ook overweging 78 AVG waarin staat dat deze producenten 'gestimuleerd' moeten worden om gegevensbescherming door ontwerp toe te passen.

[52] Overweging 78 AVG.
[53] Art. 25.1 AVG.

Hoe groter het risico, hoe meer inspanningen van een producent vereist zijn. Een AI-systeem dat minder persoonsgegevens verwerkt en/of niet rechtstreeks met mensen in contact komt, zal minder vergaand aangepast moeten worden dan een systeem dat intensief persoonsgegevens of gevoelige persoonsgegevens[54] zal verwerken en/of rechtstreeks met mensen communiceert.

Ook de **kost** speelt een rol bij de beoordeling om voldoen is aan de vereiste van gegevensbescherming door ontwerp. De kost heeft niet alleen betrekking op het financiële aspect, maar op alle inspanningen die geleverd worden inclusief de bestede tijd en personeelskosten. Van een kleine onderneming met beperkte middelen kan ook niet worden verwacht dat deze dezelfde veiligheden kan inbouwen dan een grote internationale onderneming. Er mag echter geen afbreuk worden gedaan aan de vereiste om adequate doeltreffende maatregelen te nemen. Het niet beschikken over voldoende middelen is immers geen rechtvaardiging om de AVG-vereisten niet na te leven.

Toepassing
In e-commerce vereist een reclamecampagne op basis van eenvoudige criteria zoals de opgegeven woonplaats van een bestaande klant minder vergaande veiligheidsmaatregelen en voorzichtigheid dan een campagne waarbij bijvoorbeeld ook gebruik gemaakt wordt van locatiegegevens of gegevens die via derden werden verkregen. Bij rekruteringsapplicaties die een (pre-)selectie maken uit ingediende kandidaturen moeten meer veiligheidswaarborgen ingebouwd worden dan bij applicaties die enkel een score geven en waarbij de eigenlijke selectie van kandidaten aan de gebruiker wordt overgelaten.

4.1.C. Gegevensbescherming door ontwerp heeft betrekking op het naleven van alle bepalingen in de AVG

De AVG verwijst voor gegevensbescherming door ontwerp uitdrukkelijk naar pseudonimisering[55] en het principe van gegevensbeperking.[56] Dit betekent echter niet dat enkel deze maatregelen in rekening moeten worden genomen om aan de vereiste van gegevensbescherming door ontwerp te voldoen. [57] Wel staat vast dat een **degelijk gegevensbeperkingsbeleid** enerzijds en **pseudonimisering** anderzijds onontbeerlijk zijn voor het nakomen van deze verplichting.

Gegevensbescherming door ontwerp is dan ook niet zozeer een inhoudelijke verplichting, maar eerder een **procesverplichting** die vereist om alle uit de AVG volgende verplichtingen zoals transparantie,[58] rechtmatigheid of veiligheid[59] zo veel mogelijk in processen in te bouwen.

Op welke manier deze verplichtingen geïmplementeerd moeten worden is **zelf te bepalen**. Technische

[54] Zie ook "3.2. Wat zijn bijzondere categorieën van persoonsgegevens?"
[55] Art. 25.1 AVG. Zie ook "4.1. Wat betekent gegevensbescherming door ontwerp en hoe kan dit geïmplementeerd worden in AI-systemen?"
[56] Art. 25.2 AVG. Zie ook "4.2. Wat houdt de vereiste van minimale gegevensverwerking van persoonsgegevens in voor AI-systemen?"
[57] Overweging 28 AVG. Dat blijkt onder meer uit de brede algemene verwijzingen naar het moeten nemen van "technische en organisatorische maatregelen", het waarborgen van "de" gegevensbeschermingsbeginselen en het inbouwen van waarborgen ter bescherming van "de" rechten van individuen en ter naleving van "de" voorschriften van de AVG.
[58] Zie ook "5.1. Welke transparantieverplichtingen legt de AVG op en wat zijn de specifieke aandachtspunten in een AI-context?"
[59] Zie ook "4.3. Wat zijn aandachtspunten voor de beveiliging van de verwerking van persoonsgegevens door AI-systemen?"

en organisatorische maatregelen kunnen alle mogelijke acties omvatten, gaande van uitleggen aan personeel hoe ze klantendata moeten verwerken tot het gebruik van gesofisticeerde technische geautomatiseerde oplossingen. De maatregelen moeten daarbij niet gesofisticeerd zijn. De enige vereiste is dat ze effectief en adequaat zijn bij het waarborgen van de AVG-verplichtingen.

Daarbij is wel vereist dat **rekening gehouden wordt met de stand van de techniek**.[60] Dit betekent dat technologische evoluties in rekening moeten worden gebracht. De genomen maatregelen om conformiteit met de AVG te verzekeren zijn dus dynamisch en moeten zo nodig worden aangepast aan de stand van de techniek. De verplichting om rekening te houden met de stand van de techniek is zowel van toepassing op de technische als de organisatorische maatregelen.

Toepassing
In zowel e-commerce als rekrutering vereist gegevensbescherming door ontwerp dat in elke fase veiligheden ingebouwd worden. Dit kan op verschillende niveaus: • **Technisch** in de software zelf door: bepaalde gegevens te herkennen; gegevens in versleutelde vorm op te slaan; toegang tot persoonsgegevens te registreren; rolbepaling toe te staan zodat elk gebruikersprofiel enkel die gegevens te zien krijgt die vereist zijn voor de rol; veilige paswoorden te vereisen; waarschuwingen aan de gebruikers te geven en nudges te voorzien die een correct gebruik bevorderen (bijvoorbeeld extra stappen bij toegang tot gevoelige gegevens inbouwen); bewaringstermijnen voor de verschillende types persoonsgegevens te voorzien en de gebruiker te waarschuwen wanneer deze bereikt worden. • **Organisatorisch** door bijvoorbeeld te zorgen dat het personeel dat deze software gebruikt voldoende kennis van gegevensbescherming heeft om de toepasselijke regels correct toe te passen, zowel bij gebruik van software als daarbuiten; en hiervoor de nodige interne procedures te voorzien. • **Door de software in elke fase een correcte verwerking van gegevens te laten nastreven**: bij het aanmaken en opslaan van gebruikersaccounts; bij de invoer van klantengegevens, leads, kandidaten en interessante profielen; bij het plannen van reclamecampagnes; bij het genereren van advertenties (door verplichte vermeldingen te voorzien in deze advertenties); bij het kiezen van de selectiecriteria voor het doelpubliek; bij het tonen van advertenties; bij het gebruik van gegevens voor het bepalen van een consumentenprofiel of aanwervingsprofiel; en bij het verstrekken van informatie aan het doelpubliek over de redenen waarom zij bepaalde advertenties te zien krijgen, hoe lang hun gegevens verwerkt worden en hoe zij zich kunnen uitschrijven. Specifiek voor de rekrutering kunnen afgewezen kandidaten automatisch worden ingelicht of ze in een wervingsreserve (of andere database) opgenomen worden en voor hoe lang. Eventueel kan om de twee jaar nog een geautomatiseerde bericht worden gestuurd naar kandidaten in de wervingsreserve om hen te vragen of ze hierin opgenomen willen blijven.

4.1.D. Risicoanalyse

De vereiste gegevensbescherming onder de AVG is risico-gebaseerd. Bijgevolg moet bij elke ontwikkeling, elke fase en elk product of proces, ook voor AI-systemen, in zijn geheel de vraag worden gesteld of er een **impact is op het recht op gegevensbescherming** van degenen wiens persoonsgegevens verwerkt

[60] Art. 25.1 AVG.

zullen worden. Op die manier wordt een degelijke gegevensbescherming een **(product)vereiste** van het te ontwikkelen product of proces.

Het toepassen van zogenaamde **bedreigingsmodellering** is een methode om de potentiële impact van een AI-systeem op gegevensbescherming te analyseren.[61] Daarvoor kan gebruik worden gemaakt van bedreigingsmodelleringmethodes zoals de ***LINDDUN Privacy Threat Modeling*** methode voor software ontwikkeling.[62] In functie van daarvan kan dan worden beslist om geen of bepaalde maatregelen te nemen om het niveau van gegevensbescherming te verhogen.[63]

Indien de kans bestaat dat een beoogde verwerking een groot risico inhoudt op de gegevensbescherming moet ook worden geanalyseerd of een **DPIA nodig** is.[64]

Toepassing
In e-commerce en rekrutering is het van belang om: • de nalevingsfunctionaris (*compliance officer*) van meet af aan bij productontwikkelingsprocessen te betrekken en de mogelijk impact op gegevensbescherming vanaf de ideevormingsfase/de creatieve fase te analyseren; • jaarlijks de toegepaste processen en werkwijzen en hun impact op de gegevensbescherming te evalueren; • nieuwe functionaliteiten die toegevoegd worden aan de gebruikte software te evalueren (neem niet zomaar aan dat deze gegevensbeschermingsvriendelijk zijn); • te vereisen van leveranciers dat ze de AVG naleven en hun producten aan de toepasselijke vereisten voldoen.

4.1.E. Documenteer de gemaakte evaluaties en de genomen maatregelen

Ingevolge de algemene 'verantwoordingsplicht' onder de AVG moet het nakomen van de verplichting van gegevensbescherming door ontwerp **bewezen kunnen worden**.[65]

Er moet dus bewezen kunnen worden dat **maatregelen** genomen werden én dat ze **doeltreffend zijn**. Daarbij is het noodzakelijk om in verslagen op te nemen:
- dat de impact op gegevensbescherming geanalyseerd werd;
- welke maatregelen daarbij genomen werden;
- welke afwegingen daarbij gemaakt werden; en
- wat de resultaten daarvan zijn.

[61] Bedreigingsmodellering (*threat modeling*) is een methode waarmee mogelijke risico's zoals structurele zwakheden of het ontbreken van passende beschermingsmaatregelen kunnen worden geïdentificeerd, opgesomd, geëvalueerd en geprioriteerd in functie van hun risico.
[62] Voor meer informatie: https://www.linddun.org. Zie ook: K. Wuyts, Privacy Threats in Software Architectures, Ph.D., 2015.
[63] Daarbij kan een afweging worden gemaakt en bijvoorbeeld worden beslist dat bepaalde maatregelen die een gunstig effect hebben op de gegevensbescherming niet genomen worden omdat de kost verhoudingsgewijs te groot is of de impact miniem is.
[64] Zie ook "4.4. Wanneer moet een DPIA of GEB worden uitgevoerd voor een verwerking van persoonsgegevens door AI-systemen?"
[65] Art. 5.2 AVG. De verantwoordingsplicht of het *accountability*-principe is één van de basisbeginselen inzake gegevensverwerking. Het houdt in dat verwerkingsverantwoordelijken en verwerkers moeten kunnen aantonen dat zij stappen hebben ondernomen om de verplichtingen uit de AVG na te leven.

Op die manier kan achteraf de **nodige informatie** enerzijds **worden teruggevonden** én anderzijds **worden aangetoond** op welke wijze over gegevensbescherming nagedacht werd en waarom bepaalde maatregelen al dan niet genomen werden.

Toepassing
In e-commerce en rekrutering is het van belang om: • te zorgen dat een verslag gemaakt wordt van de gemaakte afwegingen en de risicoanalyses van de productontwikkeling; • kennisgevingen aan kandidaten, klanten of prospecten te bewaren; acties te documenteren zoals de volgende: registreer zowel het inschrijven als het uitschrijven door een kandidaat of prospect op bepaalde communicatie; registreer in een bestand met het tijdstip en (in voorkomend geval) erop betrekking hebbende correspondentie.

4.1.F. Vermijd '*technical debt*' door gegevensbescherming door ontwerp toe te passen

Door rekening te houden met de toepasselijke AVG-principes vanaf de ontwikkelingsfase van AI-systemen wordt vermeden dat het eindproduct bezwaard is met zogenaamde ***technical debt***[66]. Indien daar geen rekening mee gehouden wordt, moet het systeem achteraf worden aangepast om aan de AVG-vereisten te voldoen, wat vaak een complexe of zelfs onmogelijke taak is. Daarenboven kan dat ook een grote kost met zich meebrengen.

4.1.G. Gegevensbescherming door standaardinstellingen

Bij het ontwikkelen van AI-systemen moet er voor worden gezorgd dat de **standaardconfiguratie de meest gegevensbeschermingsvriendelijke instellingen** biedt. Dit is de vereiste van gegevensbescherming door standaardinstellingen.[67]

Wanneer **AI-systemen met standaardinstellingen** gebruikt worden moeten ze dus bijvoorbeeld:
- persoonsgegevens enkel verwerken binnen de grenzen van de beoogde rechtmatige verwerking;
- enkel die persoonsgegevens verzamelen en verwerken die voor die verwerking vereist zijn;
- persoonsgegevens niet langer bewaren dan nodig is;[68]
- de toegang tot (niet-gepseudonimiseerde) persoonsgegevens beperken.
- de persoonsgegevens niet publiek verspreiden.

[66] Zie voor een bespreking van dit begrip voetnoot 38 op pagina 29.
[67] Art. 25.3 AVG.
[68] Zie ook "5.2. Welke beperkingen legt de AVG op voor de bewaring van persoonsgegevens?"

Toepassing
In e-commerce is het bijvoorbeeld aangewezen om standaard enkel communicatie te versturen naar personen wanneer bij het ingeven van hun gegevens een optieveld ingevuld werd met een waarde die toelaat om deze personen te contacteren (bijvoorbeeld 'toestemming gegeven' of 'bestaande klant'). Het is dan ook van belang dat standaard geen communicatie gericht kan worden aan personen die zich daartegen verzet hebben, bijvoorbeeld door op de uitschrijflink in een e-mail te klikken. In het geval van rekrutering is het aangewezen om standaard velden in job advertenties te voorzien voor verwijzing naar het gegevensbeschermingsbeleid en tekstvelden voor een korte toelichting over de verwerking van gegevens.

4.1.H. Certificering

Certificering[69] kan gebruikt worden als **element** om aan te tonen dat aan de vereiste van gegevensbescherming door ontwerp voldaan werd.[70] Hieruit volgt dat zelfs in geval van certificering nog steeds moeten worden aangetoond dat effectief aan deze verplichting voldaan werd. Enkel de **bewijslast zal lichter** zijn. Tot op vandaag bestaan er geen standaarden voor gegevensbescherming door ontwerp, en is het niet mogelijk om te laten certificeren dat een product of dienst daaraan zou voldoen.

4.2. Wat houdt de vereiste van minimale gegevensverwerking van persoonsgegevens in voor AI-systemen?

Essentie
Minimale gegevensverwerking moet toegepast worden bij het verzamelen, het gebruik en de opslag van persoonsgegevens. De intern verleende toegang tot deze persoonsgegevens moet ook beperkt worden. Voor elk van deze aspecten geldt dat de persoonsgegevens in essentie slechts verzameld, gebruikt, opgeslagen en geraadpleegd mogen worden voor zover dat noodzakelijk is voor de doeleinden waarvoor ze verwerkt (mogen) worden. Indien een gelijkaardig resultaat dus bereikt kan worden zonder (bepaalde) persoonsgegevens te gebruiken, mogen die persoonsgegevens daarvoor niet worden gebruikt.[71] De verwerkte persoonsgegevens en de toegang ertoe moeten beperkt worden tot het strikt noodzakelijke. Daarbij moet voorkomen worden dat onnodig kopieën van de persoonsgegevens worden gemaakt. Persoonsgegevens mogen niet langer dan nodig worden bewaard,[72] wat aansluit bij de 'opslagbeperking'.[73]

[69] Zoals voorzien in art. 42 AVG.
[70] Art. 25.3 AVG.
[71] Overweging 39 AVG.
[72] Zie ook "5.2. Welke beperkingen legt de AVG op voor de bewaring van persoonsgegevens?"
[73] Zie hierna deel D voor meer informatie over deze en andere technieken en methoden.

Actiepunten

- ◊ Ken je data: de herkomst, de eraan verbonden rechten, de eigenschappen en de attributen moeten gekend zijn om te weten of ze correct verzameld werden en voor welke doeleinden ze gebruikt mogen worden.
- ◊ Bepaal vooraf welke types (kwalitatief) persoonsgegevens en ook welk volume (kwantitatief) aan persoonsgegevens strikt noodzakelijk zijn om het AI-systeem te kunnen trainen en/of gebruiken.
- ◊ Doe daarvoor beroep op een multidisciplinair team, waaronder experten in het domein waar het AI-systeem toegepast zal worden. Motiveer en documenteer.
- ◊ Evalueer of technieken en/of systemen kunnen worden toegepast die toelaten om met minder persoonsgegevens te werken, met versleutelde persoonsgegevens te verwerken of bijvoorbeeld methoden gebruikt kunnen worden zoals *Generative Adversarial Networks* (GAN's), *federated learning of transfer learning*.[74]
- ◊ Zet een performant pseudonimiseringsbeleid op. Zorg dat de persoonsgegevens slechts in gepseudonimiseerde vorm gebruikt, geconsulteerd en verwerkt kunnen worden wanneer het niet strikt noodzakelijk is dat dit gebeurt met de 'rauwe' persoonsgegevens.
- ◊ Zet een performant anonimiseringsbeleid op. Anonimiseer of vernietig persoonsgegevens waarvan het niet meer noodzakelijk is dat de personen op wie ze betrekking hebben geïdentificeerd kunnen worden of waarvan niet langer gerechtvaardigd is dat ze verwerkt kunnen worden.
- ◊ Zorg voor een afdwingbaar rol- en toegangsbeleid in de organisatie, waarbij personen en applicaties enkel toegang hebben tot de rauwe persoonsgegevens als ze dat echt nodig hebben en na bijkomende identificatie of machtiging. Anderen krijgen slechts toegang tot gepseudonimiseerde persoonsgegevens. Zorg dat toegang tot persoonsgegevens enkel mogelijk is met een individuele account en log de toegang.
- ◊ Kuis datasets regelmatig op en verwerk data niet langer dan noodzakelijk.[75]
- ◊ Documenteer en registreer elke stap en evaluatie, conform de verantwoordingsplicht.
- ◊ Belangrijk is dus dat:
  - ◊ gemotiveerd kan worden waarom bepaalde persoonsgegevens verwerkt worden;
  - ◊ persoonsgegevens centraal op één plaats opgeslagen worden zonder onnodige kopieën;
  - ◊ een goed ICT-toegangsbeleid wordt opgezet dat rekening houdt met de verschillende rollen van personeel in het verwerkingsproces;
  - ◊ een performant anonimiserings- en pseudonimiseringsbeleid wordt opgezet;
  - ◊ waar mogelijk zo weinig mogelijk persoonsgegevens gebruikt worden;
  - ◊ technieken gebruikt worden die de nood aan persoonsgegevens, het volume en de risico's op blootstelling beperken.

4.2.A. Algemeen

Persoonsgegevens moeten **toereikend** zijn, **ter zake dienend** zijn en **beperkt zijn tot wat noodzakelijk is** voor de doeleinden waarvoor zij verwerkt worden.[76] Zij mogen enkel worden verwerkt indien het doel van de verwerking niet redelijkerwijze op een andere wijze kan worden bereikt.[77]

Minimale gegevensverwerking is nauw verwant met het principe van opslagbeperking,[78] de vereiste

[74] Zie hierna deel D voor meer informatie over deze en andere technieken en methoden.
[75] Zie ook "5.2. Welke beperkingen legt de AVG op voor de bewaring van persoonsgegevens?"
[76] Art. 5.1, c AVG.
[77] Overweging 39 AVG.
[78] Zie ook "5.2. Welke beperkingen legt de AVG op voor de bewaring van persoonsgegevens?"

'juistheid' van de persoonsgegevens en het verbod om persoonsgegevens voor andere doeleinden te (her)gebruiken ('doelbinding'). Minimale gegevensverwerking moet **ruim** bekeken worden en geldt niet enkel bij het **verzamelen** van persoonsgegevens, maar ook bij de **interne toegang** en het **gebruik** van persoonsgegevens. Minimale gegevensverwerking is een essentieel **onderdeel** van elke strategie tot gegevensbescherming door ontwerp.[79]

Minimale gegevensverwerking heeft de volgende **voordelen**:
- Bescherming van de rechten van de betrokkenen:
  - er is geen verdere inkijk in hun privéleven dan nodig;
  - zo weinig mogelijk mensen binnen een organisatie hebben zicht hebben op hun privéleven;
  - wie toegang had tot welke persoonsgegevens is traceerbaar.
- Vermindering van het risico op en bij gegevenslekken
  - informatie die niet verzameld wordt, kan ook niet gelekt worden;
  - persoonsgegevens die niet langer relevant zijn, kunnen niet gelekt worden;
  - de kans op gegevenslekken wordt kleiner. Dit geldt zowel voor bedreigingen binnen een organisatie (*insider threats*), als voor externe aanvallen (zoals *phishing* en *hacking*) en voor onbedoelde gegevenslekken;
  - het risico op schade bij gegevenslekken wordt kleiner. De schade is bijvoorbeeld kleiner wanneer enkel gepseudonimiseerde persoonsgegevens gelekt worden ingevolge een *phishing*aanval, omdat de persoon via wie het lek zich voordeed geen toegang had tot rauwe data.

4.2.B. Risico-gebaseerde benadering

Ook de vereiste tot minimale gegevensverwerking kan niet los van het risico-gebaseerde karakter van de AVG worden gezien.

Door minder persoonsgegevens te verwerken, wordt in eerste instantie het **risico verminderd** dat een **te vergaande inkijk in het privéleven** van de betrokkenen tot stand komt en dat uit deze persoonsgegevens onverwachte conclusies kunnen worden getrokken. Het degelijk opkuisen van datasets draagt bovendien ook bij aan de **kwaliteit van de data**.

Ook het **risico op een gegevenslek en de schade** bij een gegevenslek worden erdoor verminderd:
- Minder persoonsgegevens verwerken, gegevens sneller anonimiseren en pseudonimiseren, gegevens versleutelen, de toegang tot persoonsgegevens beveiligen en beperken of de persoonsgegevens op één centrale plaats bewaren zorgen er allemaal voor dat de kans op een gegevenslek verkleint.
- Indien zich toch een lek van deze gegevens zou voordoen, worden zowel het aantal getroffen personen, de bruikbaarheid van de gelekte gegevens persoonsgegevens en dus de omvang van de schade voor deze personen aanzienlijk beperkt.

Een ander gevolg hiervan is dat minimale gegevensbeperking verder moet worden geïmplementeerd en verzekerd naarmate de **risico's groter worden**. Dat is bijvoorbeeld het geval wanneer:
- grote volumes gegevens verwerkt worden;
- gegevens van meer personen verwerkt worden;
- meer types persoonsgegevens aangaande dezelfde personen verwerkt worden;
- persoonsgegevens verwerkt worden die onder de bijzondere categorieën vallen of die meer in het algemeen als gevoelig te bestempelen zijn.

[79] Zie ook "4.1. Wat betekent gegevensbescherming door ontwerp en hoe kan dit geïmplementeerd worden in AI-systemen?"

4.2.C. Minimale gegevensbescherming en de doeltreffendheid van AI-systemen

Vaak wordt aangenomen dat meer data (altijd) beter is voor de werking en de resultaten van AI-systemen.

Dit is echter niet noodzakelijk het geval. Door minimale gegevensbescherming toe te passen, wordt vermeden dat getraind wordt met irrelevante parameters. Hierdoor vermijdt men het risico dat deze parameters toch als betekenisvol beschouwd zouden worden door het AI-systeem en er onterecht conclusies op zou baseren. Goede kennis en opschoning van de dataset dragen dus bij aan de kwaliteit van de data én van de uitkomsten.

Door zorgvuldig te selecteren welke gegevens verwerkt zullen worden, wordt bovendien de zogenaamde ***curse of dimensionality***[80] **vermeden**. Daarbij wordt een (AI-)systeem geacht de beste prestaties te leveren, wanneer het gevoed wordt met een optimaal volume aan gegevens. Wordt dit optimum overschreden, dan gaat de efficiëntie naar beneden.

Ook het risico op ***overfitting* wordt vermeden** door enkel te werken met de relevante parameters. Door te veel parameters toe te voegen, kan het AI-systeem te veel aangepast zijn aan de trainingsdata. Het kan dan waarde gaan toekennen aan elementen die in de trainingsdata aanwezig zijn, maar die mogelijk minder relevant of niet altijd aanwezig zijn in andere datasets. Bijgevolg zal de efficiëntie van dergelijke AI-systeem dalen wanneer nieuwe datasets geanalyseerd moeten worden.[81]

4.2.D. Verzamel en gebruik minder persoonsgegevens (trainingsfase en gebruiksfase)

Bij de verzameling en het gebruik van persoonsgegevens zijn de hierna **volgende vragen en aspecten** relevant.

1. KAN DE VERWERKING VAN PERSOONSGEGEVENS IN ZIJN GEHEEL VERMEDEN WORDEN?
Minder persoonsgegevens verzamelen, betekent minder persoonsgegevens verwerken. De eerste stap bij het toepassen van minimale gegevensverwerking is dan ook zorgen dat **zo weinig mogelijk persoonsgegevens** verwerkt worden.

De eerste vraag is dan ook: zijn er effectief echte persoonsgegevens nodig? Met andere woorden: kan het AI-systeem ook getraind worden met andere gegevens? Bijvoorbeeld met synthetische persoonsgegevens[82] of geanonimiseerde gegevens[83]. Persoonsgegevens mogen enkel worden verwerkt indien het doel van de verwerking niet redelijkerwijze op een andere manier kan worden verwezenlijkt.[84]

[80] Vrij vertaald de 'vloek van de (multi-)dimensionaliteit'. Meer informatie: https://deepai.org/machine-learning-glossary-and-terms/curse-of-dimensionality.
[81] Meer informatie: https://www.datarobot.com/wiki/overfitting.
[82] *Synthetic data* vrij vertaald als 'synthetische gegevens' zijn (in dat geval) gegevens die persoonsgegevens nabootsen, maar die ofwel geproduceerd zijn, samengesteld zijn op basis van anonieme gegevens ofwel beide.
[83] *Zie ook "3.3. Wat is het belang van het onderscheid tussen anonimisering en pseudonimisering?"*
[84] Overweging 39 AVG.

Toepassing
Indien een e-commerce AI-systeem wordt getraind op bestaande klantengegevens van een gebruiker, is het mogelijk dat daarvoor niet op de eigenlijke gegevens, maar op een geanonimiseerde dataset wordt getraind. Ook in rekrutering kan een AI-systeem lokaal getraind worden op de persoonsgegevens van het personeel werkzaam in de organisatie die het systeem wil gebruiken. Ook daar is het aangewezen om deze dataset vooraf te anonimiseren indien mogelijk.

2. KAN DE VERWERKING GEBEUREN MET MINDER PERSOONSGEGEVENS?
Indien effectief vereist is dat 'echte' persoonsgegevens worden gebruikt, is de volgende vraag of het gebruik van de persoonsgegevens **proportioneel** is. Met andere woorden, welke van die persoonsgegevens zijn echt noodzakelijk om het vooropgestelde doel te bereiken?

Deze vraag moet op **twee vlakken** gesteld worden.

Enerzijds moet gekeken worden naar de **hoeveelheid personen** van wie gegevens moeten worden verwerkt. Anderzijds is de vraag ook hoeveel **verschillende gegevens** van elk van deze personen vereist is, dus hoeveel soorten persoonsgegevens verwerkt moeten worden. Daarbij moet ook een **redelijkheidstoets** gebeuren. Er moet worden nagegaan of het gebruik van de beoogde persoonsgegevens proportioneel is ten aanzien van het beoogde doel en ten aanzien van de risico's die de betreffende personen daarbij lopen.

Indien de training dus kan gebeuren met persoonsgegevens van x personen, dan mag geen dataset gebruikt worden die de gegevens van meer personen bevat. Indien slechts y parameters vereist zijn van elke persoon om een kwalitatief resultaat te bereiken, mogen niet meer parameters gebruikt worden. Ook niet omdat daar verbanden uit zouden kunnen volgen die nog niet geweten zijn. Indien de te gebruiken dataset op meer personen betrekking heeft en/of meer parameters bevat, moet deze **opgekuist worden** om hieraan te voldoen.

Bijgevolg moet **op voorhand** geanalyseerd en bepaald worden hoeveel en welke persoonsgegevens vereist zijn voor het verwerkingsdoel. Daarbij moet duidelijk vastgesteld en omschreven worden waarom deze vereist zijn, wat daaruit geleerd moet worden en waarom deze persoonsgegevens relevant en niet overmatig zijn. Enkel persoonsgegevens die deze toets doorstaan, mogen verwerkt worden.

Om deze analyse te kunnen maken is het noodzakelijk om **voldoende expertise** te hebben in het domein waarbinnen het AI-systeem analyses zal leveren.

Deze toets moet per **afzonderlijke verwerking** bepaald worden: het is niet omdat het gerechtvaardigd is om in een bepaalde dataset x parameters van y mensen te bewaren, dat deze volledige set gebruikt mag worden bij andere verwerkingen die gebruik maken van deze zelfde dataset.

Toepassing
Bij rekrutering en e-commerce is de vraag of met minder persoonsgegevens gewerkt kan worden relevant bij zowel het trainen van een AI-systeem, als het gebruiken van een AI-systeem. Bij het gebruiken, is de vraag dan hoeveel persoonsgegevens of parameters effectief nodig zijn om de gewenste conclusies te kunnen trekken. Een effectieve toepassing van het verminderen van de verwerkte persoonsgegevens is te zorgen dat de invulvelden die een kandidaat of lead moet invullen, enkel die velden omvatten die echt nodig zijn.

3. BEWAAR PERSOONSGEGEVENS NIET LANGER DAN NODIG

De persoonsgegevens die na voormelde toetsen verwerkt kunnen worden, mogen **niet langer dan noodzakelijk** verwerkt worden. Dit is van groot belang om te vermijden dat persoonsgegevens onbeperkt in de tijd gecumuleerd worden. Daardoor zouden grote risico's ontstaan voor de betrokken personen, zeker wanneer zich een gegevenslek zou voordoen.

Gelet op het belang van deze verplichting, is dit in de AVG in een afzonderlijk principe gegoten, namelijk dat van **opslagbeperking**.[85]

4. VERNIETIG OF ANONIMISEER PERSOONSGEGEVENS ZODRA MOGELIJK

Wanneer het niet langer noodzakelijk is om persoonsgegevens aan de betrokken personen te kunnen koppelen, moeten ze **vernietigd** worden. Dit kan door ze effectief **integraal te vernietigen (wissen)** op alle plaatsen waar ze bewaard worden. Daarbij kunnen eveneens maatregelen worden genomen zodat ze ook in **back-ups niet onbeperkt opgeslagen blijven** en minstens, indien dat onmogelijk zou zijn door een wettelijke verplichting, zeer moeilijk toegankelijk worden gemaakt.

In de praktijk zullen datasets vaak **geanonimiseerd** worden, wat vanuit het oogpunt van de AVG een vorm van vernietiging is.[86] Door degelijke anonimisering houden de persoonsgegevens op persoonsgegevens te zijn en is de AVG hierop niet langer van toepassing.[87]

Er bestaan **verschillende methodes** die toelaten om persoonsgegevens te anonimiseren. Elk hebben ze hun voor- en hun nadelen, reden waarom vaak een combinatie van verschillende methodes toegepast wordt. Een degelijk anonimiseringsbeleid vereist dat de gebruikte technieken/methodes aangepast worden aan **de stand van de techniek** en dat **her-identificatietests gebeuren**, om de kwaliteit van de anonimisering te verzekeren. Aangezien een aantal van deze technieken een vorm van aggregatie met zich meebrengen en elke anonimisering als gevolg heeft dat bepaalde gegevens verwijderd worden, zullen de accuraatheid en de omvattendheid van de dataset hierdoor in veel gevallen verminderen.

Datasets anonimiseren kan door gebruik van ***open source* applicaties** zoals ARX[88] of Amnesia[89] of één van de vele applicaties beschikbaar op de markt. Belangrijk bij de keuze van anonimiseringssoftware is dat zoals hierboven beschreven verschillende methodes naast elkaar toegepast worden en dat her-identificatietests uitgevoerd kunnen worden. Dit uiteraard naast de primaire vereiste dat een kwalitatief eindresultaat bereikt wordt.

[85] Art. 5.1, e) AVG. Zie hierna deel 5.2. over opslagbeperking.
[86] Zie ook "3.3. Wat is het belang van het onderscheid tussen anonimisering en pseudonimisering?"
[87] Zie ook "3.3. Wat is het belang van het onderscheid tussen anonimisering en pseudonimisering?"
[88] Zie: https://arx.deidentifier.org.
[89] Zie: https://amnesia.openaire.eu.

Er bestaan verschillende methoden om te meten of een dataset afdoende geanonimiseerd werd, zoals het concept van "k-anonimiteit".[90]

Hieronder worden **enkele methodes** toegelicht die apart of samen gebruikt worden om persoonsgegevens te anonimiseren en worden eveneens enkele voor- en nadelen opgesomd.[91] Bij elk van deze technieken is vereist dat de **wijzigingen definitief zijn** en de geanonimiseerde gegevens niet meer aan de oorspronkelijke dataset gekoppeld kunnen worden. Zo niet, is er geen sprake van anonimisering, maar van pseudonimisering.

***ATTRIBUTE SUPPRESSION* OF HET VERWIJDEREN VAN ALLE MELDINGEN VAN EEN BEPAALD TYPE (IDENTIFICERENDE) EIGENSCHAPPEN**	
Concept Dit is de meest eenvoudige vorm van anonimisering. De informatie die geacht wordt tot identificatie te leiden, wordt gewist. **Voorbeeld** Uit een lijst met voornaam, familienaam en een testscore wordt de familienaam gewist. Alleen de voornaam en de testscore blijven over, waardoor geen (rechtstreekse) identificatie meer mogelijk is.	**Voordelen** De overige attributen en eigenschappen blijven ongewijzigd. **Nadelen** Risico op her-identificatie: • mogelijk op basis van andere eigenschappen; • door combinatie met verschillende datasets.
***RECORD SUPPRESSION* OF HET VERWIJDEREN VAN GEGEVENS MET BETREKKING TOT BEPAALDE OPVALLENDE BETROKKENEN**	
Concept Hierbij worden alle gegevens van personen die buiten bepaalde grenzen vallen gewist. Dit om te voorkomen dat degenen die afwijken van het gemiddelde gemakkelijk geïdentificeerd kunnen worden. **Voorbeeld** Na een test worden resultaten zonder naam, maar met vermelding van woonplaats publiek bekend gemaakt.	**Voordelen** De overige attributen en eigenschappen blijven ongewijzigd. **Nadelen** De gegevens van personen die afwijken van de norm worden verwijderd, zodat de dataset niet langer accuraat is. Het gemiddelde en de mediaan worden er bijvoorbeeld door beïnvloed.

[90] Hiernaar wordt meestal verwezen met de Engelstalige term 'k-anonimity'. Zie bijvoorbeeld R. Shokri, C. Troncoso, C. Diaz, J. Freudiger en J-P Hubaux, "Unraveling an Old Cloak: k-anonymity for Location Privacy", in: K. Frikken (ed.), Proceedings of the 9th ACM workshop on Privacy in the electronic society (WPES 2010), 2010, p. 115-118.
[91] Zie voor meer informatie bijvoorbeeld Personal Data Protection Commissioner Singapore, "Guide To Basic Data Anonymisation Tech-niques", 25 januari 2018, 39p.

De meeste deelnemers komen uit gemeenten in de buurt van waar de test afgenomen werd, maar sommige deelnemers komen uit andere gemeenten. Indien de punten van deelnemers uit andere gemeenten gepubliceerd worden, kunnen anderen die weten dat zij deelnamen gemakkelijk identificeren welke score zij behaalden. Daarom worden de scores van deze deelnemers niet mee gepubliceerd.	Risico op her-identificatie: • mogelijk op basis van andere eigenschappen; • door combinatie met verschillende datasets.
CHARACTER MASKING OF HET VERWIJDEREN VAN BEPAALDE TEKENS	
Concept Bepaalde tekens worden gemaskeerd om identificatie te voorkomen. **Voorbeeld** Testresultaten worden per postcode gepubliceerd. Uit sommige postcodes komen slechts enkele deelnemers, zodat zij geïdentificeerd zouden kunnen worden. Hun postcodes verschillen echter enkel in de laatste 2 cijfers van andere postcodes. Als dan enkel de eerste 2 cijfers van de postcodes bekend gemaakt worden, dan is wel nog zichtbaar uit welke regio de deelnemers afkomstig zijn, maar niet meer uit welke stad/ gemeente. De postcodes van deelnemers uit Gent (9000) en Gentbrugge (9050) worden dan weergegeven als “90xx” of als “9xxx”. Deelnemers uit Leuven (3000) en Kessel-Lo (3010) worden weergegeven als “30xx” of “3xxx”.	**Voordelen** De overige attributen en eigenschappen blijven ongewijzigd. **Nadelen** De accuraatheid vermindert. Risico op her-identificatie: • mogelijk op basis van andere eigenschappen; • door combinatie met verschillende datasets.
GENERALISERING	
Concept De gegevens worden omgezet naar bepaalde categorieën, zodat er geen individuele gegevens meer beschikbaar zijn.	**Voordelen** Er staan geen exacte gegevens meer in de dataset en de kans op her-identificatie wordt kleiner.

Voorbeeld Testresultaten worden gepubliceerd, maar in plaats van de leeftijd, de woonplaats en de score, worden gepubliceerd: • de leeftijdscategorie: "26-30j" in plaats van 27j; • de provincie: "Vlaams-Brabant" in plaats van "Leuven" • de scoregroep: "76-80%" in plaats van "77%".	**Nadelen** De accuraatheid vermindert. Risico op her-identificatie: • mogelijk op basis van andere eigenschappen, bijvoorbeeld bij uitschieters; • door combinatie met verschillende datasets.
SWAPPING, SHUFFLING OF HET 'DOOREENSCHUDDEN' VAN GEGEVENS	
Concept Alle persoonsgegeven blijven in de set, maar ze worden gewisseld zodat een geheel aan samenhorende gegevens, niet meer van dezelfde persoon afkomstig is. **Voorbeeld** In een dataset staan de naam, voornaam, postcode, geboortejaar en het jaarlijks inkomen van de betrokken personen. Deze worden zo herschikt dat de verschillende gegevens die aan één persoon toebehoorden, niet meer samen staan. (zie volgende pagina) Indien de volgende gegevens deel uitmaken van de set: • Janssens, Lenka, 9000, 1987, 30.000 EUR; • Achmar, Petra, 3000, 1979, 35.000 EUR; • Duchateau, Kim, 1000, 1992, 25.000 EUR. Dan zien de gegevens er vervolgens als volgt uit: • Janssens, Petra, 9000, 1987, 35.000 EUR; • Achmar, Kim, 1000, 1979, 30.000 EUR; • Duchateau, Lenka, 3000, 1992, 25.000 EUR. Alle gegevens zijn nog aanwezig in de dataset, maar er valt niet meer uit op te maken welke gegevens samen horen.	**Voordelen** Alle gegevens worden behouden. **Nadelen** Enkel mogelijk voor zover het beoogde gebruik toelaat om de verschillende parameters te mengen. Risico op her-identificatie: • mogelijk op basis van andere eigenschappen, bijvoorbeeld bij uitschieters; • door combinatie met verschillende datasets.

5. VERMINDER DE LEESBAARHEID EN DE BRUIKBAARHEID VAN DATASETS VOOR DERDEN

Verschillende **technieken** kunnen toegepast worden om de l**eesbaarheid en de bruikbaarheid van persoonsgegevens door derden te verminderen**. Hierdoor wordt het moeilijker om de betrokken personen en/of bepaalde persoonsgegevens in een dataset te identificeren. Bijgevolg vermindert het risico op schade voor de betrokkenen personen, bijvoorbeeld bij een lek van deze gegevens.

Hieronder bespreken we enkele nuttige **gegevensbeschermings-bevorderende technieken**, ook wel PET's genoemd naar hun Engelstalige benaming *Privacy Enhancing Techniques*, om de leesbaarheid en de bruikbaarheid van datasets te verminderen.

Alle hierboven vermelde methodes om persoonsgegevens te anonimiseren, zijn eveneens PET's. Wanneer ze 'onvolmaakt' toegepast worden en dus geen echte anonimisering bereikt wordt, dragen ze wel nog bij aan de vermindering van de bruikbaarheid van de gegevens.

DIFFERENTIËLE PRIVACY (*DIFFERENTIAL PRIVACY*) OF VERSTORING (*PERTURBATION*)	Toepassen van differentiële gegevensbescherming betekent dat 'ruis' of noise toegevoegd wordt aan een dataset, zodat het moeilijk wordt om te bepalen welke persoonsgegevens echt zijn en welke niet.
PSEUDONIMISERING	De AVG schuift pseudonimisering uitdrukkelijk naar voor als een essentiële techniek voor zowel minimale gegevensverwerking, gegevensbescherming door ontwerp, als voor beveiliging van de verwerking.[92] Zoals eerder besproken, wordt het door sommigen als een beperkte vorm van versleuteling beschouwd.[93] Het toepassen van pseudonimisering betekent dat de originele dataset op de achtergrond kan blijven bestaan, maar moeilijker toegankelijk gemaakt wordt. Bij het raadplegen kan dan, door bepaalde parameters te verbergen, enkel een gepseudonimiseerde versie getoond worden aan de gebruiker. De pseudonimisering heeft dan geen effect op de eigenlijke dataset, enkel op de raadpleegbaarheid daarvan. Pseudonimiseer persoonsgegevens zodra dit mogelijk is en bewaar de sleutels op een aparte locatie.
ONLEESBAAR MAKEN VAN PERSOONSGEGEVENS	Persoonsgegevens worden onleesbaar gemaakt voor mensen, maar blijven leesbaar voor de computers die ze gebruiken. Dit is bijvoorbeeld het geval wanneer persoonsgegevens opgeslagen worden in *feature vectors*[94]. Dit sluit identificatie en herkenning van de persoonsgegevens niet uit, maar bemoeilijkt dit wel. Afhankelijk van de wijze van toepassing, kan deze techniek aansluiten bij pseudonimisering en versleuteling, mede afhankelijk van de gangbaarheid of de universaliteit van de gebruikte codering.

[92] Zie onder meer art. 4.5, 6.3, 25.1, 32.1, 40.2 en 89.1 AVG.
[93] Zie ook "3.3. Wat is het belang van het onderscheid tussen anonimisering en pseudonimisering?"
[94] *Feature vectors* zijn vectoren die de karakteristieken bevatten van een parameter. Zo worden bijvoorbeeld de biometrische gegevens nodig voor gezichtsherkenning opgeslagen in de wiskundige weergave hiervan. Deze vectoren laten nog steeds identificatie toe en zijn persoonsgegevens, maar ze zijn niet betekenisvol wanneer waargenomen met het blote oog.

VERSLEUTELING EN HOMOMORFE VERSLEUTELING	Persoonsgegevens worden omgezet in een formaat dat niet leesbaar is zonder de vereiste sleutel. Ook versleuteling wordt herhaaldelijk uitdrukkelijk vermeld in de AVG, onder meer bij de beveiliging van de verwerking[95] en bij gegevenslekken[96]. Versleutelde gegevens zijn nog steeds persoonsgegevens aangezien identificatie en herkenning van de persoonsgegevens mogelijk zijn, mits toepassing van de sleutel. Bij de toepassing van homomorfe versleuteling is het mogelijk om de dataset niet alleen versleuteld te bewaren, maar om een AI-systeem rechtstreeks op de versleutelde gegevens te trainen, zonder dat het AI-systeem toegang heeft tot de niet-versleutelde gegevens.[97]

6. VERMINDER HET VOLUME AAN (CENTRAAL) VEREISTE PERSOONSGEGEVENS IN DE TRAININGSFASE
Er zijn ook gegevens-bevorderende technieken om het volume aan (centraal) vereiste datasets te verminderen in de **trainingsfase**.

MINIMALE GEGEVENSVERWERKING ALS PRODUCTVEREISTE	Leg minimale gegevensbescherming op als productvereiste bij aankoop of opdracht tot ontwikkeling van een AI-systeem waarmee zelf nog getraind zal worden.
SELECTIE VAN DE PARAMETERS IN EEN DATASET	Evalueer welke aanwezige parameters in een dataset noodzakelijk zijn voor het trainingsproces en verwijder de andere parameters vooraleer de dataset te gebruiken.
GEFEDEREERD LEREN (*FEDERATED LEARNING*)	Gefedereerd leren laat toe om op verschillende databases op lokale toestellen te trainen, zonder dat de persoonsgegevens het lokale toestel verlaten. Een AI-systeem traint dan binnen verschillende data-omgevingen waar zich persoonsgegevens bevinden. Enkel de geleerde inzichten verlaten de lokale omgeving. De persoonsgegevens worden niet gedeeld met het centrale AI-systeem. Deze techniek wordt bijvoorbeeld toegepast bij tekstvoorspelling op smartphones. Op elke smartphone wordt de werking van het AI-systeem getraind. De input waarop getraind wordt, verlaat de toestellen niet. De inzichten worden wel gedeeld, zodat het centrale AI-systeem en de werking ervan over alle toestellen gespreid verbeterd worden.

[95] Art. 32.1, a) AVG.
[96] Art. 34.4, a) AVG.
[97] Voor meer informatie over versleuteling en de toegepaste technieken, zie "4.3. Wat zijn aandachtspunten voor de beveiliging van de verwerking van persoonsgegevens door AI-systemen?"

	Deze techniek wordt ook toegepast bij medisch onderzoek waarbij verschillende patiëntenbestanden gebruikt moeten worden om te trainen. Een aandachtspunt bij gefedereerd leren is dat uit de gedeelde inzichten geen persoonsgegevens afgeleid mogen kunnen worden en geen (her-)identificatie van de betrokken personen mogelijk mag zijn. Dit risico wordt groter naargelang het model complexer is en met gerichtere fijnmazigere parameters gewerkt wordt.
GENERATIVE ADVERSERIAL NETWORKS (GAN'S)	Door gebruik te maken van een GAN of *Generative Adverserial Network* wordt het vereiste volume aan (persoons-)gegevens voor training van een AI-systeem verminderd. Hierbij worden twee neurale netwerken getraind. Een netwerk doet dienst als generator, het ander netwerk als *discriminator.* De generator zal proberen om op basis van een gegevensset nieuwe gegevens te genereren, die lijken op de initiële gegevens. De *discriminator* zal proberen om de echte gegevens, namelijk deze uit de oorspronkelijke gegevensset, te onderscheiden van de gegevens gegenereerd door de generator. De generator zal trachten om de *discriminator* te misleiden en proberen om nieuwe gegevens aan te maken die de discriminator niet kan onderscheiden van de oorspronkelijke gegevens. Daarbij zal de generator de kwaliteit van de output verbeteren op basis van de feedback van de *discriminator.* De *discriminator* verbetert zijn capaciteit om de niet-originele gegevens te herkennen. Hoewel GAN's nog steeds vrij grote hoeveelheden gegevens vereisen om getraind te worden, staan ze toch toe om een AI-systeem te trainen met een beperktere initiële gegevensset, aangevuld met gegevens bekomen via een GAN. Bij gebruik van te kleine gegevenssets bestaat het risico dat de gegevens onvoldoende divers zijn en dat daardoor inherente redeneerfouten of vooroordelen (*bias*) in het systeem sluipen.
TRANSFER LEARNING	Bij overdrachtsleren ('*transfer learning*') wordt een AI-systeem niet op (persoons-)gegevens getraind, maar leert een AI-systeem van een ander reeds getraind AI-systeem. Het systeem zelf verwerkt dus geen persoonsgegevens meer, maar ergens in de keten hogerop zal wel een training gebeurd zijn met persoonsgegevens. Hierbij is uiteraard van belang dat het bestaand AI-systeem waarvan geleerd wordt betrouwbaar is en geen bias bevat.

7. VERMINDER HET VOLUME AAN (CENTRAAL) VEREISTE PERSOONSGEGEVENS IN DE GEBRUIKSFASE
Tot slot zijn er ook nog een aantal mogelijkheden om het volume aan (centraal) vereiste persoonsgegevens in de **gebruiksfase** te verminderen zoals:

- minimale gegevensverwerking als productvereiste;
- minimale gegevensbescherming als productvereiste bij aankoop of opdracht tot ontwikkeling van een AI-systeem;
- persoonsgegevens-beschermende verzoeken (*privacy preserving queries* of P2Q).

4.2.E. Een performant en afgedwongen databeleid: beperk de toegang tot persoonsgegevens

Het is ook aangewezen om een performant databeleid op te zetten waarbij een **duidelijke en logische rolbepaling** er onder meer voor zorgt dat:

- zo weinig mogelijk mensen toegang hebben tot de 'rauwe' persoonsgegevens;
- toegang enkel met individuele accounts kan gebeuren en gelogd wordt;
- bij de raadplegingen waarbij geen toegang tot (alle) 'rauwe' persoonsgegevens vereist wordt, de getoonde persoonsgegevens beperkt en/of visueel gepseudonimiseerd worden.

Toepassing
Bij zowel e-commerce als rekrutering kan gezorgd worden dat een persoon die een klant moet opbellen voor een afspraak, standaard enkel toegang heeft tot de gegevens die daarvoor nodig zijn, bijvoorbeeld de naam, het telefoonnummer en de reden van de afspraak.

4.3. Wat zijn aandachtspunten voor de beveiliging van de verwerking van persoonsgegevens door AI-systemen?

Essentie
Persoonsgegevens kunnen enkel goed beschermd worden indien er maatregelen worden genomen om hun integriteit en vertrouwelijkheid te waarborgen. Deze maatregelen moeten technisch zijn ter bescherming van de infrastructuur waarop de gegevens verwerkt worden en van de gegevens zelf. Daarnaast zijn ook organisatorische maatregelen vereist die moeten zorgen dat de personen binnen een organisatie de vereiste maatregelen correct toepassen, zelf op de juiste manier met persoonsgegevens omgaan en zich bewust zijn van het belang van gegevensbescherming.
Actiepunten
◊ Stel een gegevensveiligheidsbeleid op. ◊ Stel, indien dat er nog niet zou zijn, een register van verwerkingsactiviteiten op en zorg op basis daarvan voor een doordacht beheer van persoonsgegevens en een evaluatie van de datastromen. ◊ Neem de nodige technische maatregelen om de ICT-infrastructuur en de persoonsgegevens te beschermen tegen zowel intentionele als toevallige of onbedoelde bedreigingen. ◊ Voorzie een performante toegangscontrole en authenticatie voor ICT-systemen, specifieke omgevingen waar persoonsgegevens aanwezig zijn en gebouwen. ◊ Implementeer een afdwingbaar rol- en machtigingsbeleid dat bepaalt wie toegang krijgt tot welke omgeving en welke persoonsgegevens.

◊ Monitor de ICT-omgeving en de toegang tot persoonsgegevens.
◊ Informeer en sensibiliseer personeel binnen de organisatie en stel bindende richtlijnen op voor het gebruik van de ICT-infrastructuur en de omgang met persoonsgegevens.
◊ Maak duidelijke afspraken met leveranciers en verwerkers over veiligheid en naleving van de AVG.
◊ Documenteer alle geleverde inspanningen en daarbij gemaakte afwegingen.

4.3.A. Technische en organisatorische beveiliging van de gehele omgeving waarin persoonsgegevens verwerkt worden

Persoonsgegevens kunnen enkel goed beschermd worden als ze **effectief beveiligd** worden. Het heeft geen zin om regels, richtlijnen of een beleid op te stellen, als eenieder die dat wenst zich toegang kan verschaffen tot de persoonsgegevens. De 'beveiliging van de verwerking' is één van de hoekstenen van de AVG.

Daarmee wordt enerzijds op **technische maatregelen** gedoeld zoals het voorzien van encryptie, een firewall of paswoordcontrole. Anderzijds zijn ook **organisatorische maatregelen** vereist zoals het opleggen van bepaalde verplichtingen aan personeel en onderaannemers. Deze maatregelen moeten voorkomen dat de verwerkte persoonsgegevens ongewenst per ongeluk of onrechtmatig:

- worden gedeeld met of blootgesteld aan derden of personen die hier geen toegang toe behoren te hebben (al dan niet te kwader trouw);
- verloren gaan of worden vernietigd;
- worden gewijzigd.

Ook hier dienen deze maatregelen rekening te houden met **de stand van de techniek, de uitvoeringskosten, de context en de risico's** voor de personen wiens gegevens verwerking worden. Ook beveiliging is dus een dynamische verplichting die **risico- en context-gebaseerd** en die ook binnen eenzelfde organisatie kan evolueren.[98]

De AVG bepaalt dat de beveiligingsmaatregelen, indien van toepassing, het volgende moeten omvatten:[99]

- pseudonimisering van de persoonsgegevens;
- versleuteling van de persoonsgegevens;
- maatregelen die toestaan om voortdurend de vertrouwelijkheid, integriteit, beschikbaarheid en veerkracht van de systemen te garanderen;[100]
- maatregelen die toestaan om bij een incident de beschikbaarheid van de persoonsgegevens 'tijdig' te herstellen;
- regelmatige evaluatie en test van de effectiviteit van de genomen maatregelen.

Het aansluiten bij een **goedgekeurde gedragscode** of het laten **certificeren** van bepaalde processen kan gebruikt worden als element om aan te tonen dat de verwerking beveiligd is. Ook dan moet nog in concreto naar de adequaatheid van de genomen maatregelen gekeken worden.

De verwerkingsverantwoordelijke en de verwerker zijn verantwoordelijk voor de beveiliging van de verwerking door eenieder die onder hun gezag of in hun opdracht persoonsgegevens verwerkt.

[98] Art. 32.1 en 32.2 AVG.
[99] Art. 32.1, a) tot d) AVG.
[100] Hier wordt verwezen naar de zogenaamde CIA-triad, die staat voor Confidentiality, Integrity and Availability.

4.3.B. Risico-gebaseerde benadering

Ook de beveiliging van de verwerking kan niet los bekeken worden van het risico-gebaseerde karakter van de AVG, zoals benadrukt in de hierop betrekking hebbende bepalingen.[101]

Door de toegang te beperken, de persoonsgegevens af te schermen van de buitenwereld en de gegevens onbruikbaar te maken voor derden wordt ook hier het risico verminderd dat een te verregaande kijk in het privéleven van de betrokkenen mogelijk is.

Beveiliging van de verwerking is ook de belangrijkste verplichting om zowel het risico op een **gegevenslek als de schade bij een dergelijk lek te verminderen**. Dit zowel voor wat betreft gegevenslekken die het gevolg zijn van onbedoelde als van kwaadwillig acties, als gegevenslekken veroorzaakt door enerzijds interne en anderzijds externe dreigingen. Indien zich toch een lek van deze gegevens zou voordoen, worden zowel het aantal getroffen personen, de bruikbaarheid van de gelekte gegevens persoonsgegevens en de omvang van de schade voor deze personen aanzienlijk beperkt.

4.3.C. Gegevensveiligheidsbeleid

Om de beveiliging van de verwerking degelijk te organiseren in een organisatie die AI-systemen ontwikkelt of ermee werkt is het noodzakelijk dat een **gegevensveiligheidsbeleid** opgemaakt wordt. Hiervoor dient in eerste instantie een grondige interne analyse te gebeuren van de risico's en de vereisten en moet worden bepaald hoe deze verminderd, dan wel geïmplementeerd kunnen worden. Zoals hierboven besproken, betreft het hier zowel organisatorische als technische risico's en maatregelen.

De belangrijkste **(onderling samenhangende) domeinen** zijn:

VERSLEUTELING EN PSEUDONIMISERING
Gegevens worden onleesbaar gemaakt wanneer onbevoegden er toegang toe krijgen of wanneer opslagmedia (zoals USB-opslag, harde schijven of laptops) in onbevoegde handen vallen.
CONFIDENTIALITEIT
Er wordt bepaald wie waartoe toegang moet hebben. Ook de gepaste toegangscontrole en -beperkingen worden opgezet (zowel fysiek als elektronisch). De confidentialiteit moet zowel worden gegarandeerd ten aanzien van onbevoegde internen als ten aanzien van externen.
INTEGRITEIT TEGENOVER FOUTEN
De integriteit van gegevens kan onopzettelijk bedreigd worden door fouten van medewerkers, systemen of externen. Deze fouten moeten opgespoord en voorkomen worden.
INTEGRITEIT TEGENOVER OPZET
De integriteit kan ook opzettelijk bedreigd worden, wanneer personen of systemen de data te kwader trouw onbeschikbaar willen maken, vernietigen of aanpassen.

[101] Art. 32.1 en 32.2 AVG.

VEERKRACHT EN WEERBAARHEID TEGENOVER INCIDENTEN
Bij incidenten kan de goede werking van het netwerk en de infrastructuur (zoals netwerk, servers, laptops of gebouwen) verstoord worden. De beschikbaarheid voor het personeel kan onderbroken worden. Daarom moeten er maatregelen worden genomen om de verstoring van de werking van een organisatie te beperken bij incidenten, bijvoorbeeld door het voorzien van back-ups, reservestructuren, business continuity plannen en incident-response plannen.
INCIDENT RESOLUTIE
Incidenten zijn nooit volledig uit te sluiten. Een organisatie moet in staat zijn om snel op incidenten te reageren, deze op te lossen en, zo nodig, de vereiste kennisgevingen te doen indien zich daarbij een gegevenslek voordoet.
EVALUATIE
De adequaatheid van veiligheidsmaatregelen moet geëvalueerd en proactief gemonitord worden. Leer uit fouten en oefeningen en doe aan continue verbetering.

Een gegevensveiligheidsbeleid opstellen vereist een **multidisciplinaire en interdepartementale aanpak** en ondersteuning van het management. Op basis van het **intern gegevensveiligheidsbeleid** kan ook een **extern gegevensveiligheidsbeleid** opgesteld worden: dit zijn één of meerdere documenten waarin de (niet-vertrouwelijke) speerpunten van het gegevensveiligheidsbeleid worden toegelicht. Deze kunnen dienen om leveranciers en partners duidelijk te maken wat op dit vlak verwacht wordt en ook om zowel klanten, de overheid als het publiek te informeren.

Hierna worden een aantal maatregelen toegelicht die de bovenvermelde domeinen transversaal helpen waarborgen.

4.3.D. Databeheer en *datamapping*

Om een efficiënt gegevensveiligheidsbeleid voor persoonsgegevens op te zetten, is het noodzakelijk om te weten **welke types en volumes** gegevens verwerkt worden, **hoe de datastromen lopen**, **welke verwerkingen** ermee gebeuren en **wie toegang nodig heeft** tot welke gegevens. Een degelijk opgemaakt **register van verwerkingsactiviteiten**[102] is hiervoor noodzakelijk.

Op basis van deze informatie kunnen de optimale gegevensstromen vastgesteld worden, die toelaten om de toegang tot en de bescherming van deze gegevens op de meest efficiënte wijze te organiseren en te monitoren en de blootstelling zoveel mogelijk te beperken. Daarbij moet uiteraard ook rekening worden gehouden met de andere toepasselijke principes zoals minimale gegevensverwerking.[103]

4.3.E. Technische maatregelen

Technische maatregelen zijn **noodzakelijk** om persoonsgegevens adequaat te beschermen.

Hieronder worden enkele maatregelen toegelicht die altijd deel moeten uitmaken van een veiligheids-

[102] Art. 30 AVG.
[103] Zie ook "4.2. Wat houdt de vereiste van minimale gegevensverwerking van persoonsgegevens in voor AI-systemen?"

beleid in een organisatie die persoonsgegevens verwerkt door middel van een AI-systeem. Gelet op de niet-technische aard van deze bijdrage worden de begrippen geduid zonder verder in te gaan op de technische eigenschappen en mogelijkheden.

Bij elk van de onderstaande technieken is het steeds noodzakelijk om rekening te houden met de **stand der techniek** en de **gangbare marktgebruiken** om te bepalen welke technieken het meest aangewezen zijn om een adequate bescherming te bieden.

OPSLAG VAN PERSOONSGEGEVENS	**Intelligente organisatie van de opslag** Zorg dat persoonsgegevens en gevoelige gegevens logisch en gebruiksvriendelijk opgeslagen worden op de daarvoor voorziene plaats en dat geen onnodige kopieën gemaakt (kunnen) worden. **Versleutel persoonsgegevens** Door persoonsgegevens te versleutelen, wordt het risico op schade bij een gegevenslek aanzienlijk verminderd omdat derden er in principe geen kennis van kunnen nemen. Aangezien versleutelingsstandaarden aan sterke inflaties en evoluties onderhevig zijn, is het zeer belangrijk om hier rekening mee te houden. In elk geval is het aangeraden om zowel servers, individuele computers, als elk ander opslagmedium (zoals USB-sticks en smartphones) te versleutelen en te vereisen dat binnen de organisatie enkel versleutelde toestellen gebruikt mogen (/kunnen) worden, ook indien gebruikers eigen toestellen mogen gebruiken (‘*bring your own device*’). **Back-up** Zorg voor regelmatige en bruikbare back-ups, die op afzonderlijke locaties bewaard worden en die niet rechtstreeks met het eigen netwerk in verbinding staan. Test op regelmatige tijdstippen of de back-ups effectief gemaakt worden en bruikbaar zijn. **Markeer persoonsgegevens (‘*data tagging*’)** Markeer (‘tag’) persoonsgegevens ter bevordering van het monitoren van de datastromen en het nemen van de juiste acties op het juiste moment. **Zet een dataretentiebeleid op** Zet een waar mogelijk afgedwongen dataretentiebeleid op in overeenstemming met de vereisten van de minimale gegevensverwerking en de opslagbeperking.[104]

[104] Zie ook "4.2. Wat houdt de vereiste van minimale gegevensverwerking van persoonsgegevens in voor AI-systemen?"

TOEGANG TOT SYSTEMEN EN GEGEVENS	**Voorzie een performante toegangscontrole en authenticatie** Digitale toegang ter plaatse en vanop afstand tot zowel de ICT-omgeving, als persoonsgegevens als de gebouwen waar prestaties geleverd worden, moet gebeuren door middel van individuele authenticatie door de gebruiker. Dit laat toe om in functie van de rolvereisten te bepalen wie toegang krijgt tot welke gegevens. Toegang tot de eigenlijke persoonsgegevens wordt beperkt tot personen die daar effectief toegang toe moeten hebben ('*need-to-know*'), met individuele toegang. Anderen hebben slechts toegang tot gepseudonimiseerde of geanonimiseerde gegevens. Dit voorkomt dat onbevoegden (on)gewenst toegang krijgen tot gegevens, beperkt de risico's bij zowel externe (bijvoorbeeld *hacking* of *phishing*) als interne dreigingen (bijvoorbeeld een ontevreden werknemer die bedrijfsdata kopieert). Ook kunnen potentieel verdacht gedrag en oorzaken van gegevenslekken hiermee sneller gedetecteerd worden. Essentieel hierbij is dat een veilig en afgedwongen authenticatiebeleid toegepast wordt waarbij (i) unieke, sterkeen regelmatig wijzigende paswoorden vereist worden, (ii) twee- of drie-factorauthenticatie toegepast wordt waar nuttig, (iii) de authenticatievereisten strenger worden naarmate het toegangsniveau van de gebruiker en (iv) bijkomende authenticatie vereist kan zijn wanneer een gevoelige omgeving betreden wordt of vanop afstand ingelogd wordt. Ook de fysieke toegangscontrole mag hier niet vergeten worden: enerzijds dient uiteraard de toegang tot de werkomgeving voor externen afgesloten te zijn, anderzijds moet de toegang tot bijvoorbeeld plaatsen waar lokale servers staan beperkt worden. Zorg hierbij dat toegangsmachtigingen onmiddellijk ingetrokken worden wanneer een werknemer of dienstverlener de organisatie verlaat. **Scherm systemen af van de buitenwereld** Bescherm de netwerkomgeving tegen externe toegang door derden en onbevoegden, met detectie en de mogelijkheid tot het automatisch treffen van bewarende en beschermende maatregelen. **Vernietig fysieke (papieren) en digitale opslagmedia**

MONITOR DE NETWERKOMGEVING EN DE GEBOUWEN	**Monitor actief de ICT-omgeving en de gebouwen op:** • (pogingen tot) veiligheidsinbreuken, zowel via het netwerk als via andere kanalen (bijvoorbeeld e-mail); • verdachte toegang, bijvoorbeeld een toegang vanuit een ander land; • verdacht gedrag, bijvoorbeeld toegang op een atypisch ogenblik of herhaaldelijke mislukte authenticatiepogingen; • grote dataverplaatsingen; • kwaadwillige applicaties zoals virussen en malware; • gegevenslekken. Hier moet bijzondere aandacht aan besteed worden zodat de juiste personen gewaarschuwd worden en zo nodig snel aan de meldingsplicht voldaan kan worden. In functie van het risico gekoppeld aan bepaalde incidenten kunnen hier dan automatisch gevolgen aan gekoppeld worden zoals (i) het (opnieuw) vereisen om in te loggen of de tweede authenticatiefactor op te geven, (ii) het weergeven van een waarschuwingen (bijvoorbeeld bij het downloaden van persoonsgegevens), (iii) bepaalde diensten binnen de organisatie waarschuwen, (iv) het beperken van de toegang van de betreffende gebruiker of (v) het stilleggen van bepaalde processen.
COMMUNICATIE	**Beveiliging** Beveilig e-mailing en andere communicatie met versleuteling, spam- en *phishing*bescherming. Laat persoonsgegevens enkel delen via een eigen platform en via tijdelijke en/of gepersonaliseerde hyperlinks. **Verbinding** Zorg ook dat communicatie en inloggen op de systemen steeds gebeurt via veilige verbinding.
SOFTWAREBELEID	Voorzie een afgedwongen softwarebeleid zodat enkel vertrouwde en goedgekeurde software gebruikt wordt op systemen met toegang tot het netwerk. Zorg dat software steeds up-to-date is.

4.3.F. Bewustmaking en opleiding van al wie toegang heeft tot persoonsgegevens

Om de verwerking van persoonsgegevens te beveiligen, volstaat het niet om technische maatregelen te nemen. Deze maatregelen moeten ook **begrepen en gedragen** worden door degenen die ze uitvoeren. Het heeft immers geen zin om vijf sloten op een deur te zetten, indien de deur in de praktijk altijd open blijft staan. Wanneer hierna naar personeelsleden verwezen wordt, moet dit begrepen worden als iedereen die binnen de organisatie persoonsgegevens verwerkt en/of toegang heeft tot de netwerkomgeving, dus ook externe consulenten en bestuurders binnen de organisatie.

Hierbij zijn vooral de hierna volgende aandachtspunten van belang.

TOEGEWEZEN ROLLEN EN TOEGANGS-MACHTIGINGEN	Zorg dat de toegang tot systemen, processen en persoonsgegevens aansluit bij de taken die de betreffende personen moeten vervullen en dat zij begrijpen waarom deze rechten en beperkingen op die manier toegepast worden.

BELEID	Stel regels vast voor een correct gebruik van en omgang met de ICT-infrastructuur en persoonsgegevens. Zorg dat deze aan de arbeidsrechtelijke vereisten voldoen om afdwingbaar te zijn. De volgende richtpijnen zijn daarvoor nodig (al dan niet deel uitmakend van één enkel of meerdere documenten): • ICT-beleid: dit bepaalt hoe veilig omgegaan wordt met ICT-apparatuur, paswoorden, internetgebruik, e-mail,… • Data-beleid: dit bepaalt specifiek hoe met persoonsgegevens en gevoelige data omgegaan moet worden. In een context waarin AI-systemen gebruikt worden, moet specifiek aandacht besteed worden aan het gebruik van persoonsgegevens in combinatie hiermee. • BYOD-beleid ('*Bring Your Own Device*' of 'gebruik je eigen toestellen'): dit bepaalt de voorwaarden waaronder eigen toestellen binnen een netwerkomgeving of, algemeen, voor werkdoeleinden gebruikt mogen worden.
OPLEIDING	Voorzie algemene en gerichte opleidingen voor personeelsleden over gegevensveiligheid, gegevensbescherming, ICT-veiligheid en gebruik van persoonsgegevens in een AI-context. De intensiteit en het aantal opleidingen hangt af van de mate waarin de betreffende personeelsleden gegevens verwerken, via hun taken invloed hebben op gegevensverwerking en/of instaan voor ICT-veiligheid.
SENSIBILISERING	Regels worden enkel correct toegepast indien ze gedragen worden door de personen die ze uitvoeren. Zorg daarom dat de personeelsleden ook gesensibiliseerd worden aangaande het belang van gegevensbescherming en zich bewust worden van de risico's. Dit gebeurt onder meer door: • het bestuur en leidinggevenden het goede voorbeeld te laten geven en te laten benadrukken welk belang de organisatie hecht aan gegevensveiligheid; • het voeren van zowel ludieke als ernstige informatiecampagnes, algemeen of per onderwerp/thema; • het opzetten van testcampagnes, bijvoorbeeld door het versturen van eigen *phishing*mails.

4.3.G. Afspraken met leveranciers en verwerkers

Maak **duidelijke afspraken** met leveranciers van goederen, diensten en software die een invloed kunnen hebben op de veiligheid van de infrastructuur en de verwerking van gegevens. Vereis dat zij voldoen aan de AVG-verplichtingen en leg minimale veiligheidsvereisten op. Zorg dat zij aansprakelijk zijn voor door hun veroorzaakte schade, wanneer dat niet het geval zou zijn.

Indien leveranciers gegevens verwerken als verwerker, moet steeds **een verwerkingsovereenkomst**

afgesloten worden.[105]

4.3.H. Documenteer

Ingevolge de algemene 'verantwoordingsplicht' onder de AVG moet ook het nakomen van de verplichting om de gegevensverwerking te beveiligen bewezen kunnen worden.[106] Er moet dus bewezen kunnen worden dat de nodige **organisatorische en technische maatregelen genomen werden én dat ze doeltreffend zijn**.

Daarbij is het noodzakelijk om onder meer te **documenteren**:
- welke technische maatregelen genomen werden en waarom;
- hoe problemen en incidenten uit het verleden in aanmerking genomen werden;
- welke opleidingen gevolgd werden door het personeel en waarom zij deze dienden te volgen;
- welke sensibiliseringscampagnes gevoerd werden; en
- dat met alle leveranciers en verwerkers de nodige afspraken gemaakt en uitgevoerd werden.[107]

4.4. Wanneer moet een DPIA of GEB worden uitgevoerd voor een verwerking van persoonsgegevens door AI-systemen?

Essentie
Een DPIA of GEB is een instrument om vooraf de gegevensbeschermingsrisico's van een gegevensverwerking in kaart te brengen en om daarna maatregelen te kunnen nemen om de geïdentificeerde risico's te verkleinen. AI-systemen omvatten vaak zowel nieuwe technologieën als complexe en onverwachte uitkomsten met betrekking tot persoonsgegevens, waardoor het nodig kan zijn om een DPIA uit te voeren. Rekening houdende met de richting die het Witboek over AI uitgaat,[108] lijkt het aangeraden dat bedrijven die AI-systemen ontwikkelen voor toepassing in de context van rekrutering of e-commerce, best nu al overwegen om DPIAs uit te voeren. Door het uitvoeren van een DPIA kunnen organisaties ook aantonen dat de verwerking van persoonsgegevens door een AI-systeem proportioneel is.

[105] Zie ook "3.5. Wat zijn de verschillende rollen die een organisatie kan vervullen onder de GDPR in een AI-context?"
[106] Art. 5.2 AVG. De verantwoordingsplicht of het accountability principe is één van de basisbeginselen inzake gegevensverwerking. Het houdt in dat verwerkingsverantwoordelijken en verwerkers moeten kunnen aantonen dat zij stappen hebben ondernomen om de verplichtingen uit de AVG na te leven.
[107] Europees Comité voor gegevensbescherming, Guidelines 4/2019 on Article 25 Data Protection by Design and by Default, 13 november 2019, p. 6, nr. 15.
[108] Het Witboek over AI stelt dat: "Het gebruik van KI-toepassingen voor wervingsprocedures en in situaties die gevolgen hebben voor de rechten van werknemers wordt altijd als risicovol beschouwd, gezien de gevolgen voor personen en gezien het EU-acquis inzake gelijke behandeling in arbeid en beroep; in dat geval zijn de onderstaande eisen dus te allen tijde van toepassing. Andere specifieke toepassingen op het gebied van consumentenrechten kunnen in overweging worden genomen".

Actiepunten
◊ Overweeg welke vorm van DPIA de organisatie wenst te implementeren en maak een DPIA-template (zie verder voor een blauwdruk van stappenplan). ◊ Onderzoek welke verwerkingen als een hoog risico beschouwd worden. ◊ Onderzoek of de organisatie een verwerker dan wel een verwerkingsverantwoordelijke is ten aanzien van de beoogde verwerking. ◊ Voer de DPIA steeds zo vroeg mogelijk uit in de levenscyclus van een AI-systeem. ◊ Onderzoek of de verwerking van persoonsgegevens noodzakelijk is en ga na of het gebruik ervan in verhouding is met het uiteindelijke doel.

4.4.A. Algemeen

Elke organisatie die persoonsgegevens verwerkt, moet nagaan of daar risico's aan verbonden zijn. Indien een organisatie vermoedt dat het gebruik van een AI-systeem naar alle waarschijnlijkheid een **hoog risico** inhoudt voor de **rechten en vrijheden van natuurlijke personen**, moet zij een DPIA uitvoeren.[109]

Een DPIA is bedoeld om de **verwerking van persoonsgegevens te beschrijven**, de **noodzaak en evenredigheid ervan te beoordelen** en de **daaraan verbonden risico's** voor de rechten en vrijheden van natuurlijke personen te helpen beheren door deze risico's in te schatten en maatregelen te bepalen om ze aan te pakken.

Hoewel een DPIA enkel verplicht is bij verwerkingen met een **waarschijnlijk hoog risico**, wordt het toch aangeraden deze ook in **andere situaties uit te voeren**. Het is immers een nuttig instrument dat organisaties helpt om aan de wetgeving inzake gegevensbescherming te voldoen.

Er bestaan **verschillende methoden** om een DPIA uit te voeren. Er bestaan geen specifieke vormvereisten, maar de DPIA moet in ieder geval bevatten:

- welke gegevens zullen worden verwerkt, op welke manier ze zullen worden verwerkt en waarom ze zullen worden verwerkt;
- waarom deze verwerkingen noodzakelijk zijn en de proportionaliteit ervan;
- welke maatregelen worden genomen om de risico's aan te pakken (zoals waarborgen en veiligheidsmaatregelen).

Indien de geplande verwerkingen een hoog risico zouden inhouden en de organisatie niet in staat is om maatregelen te voorzien die dit risico beperken, dienen deze verwerkingen v**ooraf aan de GBA te worden voorgelegd**. Dat kan via dit formulier.[110]

Om het risiconiveau te beoordelen, moet een DPIA zowel de **waarschijnlijkheid als de ernst van eventuele gevolgen** voor personen in aanmerking nemen.

4.4.B. Bepalen van het risico

De AVG stelt dat een DPIA vereist is in de volgende **niet-exhaustieve gevallen**:

- systematische en uitgebreide profilering met belangrijke gevolgen;[111]

[109] Art. 35.1 AVG.
[110] Zie ook: https://www.gegevensbeschermingsautoriteit.be/voorafgaandelijke-raadpleging.
[111] Zie ook "5.4. Wat zegt de AVG over geautomatiseerde individuele besluitvorming en profilering en wat is de impact

- grootschalige verwerking van bijzondere categorieën van persoonsgegevens;[112] of
- stelselmatige en grootschalige monitoring van openbaar toegankelijke ruimten[113].

Om te **bepalen of een verwerking een hoog risico inhoudt** en een DPIA dus nodig kan zijn, moeten de volgende criteria in aanmerking worden genomen:

1. EVALUATIE OF BEOORDELING VAN MENSEN OP BASIS VAN PERSOONLIJKE ASPECTEN
Hier gaat het onder meer over profilering en het maken van prognoses, met name van kenmerken betreffende beroepsprestaties, economische situatie, gezondheid, persoonlijke voorkeuren of interesses, betrouwbaarheid of gedrag en locatie of verplaatsingen van de betrokkene. Een voorbeeld is een bedrijf dat informatie verzamelt over de bezoekers van een webshop (bijvoorbeeld aankoopgeschiedenis, surfgedrag en andere online informatie) en op basis daarvan profielen van deze personen opstelt (of laat opstellen door een AI-toepassing), zodat ze automatisch gepersonaliseerde advertenties kunnen aanbieden.
2. GEAUTOMATISEERDE BESLUITVORMING
Dit is het nemen van besluiten met technologische middelen zonder enige menselijke tussenkomst. Profilering kan hier ook onder vallen.[114] Indien een dergelijk besluit zorgt dat iemand er rechtsgevolgen van ondervindt of anderszins aanzienlijk wordt getroffen, is een DPIA verplicht. Rekrutering op basis van louter een algoritme, dus zonder menselijke tussenkomst, zal in ieder geval aanzienlijke gevolgen hebben voor een persoon, onder meer omdat het beslist of een persoon een job wel of niet krijgt en omdat er discriminatie kan optreden.
3. 'GEVOELIGE' GEGEVENS
Het gaat hierbij om bijzondere categorieën van persoonsgegevens[115] zoals informatie over politieke opvattingen, strafrechtelijke feiten, seksuele geaardheid of medische gegevens. Het kan ook gaan om gegevens die over het algemeen als privacygevoelig worden beschouwd zoals gegevens over elektronische communicatie, locatiegegevens en financiële gegevens. Via een rekruteringsproces bestaat het risico dat er gevoelige gegevens verzameld worden zoals de financiële situatie, vakbondslidmaatschap of medische gegevens al dan niet direct (via een sollicitatiegesprek) of indirect (via sociale media).

hiervan op AI-systemen?"

[112] Zie ook "3.2. Wat zijn bijzondere categorieën van persoonsgegevens?"

[113] Art. 35.33 AVG.

[114] Zie ook "5.4. Wat zegt de AVG over geautomatiseerde individuele besluitvorming en profilering en wat is de impact hiervan op AI-systemen?"

[115] Zie ook "5.4. Wat zegt de AVG over geautomatiseerde individuele besluitvorming en profilering en wat is de impact hiervan op AI-systemen?"

4. GROOTSCHALIGE GEGEVENSVERWERKING

De AVG geeft geen definitie van grootschalige gegevensverwerking. De WP29 raadt alvast aan om met de volgende factoren rekening te houden:

- het aantal personen over wie gegevens worden verwerkt;
- de hoeveelheid gegevens en/of de verscheidenheid aan gegevens die wordt verwerkt;
- de duur van de gegevensverwerking;
- de geografische reikwijdte van de gegevensverwerking

5. COMBINEREN VAN GEGEVENS UIT VERSCHILLENDE BRONNEN

Het gaat hierbij over het matchen of samenvoegen van gegevensverzamelingen. Bijvoorbeeld gegevensverzamelingen die voortkomen uit twee of meer verwerkingen voor verschillende doeleinden en/of die door verschillende verwerkingsverantwoordelijken zijn uitgevoerd, op een manier die niet overeenkomt met de redelijke verwachtingen van de betrokkene.

Een online retailer wil bijvoorbeeld het huidige klantenbestand aanvullen met essentiële informatie zodat hij beter kan inspelen op de noden van de klanten en hen gepersonaliseerde advertenties aanbieden. Hij doet hiervoor een beroep op een onderneming die hierin gespecialiseerd is. Door zijn klantenbestand met dat van de derde partij samen te voegen, heeft hij een 'verrijkt' klantenbestand.

6. KWETSBARE BETROKKENEN

Een DPIA kan nodig zijn omdat een ongelijke machtsverhouding bestaat tussen de betrokken personen en de verwerkingsverantwoordelijke. Dit heeft als gevolg dat deze betrokkenen niet in staat zijn om vrije toestemming te geven aan, bezwaar te maken tegen de verwerking van hun gegevens of om hun rechten uit te oefenen. Kwetsbare betrokkenen zijn onder andere kinderen, werknemers, patiënten en bejaarden.

7. GEBRUIK VAN NIEUWE TECHNOLOGIEËN

De AVG stelt duidelijk dat een DPIA nodig kan zijn bij het toepassen van een nieuwe technologie. Het gebruik ervan kan namelijk gepaard gaan met nieuwe vormen om gegevens te verzamelen die mogelijks grote risico's hebben op gegevensbescherming (cf. COVID contact en gezondheidsapps).

De persoonlijke en sociale gevolgen van het gebruik van een nieuwe toepassing van AI-technologie kunnen immers nog onbekend zijn. Een DPIA kan daarbij de mogelijke risico's helpen inschatten en aanpakken en bijvoorbeeld vaststellen welke bijkomende informatie aan de betrokken personen gegeven moet worden, zodat ze de impact van de verwerking van hun gegevens kunnen inschatten. Naarmate een AI-systeem meer autonoom handelt, meer vrije beslissingsruimte heeft en zelf in staat zou zijn om gegevensbronnen te selecteren, wordt het belangrijker om de mogelijke gevolgen van deze autonomie grondig te analyseren.

8. UITSLUITING
Het gaat hier om gegevensverwerking die als gevolg hebben dat personen: • een recht niet kunnen uitoefenen; • een dienst niet kunnen gebruiken; of • geen contract kunnen afsluiten. Een voorbeeld hiervan is een werkgever die sollicitanten screent op basis van een databank met CV's en referenties om vervolgens te beslissen of ze al dan niet aangenomen worden.

In de meeste gevallen kan een organisatie ervan uitgaan dat een DPIA moet worden uitgevoerd voor een verwerking die aan twee van bovenstaande criteria voldoet. Over het algemeen gaat de WP29 ervan uit dat hoe groter het aantal criteria waaraan een verwerking voldoet, hoe groter de kans dat ze een hoog risico inhoudt voor de rechten en vrijheden van de betrokkenen. Bijgevolg is een DPIA dus vereist, ongeacht de maatregelen die de verwerkingsverantwoordelijke (nog) zou nemen.

4.4.C. Lijst van de GBA

Volgens de AVG moet elke nationale toezichthoudende autoriteit een lijst opstellen en publiceren van het **soort verwerkingen waarvoor een DPIA verplicht is**.[116]

De Belgische GBA publiceerde een **niet-exhaustieve lijst** met volgende verwerkingen waarvoor een DPIA nodig is:[117]

1.	Wanneer de verwerking gebruik maakt van **biometrische gegevens met het oog op de unieke identificatie van betrokkenen** die zich in een openbare ruimte bevinden of in privéruimten die toegankelijk zijn voor het publiek. AI-systemen die bijvoorbeeld gebruik maken van gezichtsherkenning kunnen hieronder vallen.
2.	Wanneer persoonsgegevens ingezameld worden bij derden om vervolgens in aanmerking te worden genomen bij de beslissing om een welbepaalde **dienstverleningsovereenkomst met een natuurlijke persoon te weigeren of stop te zetten**. Een voorbeeld zijn financiële instellingen die aan de hand van algoritmes naar informatie zoeken over de kredietwaardigheid van een klant.
3.	Wanneer **gezondheidsgegevens** van een betrokkene **geautomatiseerd** worden ingezameld aan de hand van een actieve **inplantbare medische voorziening**. Het gaat bijvoorbeeld om elke actieve ('slimme') medische voorziening die is ontworpen om geheel of gedeeltelijk in het menselijk lichaam te worden ingeplant.

[116] Artikel 35.4 AVG.
[117] Zie: Beslissing van het Algemeen secretariaat nr. 1/2019 van 16 januari 2019, B.S. 22 maart 2019, 28512-28514 (www.gegevensbeschermingsautoriteit.be/sites/privacycommission/files/documents/01_2019_AS.pdf).

4.	Wanneer er op **grote schaal gegevens ingezameld worden bij derden** teneinde de economische situatie, gezondheid, persoonlijke voorkeuren of interesses, betrouwbaarheid of gedrag, locatie of verplaatsingen van natuurlijke personen te analyseren of voorspellen. Denk bijvoorbeeld aan een webshop die met behulp van een AI-systeem zoveel mogelijk informatie over een persoon verzamelt doorheen verschillende kanalen, om zo gedetailleerd mogelijk voorspellingen te kunnen doen.
5.	Wanneer op systematische wijze **bijzondere categorieën van persoonsgegevens of gegevens van zeer persoonlijke aard** (zoals bijvoorbeeld gegevens over armoede, werkloosheid, betrokkenheid van jeugdzorg of maatschappelijk werk, gegevens omtrent huishoudelijke en privé-activiteiten of locatiegegevens) systematisch worden uitgewisseld tussen meerdere verwerkingsverantwoordelijken.
6.	Wanneer er sprake is van een **grootschalige verwerking van gegevens** die gegeneerd worden door middel van **toestellen met sensoren die via het internet of via een ander medium gegevens versturen** (Internet of Things of IoT toepassingen zoals slimme televisies, energiemeters en huishoudelijke apparaten of smart cities) en deze **verwerking dient** om de economische situatie, de gezondheid, de persoonlijke voorkeuren of interesses, betrouwbaarheid of gedrag, locatie of verplaatsingen van natuurlijke personen te analyseren of te voorspellen.
7.	Wanneer er sprake is van een **grootschalige en/of systematische verwerking** van telefonie-, internet- of andere communicatiegegevens, metagegevens of locatiegegevens van of herleidbaar tot natuurlijke personen (bijvoorbeeld wifi-tracking of verwerking van locatiegegevens van reizigers in het openbaar vervoer) wanneer de verwerking niet strikt noodzakelijk is voor een door de betrokkene gevraagde dienst.
8.	Wanneer er sprake is van **grootschalige verwerkingen van persoonsgegevens** waarbij op systematische wijze via geautomatiseerde verwerking gedrag van natuurlijke personen geobserveerd, verzameld, vastgelegd of beïnvloed wordt inclusief voor advertentiedoeleinden. Bijvoorbeeld het opstellen van profielen op basis van bestaande klantengegevens voor het aanbieden van gepersonaliseerde advertenties.

Een **DPIA kan ook moeten worden uitgevoerd voor bestaande verwerkingen**. Dit is met name het geval wanneer:

- het risico van een verwerking verandert, bijvoorbeeld omdat een nieuwe technologie gebruikt wordt, de aard van het eerder vastgestelde verwerkingsrisico een regelmatige her-evaluatie vereist of omdat persoonsgegevens voor een ander doel gebruikt worden;
- de organisatorische of maatschappelijke context verandert, bijvoorbeeld omdat bepaalde geautomatiseerde beslissingen belangrijker zijn geworden of omdat nieuwe categorieën betrokkenen kwetsbaar worden voor discriminatie. Neem bijvoorbeeld rekrutering waar het handmatig doorzoeken van profielen om de matching van vacatures te doen vervangen wordt door een AI-systeem dat dit taak geautomatiseerd doet zonder menselijke controle.

Toepassing
Wanneer een webshop persoonsgegevens van klanten verzamelt om de leveringen te kunnen uitvoeren is er sprake van een 'bestaande verwerking' die geen DPIA vereist. Een nieuwe verwerking van diezelfde gegeven waarvoor een DPIA vereist is, zou zich voordoen indien de organisatie dit klantenbestand voor nieuwe doeleinden gaat gebruiken, bijvoorbeeld voor profilering.[118]

Er is echter **geen DPIA vereist** indien:
- de verwerking waarschijnlijk geen hoog risico inhoudt;
- de verwerking al werd gecontroleerd door de GBA en de verwerking ondertussen niet is veranderd;
- de risico's van de verwerking niet zijn veranderd.

4.4.D. Tijdstip van het uitvoeren van de DPIA

Een DPIA moet **vóór de verwerking** van de persoonsgegevens worden uitgevoerd.[119] In de praktijk betekent dit dat de DPIA best in een zo vroeg mogelijk stadium wordt uitgevoerd zoals in de ontwerpfase van het product of proces waarbij gegevens verwerkt zullen worden, zelfs als bepaalde verwerkingen nog niet gekend zijn. Door de DPIA zo vroeg mogelijk uit te voeren, zal het ook eenvoudiger zijn om aan de vereisten van gegevensbescherming door ontwerp/standaardinstellingen te voldoen.[120]

Het werd reeds aangehaald dat het uitvoeren van een DPIA geen eenmalige uitoefening is, maar een **continu proces**. De verwerkingen moeten opgevolgd worden en het kan zijn dat de DPIA dan ook op een later moment bijgewerkt moet worden.

4.4.E. Wie moet een DPIA uitvoeren?

De **verwerkingsverantwoordelijke** moet de DPIA uitvoeren.[121] Deze kan eventueel binnen of buiten de organisatie uitgevoerd worden, maar de verwerkingsverantwoordelijke behoudt wel de eindverantwoordelijkheid.

Als er een **functionaris voor gegevensbescherming** of *Data Protection Officer* (DPO) is aangewezen, moet de verwerkingsverantwoordelijke ook diens advies inwinnen.[122] In het DPIA-rapport moeten het advies en de door de verwerkingsverantwoordelijke genomen acties worden opgenomen. De DPO ziet ook toe op de uitvoering van de DPIA.[123]

---

[118] Zie ook "5.4. Wat zegt de AVG over geautomatiseerde individuele besluitvorming en profilering en wat is de impact hiervan op AI-systemen?"
[119] Art. 35.1 AVG.
120 Zie ook "4.1. Wat betekent gegevensbescherming door ontwerp en hoe kan dit geïmplementeerd worden in AI-systemen?"
121 Art. 35.1 AVG.
122 Art. 35.2 AVG.
[123] Art. 39.1, c) AVG.

4.4.F. Stappenplan

Op grond van bovenstaande onderdelen, kan een **concrete stappenplan** worden opgesteld:

1.	Bekijk en bepaal zo vroeg mogelijk in de ontwikkeling van een AI-systeem of een DPIA moet worden uitgevoerd waarbij onder andere volgende vragen relevant zijn: • welke doelen worden nagestreefd met het AI-systeem? • wordt er gebruik gemaakt van een nieuwe technologie die mogelijkerwijs een grote impact op de gegevensbescherming kan hebben? • zal het AI-systeem tot bepaalde beslissingen of acties leiden tegen individuen op een manier die op hen een aanzienlijke invloed kunnen hebben? De inzet van een AI-systeem voor de verwerking van persoonsgegevens moet worden gestuurd door het feit dat het systeem een specifiek en legitiem doel kan bereiken en dus niet louter omdat de technologie beschikbaar is. Door de noodzaak van het gebruik van een AI-systeem in een DPIA te beschrijven, kan een organisatie aantonen dat dit doel niet op een andere redelijke manier kon worden bereikt.
2.	Beschrijf de beoogde verwerkingen systematisch, met telkens de mogelijke rechtsgrond en bijzondere aandacht voor: • het in kaart brengen van de gegevensstromen en de fases waarin een AI-toepassing en geautomatiseerde beslissingen een invloed zouden kunnen hebben op personen; • het gebruiken van zoveel mogelijk anonieme gegevens of het in kaart brengen van informatiestromen naarmate het project vordert indien de verwerkingsdoeleinden nog niet zeker zijn (bijvoorbeeld door correlaties die nog niet bekend zijn); • het bepalen en vastleggen van de rol en verplichtingen van verwerkingverantwoordelijken en verwerkers. Wanneer de verwerking door middel van AI-systemen geheel of gedeeltelijk wordt uitbesteed aan een derde partij, moeten alle betrokken organisaties beoordelen of er een gezamenlijke controlestructuur is opgezet. Indien dit het geval is, moeten zij aan de DPIA meewerken.[124]

3.	Breng de gegevensbeschermings- en daarmee samenhangende risico's in kaart met aandacht voor:
	• het beoordelen van de noodzaak en de proportionaliteit van de verwerking in verhouding tot het doel;
	• het beoordelen van de risico's voor de rechten en vrijheden van de betrokkenen;
	• het feit dat het opstellen van een DPIA niet een eenmalig feite/actie is, maar eerder een 'levend' en voortdurend proces dat zich ontwikkelt tijdens (de evolutie van) een project. Hier kunnen volgende vragen aan bod komen:
	  • zijn personen bewust gemaakt van het gebruik van hun persoonsgegevens?
	  • kan de gegevensset gevoelige gegevens bevatten?
	  • wat zijn de bewaringstermijnen van de verwerkte gegevens?
	  • worden de gegevens bewaard over meerdere systemen?
	  • hebben de systemen aangepaste beveiligingssystemen?
	  • kunnen geanonimiseerde gegevens ge-her-identificeerd worden?

[124] Zie ook "3.5. Wat zijn de verschillende rollen die een organisatie kan vervullen onder de GDPR in een AI-context?"

	• andere methodes dan de analyse van big data voor dit project. Het feit dat organisaties rekening moeten houden met andere wetgeving. Zo kan het gebruik van een AI-systeem bijvoorbeeld leiden tot discriminatie op basis van historische gegevenspatronen, wat in strijd kan zijn met de discriminatie-regelgeving.
4.	Identificeer en evalueer de mogelijke technische en organisatorische oplossingen of maatregelen om gegevensbeschermingsinbreuken uit te sluiten/beperken.
5.	Sluit de DPIA af.
6.	Integreer de uitkomsten van de DPIA in het projectplan van het AI-systeem. Het is belangrijk dat de personen die het AI-project gaan uitvoeren de oplossingen en maatregelen begrijpen, waarom ze nodig zijn en hoe ze geïmplementeerd kunnen worden. Dit is de laatste stap in het proces, maar zoals reeds gezegd niet het eindpunt. Regelmatige evaluaties moeten ervoor zorgen dat de voorgestelde oplossingen werken zoals het hoort. Bovendien kunnen de doelen en de toepassingen van het project gaandeweg veranderen. Regelmatige beoordelingen kunnen helpen om deze veranderingen te bepalen en na te gaan of de DPIA aangepast moet worden.

4.5. Welke verplichtingen gelden er wanneer persoonsgegevens worden verwerkt met het oog op wetenschappelijk onderzoek of statistische doeleinden?

Essentie
Bij de verwerking van persoonsgegevens met het oog op wetenschappelijke of statistische doeleinden dienen de AVG-verplichtingen te worden nageleefd. Zowel de AVG als de Belgische wetgeving[125] bepalen echter enkele uitzonderingen voor onderzoekers met betrekking tot bijvoorbeeld de rechten van betrokkenen, bewaartermijnen of verenigbaarheid van de verdere verwerking. In België is bovendien een 'waterval'-systeem van toepassing: in principe moet onderzoek gebeuren met geanonimiseerde gegevens, en enkel indien dat niet mogelijk is mogen er gepseudonimiseerde, respectievelijk, identificerende persoonsgegevens gebruikt worden.[126] De Belgische wet behandelt ook nog vijf onderwerpen: ◊ gegevensverzameling rechtstreeks bij de betrokkene: de gegevensbeschermingsverklaring dient extra informatie te bevatten;

[125] Wet van 30 juli 2018 betreffende de bescherming van natuurlijke personen met betrekking tot de verwerking van persoonsgegevens.
[126] Zie ook "3.3. Wat is het belang van het onderscheid tussen anonimisering en pseudonimisering?"

- ◊ onrechtstreekse gegevensverzameling / verdere verwerking van gegevens: er dient een overeenkomst te worden gesloten met of een kennisgeving te worden gedaan aan de oorspronkelijke verwerkingsverantwoordelijke;
- ◊ wanneer en hoe de gegevens anonimiseren of pseudonimiseren: afhankelijk van de situatie waarin men zich bevindt, dient de oorspronkelijke verwerkingsverantwoordelijke of een derde vertrouwenspersoon de doorgegeven persoonsgegevens op een bepaald moment te anonimiseren of te pseudonimiseren;
- ◊ verspreiding van gegevens: identificerende persoonsgegevens mogen niet worden verspreid tenzij er bepaalde uitzonderingen van toepassing zijn. Gepseudonimiseerde gegevens mogen wel verspreid worden, tenzij er bepaalde wettelijke beperkingen gelden of het gaat over gevoelige gegevens;
- ◊ mededeling van de gegevens: identificerende persoonsgegevens mogen voor wetenschappelijke of statistische doeleinden aan een geïdentificeerde derde worden meegedeeld/doorgegeven. In bepaalde gevallen mogen zij echter niet reproduceerbaar zijn, tenzij er uitzonderingen van toepassing zijn (zie verder).

Actiepunten

- ◊ Overweeg of de nodige technische en organisatorische maatregelen zijn getroffen om aan het beginsel van minimale gegevensverwerking te kunnen voldoen.
- ◊ Overweeg in welke mate de (mogelijke) uitoefening door betrokkenen van de rechten die de AVG hen toekent, de verwezenlijking van de specifieke wetenschappelijke of statistische onderzoeksdoeleinden onmogelijk dreigen te maken of ernstig dreigen te belemmeren, én in welke mate het beperkt of niet (hoeven) antwoorden op dergelijke verzoeken noodzakelijk is om de doeleinden te bereiken.
- ◊ Breng in kaart in welke gevallen persoonsgegevens verder worden verwerkt met het oog op een wetenschappelijk of statistische onderzoeksdoeleinde in de zin van de AVG (bijvoorbeeld training van AI-systeem voor commercialisering). Informeer in die gevallen de betrokkenen alvorens tot die verdere verwerking over te gaan (tenzij er sprake is van een van de uitzonderingsgevallen).
- ◊ Breng in kaart waar het handig zou zijn dat persoonsgegevens langer dan de strikt noodzakelijke periode worden bewaard, in zoverre zij louter (verder) worden verwerkt met het oog op wetenschappelijk onderzoek of statistische doeleinden. Voorzie in dat geval passende technische en organisatorische maatregelen zodat onder andere de toegang tot en gebruik van die gegevens wordt beperkt.
- ◊ Voor een DPIA uit indien vereist door de AVG.
- ◊ Voeg de bijkomende informatie die specifiek door de wet wordt vereist toe aan het register van verwerkingsactiviteiten en de relevante gegevensbeschermingsverklaring.
- ◊ Sluit een overeenkomst die aan de wettelijke vereisten voldoet indien er persoonsgegevens worden verwerkt die niet rechtstreeks bij de betrokkene werden verzameld. Indien er geen overeenkomst moet worden gesloten, richt dan een kennisgeving conform de wettelijke vereisten aan de oorspronkelijke verwerkingsverantwoordelijke.
- ◊ Onderzoek onder welke van de wettelijke beschreven situaties de verwerking zich bevindt, om te weten te komen wie wanneer persoonsgegevens moet anonimiseren of pseudonimiseren.
- ◊ Maak geen niet-gepseudonimiseerde gegevens publiek tenzij het zeker is dat één van de uitzonderingen van toepassing is.
- ◊ Zorg ervoor dat niet-gepseudonimiseerde gegevens op een niet-reproduceerbare wijze worden meegedeeld aan een geïdentificeerde derde indien er sprake is van een situatie zoals beschreven in de wet.

4.5.A. Wetenschappelijke doeleinden of statistische doeleinden

In dit deel wordt besproken of en wanneer AI-ontwikkelaars en -onderzoekers beroep kunnen doen op het (nationale) **uitzonderingregime** onder de AVG voor door hun verrichte verwerkingen in het kader van (wetenschappelijk) onderzoek of met het oog op statistische doeleinden. Het is dus in de eerste plaats van belang om te weten wat de AVG verstaat onder **wetenschappelijk onderzoek en statistische doeleinden**.

WETENSCHAPPELIJK ONDERZOEK

Wetenschappelijk onderzoek wordt ruim opgevat en omvat bijvoorbeeld technologische ontwikkeling en demonstratie, fundamenteel onderzoek, toegepast onderzoek en uit private middelen gefinancierd onderzoek. Het omvat ook studies op het gebied van de volksgezondheid die in het algemeen belang worden gedaan.

Er kan worden vanuit gegaan dat AI-ontwikkelaars, -onderzoekers en zelfs -gebruikers zich in bepaalde omstandigheden bezighouden met wat de AVG als 'wetenschappelijk onderzoek' verstaat. Het trainen van een AI-systeem tijdens de ontwikkelingsfase kan bijvoorbeeld vallen onder de technologische ontwikkeling van het systeem. Ook fundamenteel AI-onderzoek, ongeacht of privaat of publiek wordt gefinancierd, kan onder dit uitzonderingsregime vallen. Het trainen van een AI-systeem tijdens de gebruiksfase, zal dan weer niet als technologische ontwikkeling worden beschouwd.

VERWERKING MET OOG OP STATISTISCHE DOELEINDEN

Onder statistische doeleinden wordt verstaan het verzamelen en verwerken van persoonsgegevens die nodig zijn voor statitische onderzoeken en voor het produceren van statistische resultaten.[127] Die statistische resultaten mogen ook voor andere doeleinden worden gebruikt, onder meer voor wetenschappelijke onderzoeksdoeleinden.

Het statistisch oogmerk betekent dat het resultaat van de verwerking voor statistische doeleinden niet uit persoonsgegevens, maar uit geaggregeerde gegevens bestaat. Dit resultaat en de gerelateerde persoonsgegevens mogen daarbij geen aanleiding geven tot maatregelen of beslissingen die een specifieke natuurlijke persoon betreffen (omgekeerd, wel een groep personen).

Ook hier kan er vanuit worden gegaan dat AI-ontwikkelaars, -onderzoekers en zelfs -gebruikers in bepaalde omstandigheden verwerkingen zullen uitvoeren die vallen onder wat de AVG onder 'statistische doeleinden' verstaat. Zo kunnen bijvoorbeeld gebruik(er)statistieken of accuraatheids-analyses zowel tijdens de training als tijdens de implementatie van een AI-systeem (in de mate dat er persoonsgegevens voor gebruikt worden) gebeuren met het oog op het verkrijgen van statistische resultaten. Daarenboven sluit de AVG niet uit dat dergelijke resultaten later voor commerciële doeleinden kunnen worden gebruikt.

[127] Overweging 162 AVG.

4.5.B. Toepasselijke AVG-principes

In de AVG staan een aantal **belangrijke elementen** met betrekking tot de (initiële en verdere) verwerking van persoonsgegevens met oog op wetenschappelijk onderzoek of statistische doeleinden. De onderstaande tabel vat ze kort samen.[128]

1.	Dergelijke verwerking dient de rechten en plichten die een betrokkene onder de AVG geniet (in principe) te respecteren. Meer specifiek, dient het beginsel van minimale gegevensverwerking door middel van technische en organisatorische maatregelen (bijvoorbeeld pseudonimisering) te worden gegarandeerd, in de mate dat dergelijke doeleinden nog steeds kunnen worden verwezenlijkt.[129]
2.	Wanneer persoonsgegevens worden verwerkt voor deze doeleinden, kan het nationale recht voorzien in afwijkingen van het recht op inzage, het recht op rectificatie, het recht op beperking van verwerking, het recht van bezwaar en het recht op wissing, voor zover die rechten de verwezenlijking van de specifieke doeleinden onmogelijk dreigen te maken of ernstig dreigen te belemmeren, én dergelijke afwijkingen noodzakelijk zijn om die doeleinden te bereiken. Deze noodzakelijkheidsvereiste geldt echter niet bij wissingsverzoeken.[130]
3.	Wanneer verwerking met oog op wetenschappelijk onderzoek of statistische doeleinden tegelijkertijd ook een ander doel dient, zijn de mogelijke afwijkingen enkel van toepassing op de verwerking voor wetenschappelijke of statistische doeleinden.
4.	De AVG bepaalt ook nog enkele bijkomende afwijkingen of uitzonderingen: • De verdere verwerking met het oog op wetenschappelijk onderzoek of statistische doeleinden dient in principe als verenigbaar met de oorspronkelijke doeleinden (van de initiële verwerking) te worden beschouwd.[131] De verwerkingsverantwoordelijke moet in dat geval in principe de betrokkene daar wel van informeren, alvorens over te gaan tot die verwerking. De verwerkingsverantwoordelijke kan echter vrijgesteld worden van deze informatieplicht (in het kader van hun onderzoeksactiviteiten) in het geval hij gegevens verwerkt die hij zelf niet rechtstreeks bij de betrokkenen heeft verzameld. Dat is echter enkel het geval indien het verstrekken van die informatie onmogelijk blijkt of onevenredig veel inspanning zou vergen.[132] Er dienen dan wel passende maatregelen te worden genomen, zoals de relevante informatie publiek mee te delen (bijvoorbeeld op een website, als de betrokken personen daarmee voldoende worden bereikt); • Ook mogen persoonsgegevens voor langere dan strikt noodzakelijke perioden worden opgeslagen indien zij louter met het oog op wetenschappelijk onderzoek of statistische doeleinden worden verwerkt, mits er passende technische en organisatorische maatregelen worden getroffen om de rechten en vrijheden van de betrokkene te beschermen (bijvoorbeeld pseudonimisering of beperkte toegang tot de gegevens).

Voor meer informatie over het onderscheid tussen **gepseudonimiseerde** en **anonieme persoonsgegevens**,

[128] Art. 89 AVG. Dat artikel behandelt ook nog archivering in het algemeen belang en historisch onderzoek. Deze categorieën worden hier niet verder behandeld.
[129] Zie ook "3.3. Wat is het belang van het onderscheid tussen anonimisering en pseudonimisering?"
[130] Art. 17.3, d) AVG.
[131] Art. 5.2, b) AVG. Dit werd ook bevestigd door de Spaanse Gegevensbeschermingsautoriteit.
[132] Art. 14.5, b) AVG.

wordt verwezen naar deel 3.3. Ook wordt nog benadrukt dat een verwerking die gebeurt voor wetenschappelijke doeleinden of met het oog op statistische doeleinden aanleiding kan geven tot het moeten invullen van **DPIA**. Dit werd nader toegelicht in deel 4.4.

4.5.C. Toepasselijke Belgische wetgeving

Deze bepalingen uit de AVG worden **verder aangevuld** in artikels 186 en verder van de Belgische wet betreffende de bescherming van natuurlijke personen met betrekking tot de verwerking van persoonsgegevens. Het gaat daarbij over zowel **algemene als specifieke aanvullingen**.

ALGEMENE AANVULLINGEN
De wet herhaalt dat er uitzonderingen mogelijk zijn op bepaalde rechten van betrokkenen indien zij de verwezenlijking van de verwerkingen onmogelijk dreigen te maken of ernstig dreigen te belemmeren, en afwijkingen noodzakelijk zijn om die doeleinden te bereiken. Vervolgens benadrukt de wet dat een verwerkingsverantwoordelijke die persoonsgegevens verwerkt voor statistische of wetenschappelijke doeleinden nog steeds een DPO dient aan te wijzen indien de verwerking van de persoonsgegevens een hoog risico kan inhouden zoals bedoeld in artikel 35 AVG. Tot slot legt de wet ook nog op dat de verantwoordelijke voor de verwerking met het oog op wetenschappelijke of statistische doeleinden, voorafgaandelijk aan de gegevensverzameling, de volgende elementen toevoegt aan diens register van verwerkingsactiviteiten:[133] • de verantwoording van het gebruik van de al dan niet gepseudonimiseerde gegevens; • de redenen volgens dewelke de uitoefening van de rechten van de betrokkene de verwezenlijking van de doeleinden onmogelijk dreigen te maken of ernstig dreigen te belemmeren; • desgevallend, de DPIA indien de verwerkingsverantwoordelijke met het oog op wetenschappelijke of statistische doeleinden gevoelige gegevens verwerkt.
SPECIFIEKE AANVULLINGEN
Van belang is ook dat de Belgische wet drie nieuwe begrippen invoert: • derde vertrouwenspersoon: de natuurlijke persoon of rechtspersoon, de feitelijke vereniging of de overheidsadministratie die de gegevens pseudonimiseert en die niet ook de verwerkingsverantwoordelijke is voor de verwerking met het oog op archivering of onderzoek of statistische doeleinden; • mededeling van gegevens: mededeling van gegevens aan een geïdentificeerde derde; • verspreiding van gegevens: bekendmaking van de gegevens, zonder identificatie van de derde.

Daarenboven stelt de wet volgende '**waterval-principe**' voor met betrekking tot het al dan niet anonimiseren van persoonsgegevens.

[133] In principe dienen alle verwerkingsverantwoordelijken en verwerkers een register van de verwerkingsactiviteiten die onder hun verantwoordelijkheid plaatsvinden, bij te houden. Artikel 30 AVG bepaalt welke gegevens dit register dient te bevatten.

PRINCIPE	De verantwoordelijke voor de verwerking met het oog op wetenschappelijk onderzoek of statistische doeleinden gebruikt anonieme gegevens.
UITZONDERING	Indien het niet mogelijk is om met anonieme gegevens het onderzoeksdoel of statistische doel te bereiken, gebruikt de verwerkingsverantwoordelijke gepseudonimiseerde gegevens.
UITZONDERING	Indien het niet mogelijk is om met gepseudonimiseerde gegevens het onderzoeksdoel of het statistische doel te bereiken, gebruikt de verwerkingsverantwoordelijke niet-gepseudonimiseerde gegeven

De Belgische wet behandelt daarbij **vijf onderwerpen** specifiek met betrekking tot verwerking voor wetenschappelijke of statistische doeleinden.

1. GEGEVENSVERZAMELING RECHTSTREEKS BIJ DE BETROKKENE

Bovenop de informatie die een verwerkingsverantwoordelijke dient mee te delen conform artikel 13 AVG[134] legt de wet op dat de betrokkene ook wordt geïnformeerd over:

- het feit dat de gegevens al dan niet worden geanonimiseerd;
- de redenen volgens dewelke de uitoefening van de rechten van de betrokkene de verwezenlijking van de doeleinden onmogelijk dreigen te maken of ernstig dreigen te belemmeren.

Deze informatie moet worden meegedeeld alvorens de persoonsgegevens bij de betrokkene worden verzameld.

Indien in de toepassingen van e-commerce en rekrutering dus rechtstreeks persoonsgegevens worden verzameld bij betrokkenen welke op een later moment ook worden gebruikt voor wetenschappelijke of statistische onderzoeksdoeleinden, dient de gegevensbeschermingsverklaring van de verwerkingsverantwoordelijke deze informatie te vermelden.

2. ONRECHTSTREEKSE GEGEVENSVERZAMELING / VERDERE VERWERKING VAN GEGEVENS

Indien persoonsgegevens niet rechtstreeks bij de betrokkene worden verzameld, moet de verwerkingsverantwoordelijke een overeenkomst sluiten met de verantwoordelijke voor de oorspronkelijke verwerking, dus de bron van de persoonsgegevens.

Een dergelijke overeenkomst dient echter niet te worden gesloten indien:

- de verwerking betrekking heeft op persoonsgegevens die publiek zijn gemaakt (bijvoorbeeld via sociale media); of
- wanneer het recht van de Europese Unie, een wet, een decreet of een ordonnantie:
  - de verwerkingsverantwoordelijke een mandaat geeft om persoonsgegevens te verwerken met het oog op wetenschappelijke of statistische doeleinden; en
  - het hergebruik van de verzamelde gegevens voor andere doeleinden verbiedt.

[134] Zie ook "5.1. Welke transparantieverplichtingen legt de AVG op en wat zijn de specifieke aandachtspunten in een AI-context?"

In geval van vrijstelling voor het afsluiten van een overeenkomst moet de verwerkingsverantwoordelijke (de onderzoeker) de beoogde gegevensverzameling en verwerking melden aan de verantwoordelijke van de oorspronkelijke verwerking (bijvoorbeeld het sociale media platform).

De bovenvermelde overeenkomst of kennisgeving dienen minstens de volgende elementen te bevatten:

- in geval van een overeenkomst, de contactgegevens van de verantwoordelijke voor de oorspronkelijke verwerking en van de verantwoordelijke voor de verdere verwerking;
- de redenen volgens dewelke de uitoefening van de rechten van de betrokkene de verwezenlijking van de doeleinden onmogelijk dreigen te maken of ernstig dreigen te belemmeren.

Deze overeenkomst of kennisgeving dient ook bij het register van de verwerkingsactiviteiten gevoegd te worden door de verwerkingsverantwoordelijke.

Het is mogelijk dat bedrijven die werkzaam zijn binnen rekrutering of e-commerce beroep zullen doen op externe bronnen (bijvoorbeeld *data brokers*) om gegevens te vergaren om hun AI-systemen mee te trainen of accuraatheidsanalyses te doen. Daarom dienen zij er zich van bewust te zijn dat zij een bijkomende overeenkomst zullen moeten sluiten om in overeenstemming met de Belgische wetgeving te zijn.

3. WANNEER DIENT MEN DE GEGEVENS VERWERKT MET HET OOG OP WETENSCHAPPELIJKE OF STATISTISCHE DOELEINDEN TE ANONIMISEREN OF PSEUDONIMISEREN EN DOOR WIE MOET DAT GEBEUREN?[135]

Situatie 1: conform bovenstaande 'waterval-principe' dient de verwerkingsverantwoordelijke, indien het gaat over een rechtstreekse gegevensverzameling bij de betrokkene, over te gaan tot de anonimisering of pseudonimisering van de gegevens na de verzameling ervan.

Situatie 2: indien de verwerkingsverantwoordelijk reeds persoonsgegevens in zijn bezit heeft (in verband met een eerdere verwerking) en die zelf wil verwerken met het oog op wetenschappelijke of statistische doeleinden anonimiseert of pseudonimiseert de verwerkingsverantwoordelijke de gegevens *voorafgaandelijk aan hun verdere verwerking*.

Dergelijke verwerkingsverantwoordelijke mag deze persoonsgegevens slechts de-pseudonimiseren indien dat noodzakelijk is voor het onderzoek of de statistische doeleinden, en desgevallend na advies van de DPO, wat gedocumenteerd moet worden.

Situatie 3: indien een verwerkingsverantwoordelijke de persoonsgegevens doorgeeft aan een andere verwerkingsverantwoordelijke, pseudonimiseert of anonimiseert de oorspronkelijke verwerkingsverantwoordelijke de gegevens voorafgaandelijk aan de mededeling ervan aan de verantwoordelijke voor de verdere verwerking.

De verantwoordelijke voor de verdere verwerking mag geen toegang tot de sleutels van de pseudonimisering hebben.

[135] Indien de betrokken verwerkingsverantwoordelijken een DPO hebben aangewezen, dient deze te adviseren over het gebruik van verschillende pseudonimiserings- en anonimiseringsmethoden in het bijzonder over de doeltreffendheid van de bescherming van gegevens.

Situatie 4: indien er meerdere oorspronkelijke verwerkingen worden gekoppeld, laten de oorspronkelijke verwerkingsverantwoordelijken voorafgaandelijk aan de mededeling van de gegevens aan de verantwoordelijke voor de verdere verwerking, de gegevens anonimiseren of pseudonimiseren door één van de verantwoordelijken voor de oorspronkelijke verwerking of door een derde vertrouwenspersoon.[136]

Indien een van de oorspronkelijke verwerkingsverantwoordelijke gevoelige gegevens doorgeeft in een dergelijke situatie, mag enkel deze verwerkingsverantwoordelijke, voorafgaandelijk aan de mededeling van de gegevens aan de verdere verwerkingsverantwoordelijke, de gegevens anonimiseren of pseudonimiseren (of een derde vertrouwenspersoon).

Enkel de verantwoordelijke voor de oorspronkelijke verwerking die de gegevens heeft gepseudonimiseerd of de derde vertrouwenspersoon heeft toegang tot de pseudonimiseringssleutels.

4. (PUBLIEKE) VERSPREIDING VAN GEGEVENS

Tenzij bepaalde wetgeving strengere voorwaarden oplegt voor de verspreiding van de gegevens[137] die verwerkt zijn met het oog op het wetenschappelijke of statistische doeleinden, verspreidt de verwerkingsverantwoordelijke geen niet-gepseudonimiseerde gegevens, tenzij:

- de betrokkene zijn toestemming heeft verleend; of
- de gegevens door de betrokkene zelf openbaar zijn gemaakt; of
- de gegevens nauw samenhangen met het openbare of historische karakter van de betrokkene; of
- de gegevens nauw samenhangen met het openbare of historische karakter van feiten waarbij de betrokkene betrokken was.

Gepseudonimiseerde gegevens mogen wel door de verwerkingsverantwoordelijke worden verspreid tenzij bepaalde wetgeving hieraan in de weg zou staan of indien het gaat over gevoelige gegevens. Geanonimiseerde gegevens mogen wel worden verspreid.

5. MEDEDELING VAN DE GEGEVENS

Tenzij wetgeving strengere voorwaarden zou opleggen, zorgt de verwerkingsverantwoordelijke die niet-gepseudonimiseerde gegevens aan een geïdentificeerde derde meedeelt/doorgeeft voor wetenschappelijke of statistische doeleinden ervoor dat de geïdentificeerde derde de meegedeelde gegevens niet kan reproduceren, behalve op handgeschreven wijze, indien:

- het om gevoelige persoonsgegevens gaat;[138] of
- de overeenkomst tussen de verantwoordelijken voor de oorspronkelijke verwerking en de verdere verwerking zulks verbiedt; of
- die reproductie de veiligheid van de betrokkene in het gedrang kan brengen.

[136] De derde vertrouwenspersoon dient (1) onderworpen te zijn aan een beroepsgeheim in de zin van artikel 458 Strafwetboek; en (2) mag afhankelijk zijn van noch de oorspronkelijke noch de verdere verwerkingsverantwoordelijke.
[137] Zie hierboven voor de definitie van verspreiding van gegevens.
[138] Zie ook "4.2. Wat houdt de vereiste van minimale gegevensverwerking van persoonsgegevens in voor AI-systemen?"

Deze verplichting is niet van toepassing indien:

- de betrokkene zijn toestemming heeft verleend; of
- de gegevens door de betrokkene zelf openbaar zijn gemaakt; of
- de gegevens nauw samenhangen met het openbare of historische karakter van de betrokkene; of
- de gegevens nauw samenhangen met het openbare of historische karakter van feiten waarbij de betrokkene betrokken was.

Als een in rekrutering betrokken bedrijf verkregen persoonsgegevens (bijvoorbeeld van op een CV) zou wensen mee te delen aan een derde partij met het oog op wetenschappelijke of statistische doeleinden, dient het zich er van te verzekeren dat indien er ook gevoelige gegevens worden meegedeeld, deze niet reproduceerbaar zijn. Omgekeerd houdt dit in dat als een bedrijf betrokken bij e-commerce niet-gepseudonimiseerde gegevens ontvangt met het oog op wetenschappelijke doeleinden, het zich ervan dient te verzekeren dat het eventuele gevoelige gegevens niet kan reproduceren.

Hoofdstuk 5: Gegevensbescherming tijdens de gebruiksfase van AI-systemen

5. HOE KAN GEGEVENSBESCHERMING WORDEN VERZEKERD TIJDENS DE GEBRUIKSFASE (*DEPLOYMENT*) VAN AI-SYSTEMEN?

In de volgende delen worden een aantal zaken besproken uit de AVG die van belang zijn bij het gebruik van AI-systemen: transparantieverplichtingen (*transparency*), opslagbeperkingen (*storage limitation*), de rechten van de betrokkenen en geautomatiseerde individuele besluitvorming waaronder profilering (*automated individual decision-making* en *profiling*).

5.1. Welke transparantieverplichtingen legt de AVG op en wat zijn de specifieke aandachtspunten in een AI-context?

Essentie
Verwerkingsverantwoordelijken moeten volgens de AVG bepaalde informatie verschaffen aan betrokkenen. Er zijn zowel algemene transparantieverplichtingen die op alle soorten verwerkingen van toepassing zijn, als specifieke transparantieverplichtingen die in acht moeten worden genomen bij bepaalde verwerkingen die gebruik maken van AI-systemen. Dit geldt des te meer indien er sprake is van geautomatiseerde individuele besluitvorming.[139]
Actiepunten
◊ Zorg ervoor dat de organisatie een gegevensbeschermingsbeleid heeft dat alle door de AVG opgelegde informatie bevat en op het juiste moment wordt meegedeeld aan de betrokkenen; ◊ Overweeg om te werken met een gelaagd gegevensbeschermingsverklaring, zeker indien er nuttige informatie met betrekking tot de onderliggende logica van het AI-systeem moet worden gegeven; ◊ Overweeg om visuele en interactieve technieken te gebruiken om deze informatie op een heldere en verstaanbare manier aan betrokkenen mee te delen; ◊ Breng in kaart bij welke verwerkingen door middel van AI-systemen er sprake is van geautomatiseerde besluitvorming en of deze verwerkingen juridische of andere belangrijke gevolgen met zich meebrengen ten aanzien van de betrokkenen; ◊ Tracht bij de ontwikkeling van AI-systemen een 'verklaarbaarheid door ontwerp' (*explainability by design*) aanpak te hanteren en te streven naar een zo transparant mogelijk ontwerp van AI-systemen; ◊ Informeer betrokkenen zodra ze in contact treden met een AI-systeem waarbij er sprake is van geautomatiseerde besluitvorming; ◊ Denk na over welke informatie de organisatie wilt meedelen indien er nuttige informatie omtrent de onderliggende logica van een AI-systeem moet worden verschaft, en over hoe deze informatie zal verschaft worden;

[139] Zie ook "5.4. Wat zegt de AVG over geautomatiseerde individuele besluitvorming en profilering en wat is de impact hiervan op AI-systemen?"

◊ Informeer de betrokkene over de beoogde of te verwachten verwerking door het AI-systeem en de gevolgen die de geautomatiseerde besluitvorming voor de betrokkene kan teweegbrengen door middel van tastbare voorbeelden.

Personen hun privéleven wordt ingevolge het gebruik van AI-systemen (en ruimer big data analyses) steeds transparanter voor de organisaties die dergelijke systemen en analyses gebruiken. Omgekeerd worden deze systemen zelf vaak net gekenmerkt door een gebrek aan transparantie, of zelfs ondoorzichtigheid ('*opacity*') naar de betrokkenen en toezichthouders toe. Niettemin voorziet de AVG transparantieverplichtingen die ook in een AI-context dienen te worden nageleefd.

5.1.A. Externe en interne transparantie

Met het oog op de discussie van de transparantieverplichtingen onder de AVG, wordt eerst kort stilgestaan bij het onderscheid tussen **externe en interne transparantie**. Dit is een onderscheid dat op zich niet uitdrukkelijk voortvloeit uit de AVG, maar in de praktijk wel één en ander kan verhelderen.

EXTERNE TRANSPARANTIE
Dit begrip verwijst naar de transparantie die een verwerkingsverantwoordelijke moet verschaffen aan de buitenwereld met betrekking tot de verwerkingen van persoonsgegevens die hij doet. Het komt dus neer op het vertalen van wat er op technisch vlak gebeurt in een AI-systeem naar voor de betrokkene begrijpelijke bewoordingen en redeneringen. Dergelijke transparantie is niet alleen van belang voor de betrokkenen zelf, maar ook voor andere belanghebbenden zoals de media, belangenorganisaties in het maatschappelijk veld en toezichthouders. In de artikels 12-14 bepaalt de AVG welke informatie gegeven moet worden. Het gaat dus niet zozeer om het gedetailleerd uitleggen van wat een AI-systeem kan en doet, maar eerder en voornamelijk over het beheren van de verwachtingen van derde partijen.
INTERNE TRANSPARANTIE[140]
Dit verwijst naar de transparantie met betrekking tot de werking van een AI-systeem die binnen een organisatie best verzekerd wordt. Niet alleen IT-teams moeten kunnen begrijpen tot wat de gebruikte AI-systemen in staat zijn en hoe zij werken, maar ook product- en account managers, een functionaris voor gegevensbescherming, bedrijfsjuristen of kaderleden dienen over relevante informatie te kunnen beschikken, zodat zij deze systemen op een gepaste en weloverwogen manier gebruiken.

[140] Dit luik, en het gerelateerde explainability-vraagstuk, wordt wegens het technischer karakter en het feit dat dit niet expliciet door de AVG wordt opgelegd niet in detail besproken. Voor meer informatie, wordt verwezen naar het 'Explain AI'-project van de Britse ICO en het Alan Turing Institute, of naar het rapport 'Robustness and Explainability of Artificial Intelligence' van de Europese Commissie.

Onder de AVG wordt dan eerder gesproken over '*accountability*' (verantwoordingsplicht). In deze context is het wel aanbevolen om gedetailleerd te documenteren over hoe een AI-systeem bijvoorbeeld gegevens verwerkt, welk technische principes er aan grondslag liggen, (een beschrijving van) welke gegevens verwerkt worden, wie welke rol op welk moment in het ontwikkelingsproces speelde en welke trainingsmethodologieën werden toegepast. Deze informatie kan worden opgenomen in interne richtlijnen, reglementen of memo's die aan de verschillende functies binnen een organisatie zijn gericht en informatie verschaffen omtrent de werking van de gebruikte AI-systemen.

De volgende delen gaan dieper in op het eerste luik en onderzoeken welke informatie er volgens de AVG in een gegevensbeschermingsbeleid moet worden opgenomen.

5.1.B. Transparantieverplichtingen onder de AVG

Een eerste deel bespreekt de **algemene informatieverplichtingen** die onder de AVG moeten worden nageleefd. Vervolgens wordt uitvoeriger stilgestaan bij de **informatieverplichtingen die specifiek met betrekking tot AI-systemen** verduidelijking kunnen gebruiken. Belangrijk om te beseffen in een AI-context is dat de transparantieverplichtingen betrekking hebben op **al de fasen van de gegevensverwerking**, dus op zowel de trainings- en testfase als de fase van de toepassing van de AI-systemen.

In een AI-context moet transparantie de betrokkenen in staat stellen om te begrijpen wat de gevolgen zijn van dergelijke AI-systemen. Transparantie is zowel gericht naar de betrokkenen als naar de verwerkingsverantwoordelijken. Meer in het bijzonder houdt transparantie verband met accurate informatie over de **reële mogelijkheden en de reële beperkingen van AI-systemen** zodat **valse verwachtingen** bij de betrokkenen en verkeerde interpretaties van de resultaten worden vermeden. Transparantie omvat ook het verschaffen van informatie over de context van de verwerking, de betrokkenheid van derden,...

De hieronder opgenomen informatieverplichtingen lijsten de algemene transparantieverplichtingen niet exhaustief op. Er zijn reeds meerdere documenten verschenen die dit meer in detail uitleggen.[141]

Algemene informatieverplichtingen

De artikels 12, 13 en 14 AVG bevatten de voornaamste algemene transparantieverplichtingen die verwerkingsverantwoordelijken dienen na te leven. Ze worden hieronder kort besproken.

[141] Voor meer informatie wordt verwezen naar bijvoorbeeld de vrij gedetailleerde en toonaangevende richtlijnen over transparantie van de voormalige artikel 29 Werkgroep (WP29).

ARTIKEL 12 AVG

Artikel 12 AVG bepaalt op algemene wijze dat de betrokkene de bedoelde informatie in een beknopte, transparante, begrijpelijke en gemakkelijk toegankelijke vorm en in duidelijke en eenvoudige taal moet ontvangen.

Hiermee wordt bedoeld dat informatie op dusdanige wijze moet worden gepresenteerd dat een doorsnee opgeleid persoon de informatie vlot kan terugvinden en begrijpen, waarbij deze duidelijk afgescheiden dient te zijn van informatie die niet gericht is op gegevensbescherming (bijvoorbeeld algemene voorwaarden). Online kan er daarenboven worden overwogen om met een zogenaamde 'gelaagde' gegevensbeschermingsverklaring te werken. Dat is een gegevensbeschermingsverklaring waarbij op de eerste laag er eerder algemene informatie te vinden is over hoe de persoonsgegevens verwerkt worden en die betrokkenen in staat stelt om verder te klikken naar verdere lagen waar zij meer gedetailleerde informatie kunnen consulteren, zonder door een lang tekstdocument te moeten scrollen.

ARTIKEL 13 EN 14 AVG

Artikels 13 en 14 bepalen meer in detail wanneer welke informatie aan een betrokkene dient te worden meegedeeld. Doorgaans wordt deze informatie opgenomen in een gegevensbeschermingsverklaring die online wordt gepubliceerd of als afgedrukt exemplaar wordt verspreid.

Bij de rechtstreekse gegevensverzameling bij de betrokkene zelf (art. 13 AVG) moet men onder andere volgende informatie meedelen:

- de contactgegevens van een DPO (indien aangesteld);
- de doeleinden waarvoor persoonsgegevens worden verwerkt en bijhorende rechtsgrond die men inroept, per verwerking; de ontvangers of categorieën van ontvangers van persoonsgegevens[142];
- de periode gedurende welke de persoonsgegevens zullen worden opgeslagen, of indien dat niet mogelijk is, de criteria ter bepaling van die termijn. Een betrokkene zou in staat moeten zijn om, rekening houdende met de individuele situatie, een idee te hebben over hoelang bepaalde persoonsgegevens zullen worden bewaard;
- welke rechten de betrokkene heeft, waaronder ook de rechten die een betrokkene heeft indien er sprake is van geautomatiseerde besluitvorming (namelijk recht op menselijke tussenkomst, recht om een standpunt kenbaar te maken en recht om het besluit aan te vechten) en hoe die uitgeoefend kunnen worden;
- het recht om klacht in te dienen bij een gegevensbeschermingsautoriteit.

[142] Ontvangers zijn alle partijen die persoonsgegevens verkrijgen via de verwerkingsverantwoordelijken, bijvoorbeeld andere verwerkingsverantwoordelijken of verwerkers. In principe dient men deze individueel te benoemen, maar men kan er ook mee volstaan om deze per categorie op te lijsten waarbij men bijvoorbeeld de sector en locatie van de ontvangers duidelijk maakt.

Onder artikel 14 (onrechtstreekse gegevensverzameling bijvoorbeeld via een derde partij[143]) moet dezelfde informatie worden meegedeeld, maar ook nog:[144]

- welke categorieën gegevens verwerkt worden (de betrokkene is zich in dergelijk geval immers niet bewust van welke gegevens er over hem/haar zijn verzameld);
- de bron waar men de gegevens heeft verkregen en, indien toepasselijk, of zij van publiek beschikbare bronnen komen. Hier dient men in principe de specifieke bron van de gegevens (enquêtes, online of mobiele toepassingen, open data, sociale media, koppelingen tussen verschillende databronnen,…) op te noemen tenzij dat onmogelijk is.[145] In dat geval moet men de aard van de bron (bijvoorbeeld publieke of private bron) en het type of de sector van de bron benoemen.

Met betrekking tot het **moment waarop deze informatie moet worden meegedeeld**, moet met volgende zaken rekening worden gehouden:

- indien de gegevens rechtstreeks worden verzameld bij de betrokkene dient de informatie verschaft te worden bij de verzameling van de gegevens (bijvoorbeeld als een betrokkene online een webshopformulier invult, zich inschrijft op een nieuwsbrief of een CV uploadt);
- indien de gegevens onrechtstreeks worden verzameld, hangt het af van de situatie:
  - het principe is dat de informatie binnen een redelijke termijn moet worden meegedeeld, afhankelijk van de concrete omstandigheden waarin de persoonsgegevens worden verwerkt, maar uiterlijk binnen één maand na de verkrijging van de persoonsgegevens;
  - indien de persoonsgegevens worden gebruikt voor communicatie met de betrokkene (bijvoorbeeld om iemand uit te nodigen om voor een bepaalde functie te solliciteren of bepaalde goederen aan te kopen), moet de informatie verstrekt worden uiterlijk op het moment van het eerste contact met de betrokkene; of
  - indien verstrekking van de gegevens aan een andere ontvanger wordt overwogen, uiterlijk op het tijdstip waarop de persoonsgegevens voor het eerst worden verstrekt.

Er zijn echter **enkele uitzonderingen op deze informatieplicht bij onrechtstreekse gegevensverzameling** zoals vooropgesteld in artikel 14 AVG. Zo moet deze informatie niet worden verstrekt in gevallen waarbij persoonsgegevens worden verwerkt met het oog op wetenschappelijke of statistische doeleinden en (i) het verstrekken van die informatie onmogelijk blijkt of onevenredig veel inspanning zou vergen, of (ii) het verstrekken van die informatie de verwezenlijking van de verwerkingsdoeleinden onmogelijk dreigt te maken of ernstig in het gedrang dreigt te brengen.[146]

[143] Dit kunnen bijvoorbeeld andere verwerkingsverantwoordelijken zijn zoals een vorige werkgever van de betrokkene, data brokers, andere betrokkenen of publieke bronnen waaronder het Staatsblad, KBO,….
[144] Het is echter een goede en transparante praktijk om deze informatie ook mee te delen indien gegevens rechtstreeks bij de betrokkene worden verzameld.
[145] Het feit dat in een database verschillende gegevens over betrokkenen van verschillende bronnen komen, maakt het niet onmogelijk dat de bron wordt benoemd, ongeacht de vereiste tijdsinvestering of werklast voor de verwerkingsverantwoordelijke
[146] Zie ook "4.5. Welke verplichtingen gelden er wanneer persoonsgegevens worden verwerkt met het oog op wetenschappelijk onderzoek of statistische doeleinden?"

Specifiek voor AI-systemen relevante informatieverplichtingen

In artikels 12 tot 14 en in mindere mate in artikels 15 en 22 legt de AVG verschillende transparantie- of informatieverplichtingen op die voor AI-toepassingen van groot belang zijn. Schematisch kunnen deze voorgesteld worden als volgt:

AVG-BEPALING	INFORMATIE
Art. 13.2.(f) – 14.2.(g) – 15.1.(h)	Verplichting om te informeren omtrent het **bestaan en gebruik** van geautomatiseerde (individuele) besluitvorming en profilering
	Verplichting om nuttige informatie omtrent de onderliggende **logica** te geven
	Verplichting om te informeren omtrent het **belang** en de verwachte **gevolgen** van deze verwerking voor de betrokkene.
Art. 22 – overweging 71	Verplichting om **uitleg** te geven bij een geautomatiseerd individueel besluit

Twee voorafgaande opmerkingen moeten worden gemaakt alvorens deze bepalingen uit de AVG grondiger te bespreken:

- Er is een **verschil tussen de informatieverplichtingen uit artikels 12-14 en die uit artikels 15 en 22**. In het eerste geval gaat het namelijk over informatie die in principe vooraf of bij de verwerking van de persoonsgegevens dient te worden verstrekt. In het tweede geval gaat het over informatie die doorgaans pas zal worden verstrekt nadat de betrokkene er achter vraagt. Dit laatste komt ook meer overeen met wat in AI-kringen onder het verklaarbaarheid-vraagstuk wordt begrepen. We bespreken verder in deze gids artikels 15 en 22 van de AVG. Zoals eerder werd vermeld wordt er niet dieper ingegaan op het *explainability*-vraagstuk. Deze gids beperkt zich daarom tot de informatie die vooraf moet worden meegedeeld.[147]
- Deze specifieke informatieverplichtingen zijn (in principe) **enkel van toepassing indien** er sprake is van 'geautomatiseerde besluitvorming' al dan niet gebruikmakend van profilering[148], én er voor de betrokkene rechtsgevolgen zijn aan verbonden of indien de betrokkene er anderszins *in aanmerkelijke mate* wordt door getroffen. Niettemin wordt aangeraden om de hierna besproken informatie ook mee te delen in gevallen waarin de geautomatiseerde besluitvorming geen juridische of andere belangrijke gevolgen veroorzaakt. Door voldoende informatie te verschaffen aan de betrokkene, rekening houdende met de concrete context van de verwerking, zorgt een verwerkingsverantwoordelijke er immers voor dat de verwerking behoorlijk en transparant verloopt.[149]

Tot slot is het aangeraden om innovatieve **visuele en interactieve technieken** te gebruiken bij het streven

[147] Het lijkt het ons aangewezen om het explainability-vraagstuk (dus informatieverstrekking nadat een beslissing is genomen) afzonderlijk te bespreken.
[148] De AVG definieert profilering als elke vorm van geautomatiseerde verwerking van persoonsgegevens waarbij aan de hand van persoonsgegevens bepaalde persoonlijke aspecten van een natuurlijke persoon worden geëvalueerd, met name met de bedoeling zijn beroepsprestaties, economische situatie, gezondheid, persoonlijke voorkeuren, interesses, betrouwbaarheid, gedrag, locatie of verplaatsingen te analyseren of te voorspellen. Zie ook deel 5.4 over geautomatiseerde besluitvorming en profilering.
[149] Overweging 60 AVG. Dit sluit echter niet uit dat deze informatie in geval van een audit of onderzoek wel dient te worden meegedeeld aan bv. de toezichthoudende autoriteit of gerechtelijke instantie.

naar algoritmische transparantie aangezien het vaak gaat over het meedelen van complexe informatie. Moeilijk te doorgronden tekst werkt vaak ontmoedigend. Verwerkingsverantwoordelijken kunnen dit wellicht voorkomen door visuele hulpmiddelen te gebruiken bij het meedelen van informatie. Ook een **gelaagde gegevensbeschermingsverklaring** kan een nuttig instrument zijn. Er kan worden aangeraden om minstens de informatie besproken bij onderstaande punten 1 en 3 in de eerste laag op te nemen en de informatie onder punt 2 in een verdere laag.

Specifiek voor AI-systemen zijn **volgende informatieverplichtingen** relevant:[150]

1. VERPLICHTING OM TE INFORMEREN OMTRENT HET BESTAAN EN GEBRUIK VAN GEAUTOMATISEERDE (INDIVIDUELE) BESLUITVORMING EN PROFILERING

Indien verwerkingsverantwoordelijken (of de verwerkers die zij aanstellen) gebruik maken van geautomatiseerde besluitvorming bij de verwerking van persoonsgegevens, moeten zij dit aan de betrokkene **meedelen**. Betrokkenen dienen dus te worden geïnformeerd wanneer ze rechtstreeks met een AI-system interageren of wanneer ze persoonsgegevens meedelen aan dergelijke systemen.

Toepassing
Dit is bijvoorbeeld het geval indien een webshop gebruik zou maken van een *chatbot* waarmee bezoekers mee kunnen communiceren en die, indien bezoekers aan bepaalde voorwaarden voldoen, autonoom een korting of een ander voordeel kan geven. Bezoekers moeten er dus over worden geïnformeerd bij het begin van de interactie dat de conversatie met de *chatbot* een geautomatiseerd proces is zonder menselijke tussenkomst en dat ze verdere informatie in de gegevensbeschermingsverklaring kunnen vinden.

2. VERPLICHTING OM NUTTIGE INFORMATIE OMTRENT DE ONDERLIGGENDE LOGICA TE GEVEN

De complexiteit van AI-systemen kan een uitdaging zijn om beknopt weer te geven hoe een geautomatiseerd besluitvormingsproces of profilering werkt. Van een verwerkingsverantwoordelijke wordt daarom ook niet verwacht om een complexe uitleg te geven betreffende het gebruikte AI-systeem, en al zeker niet om een algoritme, de onderliggende broncode of bedrijfsgeheimen bekend te maken of in detail te beschrijven.[151] Wel komt het erop aan om een betrokkene op een **gemakkelijk verstaanbare, maar toch voldoende specifieke manier, nuttig en betekenisvol te informeren** omtrent deze vaak complexe verwerking en de onderliggende logica.[152]

Een verwerkingsverantwoordelijke dient dus geen complexe wiskundige uitleg te verschaffen over hoe de gebruikte algoritmes of ML werken. Wel dient een betrokkene op een eenvoudige, duidelijke en verstaanbare manier nuttig geïnformeerd te worden zodat de basis waarop het AI-systeem beslissingen neemt transparant wordt en de betrokkene weet welke resultaten hij/zij kan verwachten.

[150] Mochten er in de toekomst specifieke AI-audits of -certificeringen bestaan, kan aangeraden zijn om ook mee te delen wanneer dergelijke audit laatst is gebeurd of men dergelijke certificering bezit.
[151] Zie ook overweging 63 AVG.
[152] Merk in dat verband trouwens op dat overweging 58 van de AVG stelt dat dit in het bijzonder geldt voor situaties, waarin het vanwege zowel het grote aantal actoren als de technologische complexiteit van de praktijk voor een betrokkene moeilijk is te weten en te begrijpen of, door wie en met welk doel zijn persoonsgegevens worden verzameld.

Dit kan bijvoorbeeld door **volgende informatie** mee te delen:

- De categorieën van gegevens/informatie (en gerelateerde kenmerken) die werden of zullen worden gebruikt bij de (her)training, het testen of operationeel gebruik van de profilerings- of automatische besluitvormingssystemen. Dit zijn bijvoorbeeld de verzamelde persoonsgegevens en hoe die worden verzameld, de gegevenskwaliteit of ouderdom van de gegevens. Het is ook aangeraden mee te delen hoe men de nodige voorzorgen heeft genomen zodat de training- en testdata representatief waren (en blijven) voor de doelgroep(en) waarvoor men voorspellingen wil doen of beslissingen wil nemen.
- Waarom deze categorieën als relevant worden beschouwd en hun respectievelijk gewicht.[153]
- Hoe het model/elk profiel dat in het geautomatiseerde besluitvormingsproces wordt gebruikt, wordt opgebouwd, met inbegrip van eventuele relevante statistieken die in de analyse worden gebruikt.[154]
- Waarom dit profiel relevant is voor het geautomatiseerde besluitvormingsproces of welk doel er wordt nagestreefd.
- Hoe het profiel voor een beslissing over de betrokkene wordt gebruikt en welke criteria daarbij worden gehanteerd. Indien mogelijk kan een verwerkingsverantwoordelijke een aantal zaken toelichten zoals de voornaamste methodologische keuzes over onder andere de gebruikte algoritmen en/of modelstructuur, de wijze van bepalen van eventuele parameters[155] en hoe deze bijdragen aan een beslissing. Daarbij wordt ook aangeraden om de performantie of nauwkeurigheid van het onderliggende model mee te delen, zoals getest op onafhankelijke en representatieve testdata.
- In welke mate er menselijke controle en/of interventie (mogelijk) is op de verwerking.

Deze eerder eenvoudige informatie zal voor de betrokkene immers relevanter zijn dan de achterliggende wiskundige mechanismen en op die manier bijdragen tot de transparantie van de verwerking. Omgekeerd, indien betrokkenen of toezichthouders maar een beperkte en complexe uitleg zouden krijgen over de logica die achter dergelijke beslissingen ligt, is het zeer moeilijk voor hen om suboptimale (of foute) beslissingen of weinig betrouwbare modellen/systemen te herkennen waardoor ze in feite hun rechten of bevoegdheden niet kunnen laten gelden.[156] De hierboven opgesomde informatie is niettemin informatie die men typisch niet in de eerste lagen van een gelaagde gegevensbeschermingsverklaring zal vinden.[157]

Belangrijk is ook te weten dat dit zowel geldt in geval van **'statische' AI-systemen als 'lerende' AI-systemen** waar de onderliggende logica doorheen de tijd dus kan veranderen. In dat laatste geval zal de gegevensbeschermingsverklaring dus regelmatig bijgewerkt dienen te worden.

[153] Inclusief welke variabelen of eigenschappen (features) in overweging werden genomen voor inclusie in het model en welke variabelen/eigenschappen voor het definitieve model werden geselecteerd.
[154] Dit heeft bijvoorbeeld betrekking op de definitie en bepaling van klasse labels, eventueel toegepaste pre-processing, de wijze waarop betekenisvolle punten in de training- en testset werden bepaald uit de totale populatie waarover men voorspellingen wenst te maken, onzekerheden, de gebruikte trainingsmethodologieën en de frequentie van her-trainen,...
[155] Zoals bijvoorbeeld de *cut-off* van het model.
[156] Zie hierover ook deel 5.3 over de rechten van betrokkenen. In individuele gevallen en eerder in het kader van artikel 22 van de AVG kan men dit ook trachten te bereiken door een contra-feitelijke uitleg te geven waarin wordt aangegeven welke factoren zouden moeten veranderen om tot een andere uitkomst te komen.
[157] De Spaanse gegevensbeschermingsautoriteit is echter van mening dat ook dergelijke informatie in de eerste laag van een gelaagde gegevensbeschermingsverklaring zou moeten worden opgenomen.

Toepassing
In rekrutering kan een AI-systeem bijvoorbeeld een score aan CV's toekennen en daarmee de relevantie van een bepaalde CV bij een vacature weergeven. De verwerkingsverantwoordelijke dient dan de logica achter deze score duidelijk te maken. Zo kan er worden uitgelegd dat dit proces hen helpt om eerlijke en verantwoorde rekruteringsbeslissingen te nemen en kan de hierboven opgesomde informatie worden meegedeeld. Daarenboven kan de verantwoordelijke ook meedelen dat de gebruikte CV-scoringmethoden regelmatig worden getest om ervoor te zorgen dat ze eerlijk, effectief en onbevooroordeeld blijven.

3. VERPLICHTING OM TE INFORMEREN OMTRENT HET BELANG EN DE VERWACHTE GEVOLGEN VAN DEZE VERWERKING VOOR DE BETROKKENE

Deze verplichting houdt in dat een verwerkingsverantwoordelijke de betrokkene dient te informeren met betrekking tot **de beoogde of te verwachten verwerking en de wijze waarop** de geautomatiseerde besluitvorming de betrokkene kan beïnvloeden/welke gevolgen dit voor hem/haar kan hebben. Om deze informatie zinvol en verstaanbaar te maken, is het aangewezen echte, tastbare voorbeelden van het soort mogelijke effecten te geven.

Toepassing
Om het belang en de beoogde gevolgen van de verwerking in de e-commerce case te illustreren, kan een verwerkingsverantwoordelijke uitleggen dat het bezoeken van verschillende webwinkels die hetzelfde type goederen, of net een verscheidenheid aan goederen verkopen, zal resulteren in advertenties die respectievelijk een bepaald type goederen/een verscheidenheid aan goederen zullen aanbieden. Dit kan verder worden geïllustreerd door een instructieve video waarin fictieve mensen die op het internet surfen, met verschillende surfgewoontes, met elkaar worden vergeleken. Zo kan ook worden uitgelegd dat gediversifieerde surfgewoontes tot de opbouw van een vollediger persoonlijk profiel zullen leiden, wat dan weer zou kunnen resulteren in een grotere verscheidenheid aan gepersonaliseerde advertenties, maar ook bijvoorbeeld kortingen.

5.1.C. Kort toegelicht: explainability

Gelet op bovenstaande informatie is het aangeraden dat AI-ontwikkelaars een verklaarbaarheid door ontwerp-aanpak trachten te hanteren en een controleerbaarheid (*auditability*) van systemen nastreven, waarbij ze door middel van het ontwerp van een AI-systeem de (algoritmische) transparantie verzekeren en het mogelijk maken om beslissingen vlot te reconstrueren. Dit dient zeker overwogen te worden indien de betrokken AI-systemen een mogelijk negatieve impact hebben op de fundamentele rechten van personen.

Indien dergelijke aanpak niet mogelijk of wenselijk is, dienen AI-ontwikkelaars andere methoden voor handen te hebben (zoals bijvoorbeeld *reverse engineering*[158]) om zich er van te **verzekeren dat ze op een of andere manier een verklaring uit het AI-systeem** kunnen halen **nadat een bepaalde uitkomst zich heeft voorgedaan** (post-hoc interpretability). Het kunnen verklaren van een beslissing van een AI-

[158] Dit is het onderzoeken van een product (meestal software of een communicatieprotocol) om daaruit af te leiden wat de eisen zijn waaraan het product probeert te voldoen, of om de precieze interne werking ervan te achterhalen.

systeem is immers onlosmakelijk verbonden met hoe betrouwbaar mensen de technologie zullen achten.

Daarenboven kunnen er **interne stappen** worden genomen om de transparantie van AI-systemen te verhogen. Zo kan documentatie worden opgesteld met betrekking tot de ontwikkelde of gebruikte AI-modellen en die volgende informatie bevat: de vooropgestelde toepassingen, de technische principes van het model en daarmee verband houdende parameters, de gebruikte training datasets en methodologieën, wie welke rol op welk moment in het ontwikkelingsproces speelde en hoe en wanneer de prestaties van het AI-systeem werden geëvalueerd. Een eerste stap in deze richting kan het opstellen zijn van **data sheets** met betrekking tot de gebruikte datasets of **infofiches** over de beschikbare AI-modellen. Deze documentatie is best aangepast aan vooropgestelde publiek (zoals hoger management, bedrijfsjuristen,…) en strekt ertoe om interpretatie- of gebruiksfouten te vermijden. Een verwerkingsverantwoordelijke die een verwerker aanstelt die AI-systemen in zijn dienstverlening gebruikt, moet er zich van verzekeren dat ook zijn personeel toegang heeft tot dergelijke informatie.

5.2. Welke beperkingen legt de AVG op voor de bewaring van persoonsgegevens?

Essentie
Het principe van de opslagbeperking houdt in dat persoonsgegevens verwijderd of geanonimiseerd dienen te worden van zodra zij niet meer nodig zijn voor de doeleinden waarvoor zij verzameld werden.[159]
Actiepunten
◊ Bepaal bewaartermijnen voor elk type van gegevens, eventueel per verwerkingsactiviteit. Indien het niet mogelijk is om bewaartermijnen vast te leggen, bepaal dan de parameters waarmee de termijn kan worden bepaald. ◊ Overweeg regelmatig of de verwerkte persoonsgegevens nog wel nodig zijn, en zo niet, verwijder en/of anonimiseer ze. ◊ Voor archiverings-, onderzoeks- of statistische doeleinden kunnen persoonsgegevens wel langer bewaard worden, maar zij mogen dan ook enkel voor dit doel verder gebruikt worden. ◊ Houd er rekening mee dat personen het recht hebben om hun gegevens in bepaalde omstandigheden te laten verwijderen.[160] Schat in welke impact dit recht zou kunnen hebben op de werking, ontwikkeling en uitrol van de AI-toepassing en hou daar rekening mee. ◊ Dit principe heeft een zeer technisch karakter, werk daarom samen met andere afdelingen binnen de organisatie.

De AVG zelf legt **geen specifieke bewaartermijnen** op voor verschillende types van gegevens. Het is **aan de verwerkingsverantwoordelijke om dit zelf te bepalen** en zal afhankelijk zijn van hoelang de gegevens nodig zijn voor de specifieke verwerkingen. Soms zullen er evenwel wettelijke termijnen relevant zijn, bijvoorbeeld boekhoudkundige termijnen of verjaartermijnen.

[159] Art. 5, e) AVG
[160] Zie ook "5.3. Welke rechten hebben de betrokkenen als hun gegevens door AI-systemen worden verwerkt?"

De wijze van bewaring kan ook een invloed hebben op de bewaartermijnen. Zo zal de bewaartermijn voor lokaal opgeslagen gegevens, zoals in een robot of spraakassistent, minder lang moeten zijn dan voor gegevens die opgeslagen zijn op een centrale locatie, omwille van de verschillen in opslagcapaciteit en de beveiligingsmogelijkheden. Het toepassen van bewaartermijnen is deels een **technische aangelegenheid**, waardoor er **samenwerking** nodig is tussen verschillende departementen, zoals IT, business en compliance.

Toepassing
Werkgevers verzamelen tijdens sollicitaties veel gegevens van sollicitanten (CV's, motivatie- of aanbevelingsbrieven, medische keuring,...). Na de beëindiging van het aanwervingsproces moeten deze gegevens in principe worden vernietigd want het doel waarvoor ze verwerkt werden, namelijk iemand aanwerven, is bereikt. Indien de werkgever een wervingsreserve wil aanleggen, moet hij hiervoor de toestemming vragen aan de sollicitant. Ook aan de bewaring van deze reserve moet een bewaartermijn gekoppeld worden. Indien hij de gegevens van kandidaten wil bijhouden om toekomstige sollicitaties van dezelfde personen te kunnen weigeren, mag dit slechts gedurende een bepaalde in het gegevensbeschermingsbeleid bepaalde periode (bijvoorbeeld 2 jaar).

De algemene regel is dus dat persoonsgegevens **niet oneindig bewaard** mogen worden en ook niet enkel bewaard mogen omdat ze in de toekomst nog 'nuttig zouden kunnen zijn'. Er is wel een uitzondering voor archiverings-, onderzoeks- of statistische doeleinden.[161] In dit laatste geval moeten wel de gepaste technische en organisatorische maatregelen in acht worden genomen, bijvoorbeeld anonimisering, pseudonimisering of toegangsbeperking. Indien gegevens op deze basis voor onbepaalde tijd bewaard worden, mogen ze later niet voor een ander doel gebruikt worden.

Door gegevens die niet langer nodig zijn te verwijderen, **vermindert het risico dat ze irrelevant, buitensporig, onjuist of gedateerd zijn**. Ze kunnen ook niet meer het voorwerp uitmaken van een gegevenslek.

Toepassing
Een mogelijk voorbeeld heeft betrekking op rankings van kandidaturen die werden opgesteld door een AI-systeem. Dergelijke rankings verliezen (in principe) hun relevantie eenmaal de aanwervingsprocedure is afgelopen en dienen dan dus te worden verwijderd. (De nauwkeurigheidsscore die een medewerker aan een bepaalde ranking heeft gegeven, kan daarentegen langer worden bewaard.)

Het voeren van een degelijk beleid met betrekking tot het bewaren van gegevens, helpt ook om de principes van **minimale gegevensverwerking**[162] en **juistheid**[163] na te leven. Eveneens vermindert het risico dat dergelijke gegevens **per vergissing gebruikt worden**, ten nadele van de betrokkenen.

[161] Zie ook "4.5. Welke verplichtingen gelden er wanneer persoonsgegevens worden verwerkt met het oog op wetenschappelijk onderzoek of statistische doeleinden?"
[162] Volgens artikel 5.1., c) AVG moeten persoonsgegevens toereikend zijn, ter zake dienend en beperkt tot wat noodzakelijk is voor de doeleinden waarvoor zij worden verwerkt ('minimale gegevensverwerking'). Zie ook deel 4.2.
[163] Volgens artikel 5.1., d) AVG moeten persoonsgegevens juist zijn en zo nodig worden geactualiseerd; alle redelijke maatregelen moeten worden genomen om de persoonsgegevens die, gelet op de doeleinden waarvoor zij worden verwerkt, onjuist zijn, onverwijld te wissen of te rectificeren ('juistheid').

Indien een organisatie voor het verwerken van persoonsgegevens beroep doet op een **verwerker** moet deze laatste uiteraard ook voldoen aan de verplichtingen van de AVG. De verwerker moet na afloop van de verwerkingsdiensten, de persoonsgegevens ook wissen of aan de organisatie terugbezorgen. Bestaande kopieën moeten worden verwijderd, tenzij de verdere opslag verplicht is door een wet of deel uitmaakt van de dienstverlening. De organisatie heeft ook het recht om een **audit** te laten uitvoeren bij de verwerker om na te gaan of deze zijn/haar verplichtingen nakomt.

De betrokkenen hebben in bepaalde gevallen het **recht op gegevenswissing**. Bijvoorbeeld wanneer de gegevens niet langer nodig zijn voor het doel waarvoor ze verzameld zijn of wanneer de gegevens verzameld zijn op basis van toestemming, en de betrokkene zijn toestemming wenst in te trekken.

Toepassing
Personen die bij een bepaalde onderneming gesolliciteerd hebben en in eerste instantie toestemming gaven om in de werfreserve opgenomen te worden, kunnen hun toestemming op ieder moment intrekken. Bijgevolg moet de gerelateerde gegevens verwijderd worden.

5.3. Welke rechten hebben de betrokkenen als hun gegevens door AI-systemen worden verwerkt?

Essentie
De AVG verleent personen een aantal rechten waardoor ze de controle over hun gegevens kunnen behouden. Deze rechten zijn van toepassing op persoonsgegevens die in al de verschillende stappen van ontwikkeling en implementatie in een AI-systeem worden gebruikt, inclusief persoonsgegevens (1) opgenomen in de trainingsgegevens, (2) gebruikt om een voorspelling tijdens het gebruik te maken en (3) die in het model zelf zouden zitten.
Actiepunten
◊ Stel interne procedures op om te kunnen reageren op alle soorten verzoeken van personen die hun rechten willen uitoefenen. ◊ Geef gevolg aan dergelijk verzoeken van zodra het duidelijk is dat iemand zijn/haar rechten wil uitoefenen. ◊ Hou rekening met het feit dat de termijn voor het antwoorden op een verzoek (in principe) een maand is, maar dat deze in bepaalde gevallen verlengd kan worden met twee maanden. Breng de persoon hiervan op de hoogte. ◊ Besef dat verzoeken in specifieke gevallen geweigerd kunnen worden, maar dat dergelijke weigering gemotiveerd moet worden door de betrokken organisatie. Daarnaast moet er ook steeds meegedeeld worden dat een persoon het recht heeft om klacht neer te leggen bij de GBA en de mogelijkheid heeft om beroep bij de (burgerlijke) rechter in te stellen. ◊ Hou rekening met het feit dat het verstrekken van de informatie in principe kosteloos gebeurt, maar dat als het verzoek kennelijk ongegrond of buitensporig is, er een redelijke vergoeding aangerekend mag worden. Het verzoek kan ook geweigerd worden. Hierbij moet kunnen worden aangetoond dat het verzoek kennelijk ongegrond of buitensporig is. ◊ Controleer de identiteit van de verzoeker indien nodig, doch niet op onredelijke wijze.

5.3.A. Gemeenschappelijke bepalingen

De AVG specificeert niet hoe een persoon een geldig verzoek kan indienen. Het is daarom van belang om **zelf een (interne) procedure uit te stippelen** en bijvoorbeeld in een gegevensbeschermingsverklaring duidelijk aan te geven hoe personen hun rechten kunnen uitoefenen ten aanzien van de organisatie (zoals hoe zij hun verzoek moeten indienen: mail, online formulier,…). Zodra het duidelijk is dat een persoon zijn/haar rechten wenst uit te oefenen, moet de organisatie hier **binnen de maand op reageren**.

Indien deze weigert in te gaan op een verzoek moet ze dergelijke **weigering** ten aanzien van de verzoeker motiveren. Dit moet tevens binnen de maand.[164] Indien het gaat om een complex verzoek of om meerdere verzoeken tegelijkertijd , kan de termijn indien nodig **met twee maanden worden verlengd**. De organisatie moet de persoon van een verlenging op de hoogte brengen.[165]

Het verstrekken van de opgevraagde informatie aan de persoon gebeurt (in principe) **kosteloos**.[166]

Indien verzoeken van een persoon k**ennelijk ongegrond of buitensporig zijn**, bijvoorbeeld door hun repetitieve karakter, kan de organisatie de keuze maken tussen:
- een redelijke vergoeding aanrekenen in het licht van de kosten;
- weigeren gevolg te geven aan het verzoek.[167]

Het is wel aan de organisatie om aan te tonen dat het verzoek degelijk manifest ongegrond of buitensporig is.

Wanneer de organisatie een grote hoeveelheid persoonsgegevens verwerkt, moet ze de persoon voorafgaand aan de informatieverstrekking kunnen verzoeken om te preciseren **op welke informatie of welke verwerkingsactiviteiten** het verzoek betrekking heeft.[168]

De organisatie dient alle **redelijke maatregelen** te nemen om de **identiteit** van een persoon die om inzage verzoekt te **controleren**. De vereiste tot bewijs van de identiteit moet redelijk zijn, mag niet disproportioneel zijn en mag niet gebruikt worden om de uitoefening van het recht door de betrokken persoon te vertragen of te bemoeilijken. Een organisatie mag persoonsgegevens niet uitsluitend bewaren om later op eventuele verzoeken te kunnen reageren.[169]

In een AI-context bestaat de mogelijkheid dat een persoon zijn/haar r**echten uitoefent in één van de verschillende fases** in de levenscyclus van een AI-systeem dat persoonsgegevens verwerkt. Hieronder wordt kort stilgestaan bij deze verschillende fases.

[164] Art. 12.4 en overweging 59 AVG.
[165] Art. 12.3 AVG.
[166] Art. 12.5 AVG.
[167] Art. 12.5 AVG.
[168] Overweging 63 AVG.
[169] Overweging 64 AVG.

TRAINING
Met betrekking tot trainingsgegevens die bijvoorbeeld geconverteerd zijn naar een andere vorm is het minder makkelijk om ze aan een specifieke persoon te linken. Dit betekent echter niet automatisch dat het om niet-persoonsgebonden gegevens gaat. Zelfs als de gegevens geen specifieke identificatoren of contactgegevens hebben, kunnen trainingsgegevens nog altijd als persoonsgegevens beschouwd worden. Deze gegevens kunnen immers steeds gebruikt worden om een persoon te onderscheiden, op zichzelf of in combinatie met andere gegevens in het bezit van de organisatie. Zo kunnen de trainingsgegevens in een model voor de voorspelling van aankopen bijvoorbeeld een patroon van aankopen bevatten die uniek zijn voor één specifieke klant. Daarom is het belangrijk om ook met deze gegevens rekening te houden wanneer er verzoeken zijn van personen die hun rechten onder de AVG willen uitoefenen.
OUTPUT
De output van een AI-systeem kan in een profiel van een persoon worden opgeslagen en gebruikt worden om bepaalde acties met betrekking tot deze persoon te ondernemen. Zo kan het productaanbod dat een klant op een website te zien krijgt, ingegeven zijn door de output van het AI-systeem dat geïntegreerd is in zijn/haar profiel en waarop het systeem voorspellingen uitvoert. Wanneer dergelijke profielgegevens persoonsgegevens zijn, zijn ze onderworpen aan het recht van inzage, rectificatie en wissing. Waar individuele onnauwkeurigheden in de trainingsgegevens mogelijks slechts een verwaarloosbaar effect hebben op het resultaat, kan een onnauwkeurige output van een model een persoon wel rechtstreeks beïnvloeden. Een fout in de sollicitatiegegevens door de organisatie heeft wellicht geen impact op de training van het model, maar kan wel een grote impact hebben op de persoon zelf, bijvoorbeeld omdat er een fout diploma aan de persoon gekoppeld is.
MODEL
Soms kan een model een set van individuele voorbeelden bevatten die deel uitmaken van de interne logica. Dit wordt gedaan zodat het AI-systeem tijdens de operationalisering een onderscheid kan maken met of tussen nieuwe voorbeelden. Ondanks het feit dat een dergelijk model slechts een klein percentage van dergelijke voorbeelden bevat, bestaat toch de kans dat een persoon zijn/haar rechten wil uitoefenen. Daarom is het belangrijk dat dergelijke modellen het mogelijk maken om op een gemakkelijke manier deze trainingsvoorbeelden op te halen, zodat er snel gereageerd kan worden op dergelijke verzoeken.

Dergelijke verzoeken kunnen een **grote impact** hebben op een AI-systeem. In het geval van het recht tot inzage, zal er weinig tot niets aan het model moeten veranderen. Is het echter een verzoek tot rectificatie of wissen, bestaat een kleine kans dat het model waarschijnlijk opnieuw getraind moeten worden of zelfs vernietigd moeten worden, bijvoorbeeld indien de verwerkte persoonsgegevens onlosmakelijk deel uitmaken van het model.

Er bestaat ook de mogelijkheid dat persoonsgegevens **'toevallig' openbaar worden gemaakt**. In dit geval kunnen derde partijen toegang krijgen tot bepaalde elementen van de trainingsgegevens of kunnen ze

afleiden wie in deze trainingsgegevens zit, doordat ze een analyse maken van de manier waarop het model zich gedraagt. Het zal dan ook moeilijk zijn om in te gaan op verzoeken van personen. Daarom is het aangeraden om **regelmatig en proactief te evalueren** of persoonsgegevens afgeleid kunnen worden uit de modellen, zodat het risico op een toevallige openbaarmaking wordt geminimaliseerd.

5.3.B. Recht van inzage

Essentie
Het recht van inzage geeft een persoon het recht om informatie te verkrijgen over zijn/haar verwerkte persoonsgegevens en om een kopie hiervan te ontvangen. Een persoon heeft onder andere het recht om te weten voor welke doelen de gegevens verwerkt worden, over welke persoonsgegevens het gaat en aan wie de gegevens eventueel worden doorgestuurd. Het recht moet op een eenvoudige manier uitgeoefend kunnen worden. De verstrekte informatie moet beknopt, transparant en begrijpelijk zijn. Het moet in een gemakkelijk toegankelijke vorm en in duidelijke en eenvoudige taal overgebracht worden.
Actiepunten
◊ Ga na of persoonsgegevens in de trainingsgegevens en het model aanwezig zijn. Indien dit het geval is, verifieer of deze persoonsgegevens gemakkelijk opgehaald kunnen worden. ◊ Hou er rekening mee dat verzoeken tot inzage van trainingsgegevens niet als manifest ongegrond of buitensporig kunnen worden beschouwd louter omdat het moeilijker is om er op in te gaan. ◊ Hou er rekening mee dat het daarentegen niet nodig is om extra informatie te verzamelen of te bewaren alleen maar om de organisatie in staat te stellen personen binnen de trainingsgegevens te identificeren, met als enig doel conform de AVG te handelen. Het kan dus gebeuren dat de organisatie niet in staat is om de persoon in de trainingsgegevens te identificeren (en dat de persoon geen bijkomende informatie kan verstrekken die de identificatie ervan mogelijk maakt), waardoor de organisatie dus niet in de mogelijkheid is om een verzoek tot inzage in te willigen. De organisatie brengt de betrokken persoon hiervan op de hoogte. ◊ Hou er rekening mee dat een persoon ook recht heeft op inzage in de zogenaamde afgeleide gegevens, bijvoorbeeld het profiel dat u over hem/haar heeft opgesteld.

Het recht op inzage geeft personen dus het recht om **informatie te verkrijgen** over zijn/haar verwerkte persoonsgegevens en om een **kopie** hiervan te verkrijgen.[170] Een persoon heeft recht om informatie te ontvangen over het feit of de organisatie al dan niet zijn/haar persoonlijke gegevens verwerkt. Indien dit gebeurt, heeft een persoon het **recht op de volgende informatie**:

- de verwerkingsdoeleinden;
- de ontvangers of categorieën van ontvangers aan wie de persoonsgegevens zijn of zullen worden verstrekt, met name ontvangers in derde landen of internationale organisaties;
- indien mogelijk, de periode gedurende welke de persoonsgegevens naar verwachting zullen worden opgeslagen, of indien dat niet mogelijk is, de criteria om die termijn te bepalen;
- dat de persoon het recht heeft dat persoonsgegevens worden gerectificeerd of gewist, of dat de verwerking van hem betreffende persoonsgegevens wordt beperkt, alsmede het recht tegen die verwerking bezwaar te maken;
- dat de persoon het recht heeft klacht in te dienen bij een toezichthoudende autoriteit;

[170] Art. 15.1 AVG.

- wanneer de persoonsgegevens niet bij de persoon worden verzameld, alle beschikbare informatie over de bron van die gegevens;
- het bestaan van geautomatiseerde besluitvorming, met inbegrip van profilering[171];
- de beschermingsmaatregelen die van toepassing zijn indien persoonsgegevens aan een derde land of internationale organisatie worden doorgegeven[172].

Er moet ook inzage verleend worden in **afgeleide gegevens**. Dit zijn gegevens over een persoon die door de organisatie zelf zijn gegenereerd, bijvoorbeeld door data-analyse. Het kan hierbij gaan om een profiel dat de organisatie heeft opgesteld in het kader van rekrutering.

De verschafte informatie moet **beknopt, transparant en begrijpelijk zijn**. Het moet in **een gemakkelijk toegankelijke vorm** en in **duidelijke en eenvoudige taal** opgesteld zijn.

Dit recht moet bovendien op een **eenvoudige manier uitgeoefend** kunnen worden.[173] Het is daarom aangeraden om een **digitale procedure** te voorzien. De AVG staat immers toe dat als een persoon een elektronisch verzoek indient, de informatie elektronisch mag worden meegedeeld, tenzij de persoon anders verzoekt.[174]

De AVG erkent dat het recht op toegang negatieve gevolgen kan hebben voor de rechten van anderen door te stellen dat dit recht **geen afbreuk mag doen aan de rechten en vrijheden van anderen**. Overweging 63 stelt dat dit doorgetrokken kan worden voor het zakengeheim of de intellectuele eigendom. Het mag er evenwel niet toe leiden dat de toegang tot alle informatie wordt geweigerd.

5.3.C. Recht op rectificatie

Essentie
Personen hebben het recht om onjuiste persoonsgegevens te laten verbeteren en om onvolledige gegevens te laten aanvullen.
Actiepunten
◊ Hou er rekening mee dat welke stappen en procedures nodig zullen zijn, afhangt van de aard van de persoonsgegevens en het doel waarvoor zij gebruikt worden. Hoe groter het belang van de gegevens voor het trainen van een AI-systeem, hoe groter de inspanning moet zijn om na te gaan of deze wel correct zijn en indien nodig, aangepast moeten worden. ◊ Besef dat het recht om onjuiste gegevens te corrigeren ook betrekking kan hebben op trainingsgegevens gebruikt voor AI-systemen. ◊ Een verzoek tot rectificatie kan in de praktijk niet worden geweigerd omdat de organisatie van mening is dat de gebruikte gegevens minder invloed hebben op de uiteindelijke doeleinden. ◊ Informeer de persoon indien de persoonsgegevens toch correct zijn. Geef daarbij de redenen aan waarom de rectificatie werd geweigerd.

[171] Zie ook deel "5.4. Wat zegt de AVG over geautomatiseerde individuele besluitvorming en profilering en wat is de impact hiervan op AI-systemen?"
[172] Art. 15.2 AVG.
[173] Overweging 63 AVG.
[174] Art. 15.3 AVG.

Personen hebben het recht om **onjuiste persoonsgegevens** te laten **corrigeren** en om **onvolledige gegevens** te laten **aanvullen**.

Dit recht is nauw verbonden met het **principe van juistheid**.[175] Hoewel een organisatie normaal gezien de nodige stappen met betrekking tot juistheid al moet hebben genomen bij het verkrijgen van de persoonsgegevens, impliceert dit recht een specifieke verplichting om de juistheid te herbekijken wanneer daar door de betrokkene wordt om gevraagd.

Als de organisatie gegevens aan een derde partij heeft doorgegeven, moet zij deze **op de hoogte brengen** van het verzoek, tenzij dit onmogelijk is of onevenredig veel inspanning vergt.[176]

De organisatie moet de nodige stappen ondernemen om **na te gaan of de persoonsgegevens juist zijn en aanpassen** indien nodig. Hierbij moet rekening worden houden met de argumenten en het bewijs dat de persoon heeft gegeven.

Welke stappen nodig zijn, zal zoals aangehaald afhangen van de aard van de persoonsgegevens en het doel waarvoor zij gebruikt worden. Hoe **belangrijker de gegevens zijn voor het trainen** van een AI-systeem, hoe **groter de inspanning** moet zijn om na te gaan of deze wel correct zijn en/of aangepast moeten worden.

Indien een organisatie besluit dat de persoonsgegevens **correct** zijn, moet de verzoekende persoon hierover worden ingelicht. Daarbij geeft de organisatie de redenen voor de weigering.[177]

5.3.D. Recht op gegevenswissing ('recht op vergetelheid')

Essentie
Indien een organisatie geen goede reden (meer) heeft om persoonsgegevens nog langer te verwerken, moeten ze gewist worden.
Actiepunten
◊ Pseudonimiseer of anonimiseer persoonsgegevens, zodra dit mogelijk is, dan wel zodra zij niet meer verwerkt moeten worden. ◊ Hou rekening met elk verzoek tot gegevenswissing. Dit is echter geen absoluut recht, want de AVG voorziet enkele uitzonderingen.

Het **recht op gegevenswissing** betekent dat een persoon het recht heeft dat zijn/haar persoonsgegevens gewist worden door de organisatie die de persoonsgegevens verwerkt.[178] Dit recht is evenwel **niet absoluut** en geldt enkel in bepaalde omstandigheden zoals wanneer:[179]

- de persoonsgegevens niet langer nodig zijn voor de doeleinden waarvoor zij zijn verzameld/verwerkt worden;
- de persoon zijn/haar toestemming intrekt en er geen andere rechtsgrond is voor de verwerking;

[175] Art. 5.1., d) AVG.
[176] Art. 19 AVG.
[177] Art. 12.4 AVG.
[178] Art. 17 AVG.
[179] Art. 17.1 AVG.

- de persoon bezwaar maakt tegen de verwerking;
- de persoonsgegevens onrechtmatig verwerkt zijn;
- de organisatie wettelijk verplicht is om de gegevens na een bepaalde tijd te wissen;
- de persoonsgegevens bij kinderen verzameld zijn.

Toepassing
Als trainingsgegevens bijvoorbeeld niet langer nodig zijn omdat het model al getraind is, moet de organisatie aan een mogelijk verzoek voldoen. In sommige gevallen, bijvoorbeeld wanneer de ontwikkeling van het AI-systeem nog aan de gang is, kan het nodig zijn om bepaalde trainingsgegevens te bewaren met het oog op her-training, kwaliteitsonderzoek of het verfijnen en evalueren van het AI-systeem. In een dergelijke situatie moet de organisatie geval per geval nagaan of het aan een verzoek kan voldoen.

Er dient **niet te worden ingegaan** op een verzoek tot gegevenswissing in de volgende gevallen:
- indien de verwerking nodig is voor het uitoefenen van het recht op vrijheid van meningsuiting en informatie (bijvoorbeeld in een krantenartikel);
- indien de organisatie wettelijk verplicht is om de gegevens te verwerken of om een taak van algemeen belang te vervullen (bijvoorbeeld fiscale gegevens);
- indien de verwerking nodig is om redenen van algemeen belang op het gebied van volksgezondheid;
- indien de verwerking nodig is met het oog op archivering in het algemeen belang, wetenschappelijk of historisch onderzoek of statistische doeleinden (bijvoorbeeld staatsarchieven);
- indien de verwerking nodig is voor de instelling, uitoefening of onderbouwing van een rechtsvordering.[180]

Toepassing
Deze uitzonderingen zijn vermoedelijk niet relevant bij het trainen van de AI-systemen besproken in de toepassingen in deze gids. Ze kunnen wel van toepassing zijn bij andere AI-systemen en moeten dus geval per geval bekeken worden. De uitoefening van dit recht kan grote gevolgen hebben voor de ontwikkeling van AI-systemen in het algemeen. Het wissen van persoonsgegevens uit een grote set van gegevens kan immers een invloed hebben op de accuraatheid en betrouwbaarheid van het systeem. Het is daarom belangrijk om te overwegen gegevens te pseudonimiseren of anonimiseren alvorens ze gebruikt worden. Daarenboven dient ook te worden nagedacht over de mate waarin het AI-systeem en de daardoor verwerkte persoonsgegevens van elkaar gescheiden zijn of kunnen worden.

5.3.E. Recht op beperking van de verwerking

Essentie
In bepaalde omstandigheden kunnen personen vragen dat een organisatie stopt met het actief verwerken/gebruiken van hun persoonsgegevens, zonder deze gegevens te wissen.

[180] Art. 17.3 AVG.

Actiepunten
◊ Besef dat het verwerken van persoonsgegevens een breed scala aan handelingen omvat zoals het verzamelen, structureren, verspreiden en wissen van gegevens. ◊ Daarbij kan gebruik worden gemaakt van één van de volgende opties: ◊ breng de geselecteerde persoonsgegevens tijdelijk over naar een ander verwerkingssysteem; ◊ maak de geselecteerde gegevens niet beschikbaar voor gebruikers; ◊ haal gepubliceerde gegevens tijdelijk van een website.

In bepaalde omstandigheden kunnen personen vragen om de **verwerking van hun persoonsgegevens te beperken**. Ze kunnen dit recht uitoefenen in de volgende gevallen:

- een persoon betwist de juistheid van de persoonsgegevens en de organisatie controleert dit tijdens een bepaalde periode;
- de verwerking is onrechtmatig en de persoon verzet zich tegen het wissen van de persoonsgegevens en verzoekt in de plaats daarvan om een beperking van het gebruik ervan;
- de organisatie heeft de persoonsgegevens niet meer nodig voor de verwerkingsdoeleinden, maar de persoon heeft deze nodig in het kader van een juridische procedure;
- de persoon heeft bezwaar gemaakt tegen de verwerking, in afwachting van het antwoord op de vraag of de gerechtvaardigde belangen van de organisatie zwaarder wegen dan die van de persoon.[181]

Het is belangrijk dat organisaties (technische) **procedures** voorzien waardoor ze in staat zijn om de verwerking van persoonsgegevens te beperken indien dat nodig is. Ze kunnen er ook voor zorgen dat de tools die ze aankopen dit mogelijk maken. Daarbij mag niet worden vergeten dat verwerken een groot aantal handelingen omvat zoals het verzamelen, structureren, verspreiden en wissen van gegevens.

In veel gevallen zal de beperking **tijdelijk zijn** en kan deze dus opgeheven worden. De personen moeten wel op de hoogte worden gebracht alvorens de beperking van de verwerking wordt opgeheven.[182]

5.3.F. Recht op overdraagbaarheid van gegevens

Essentie
Personen hebben het recht om hun persoonsgegevens in een gestructureerde, gangbare en machine-leesbare vorm te verkrijgen. In bepaalde omstandigheden kan een persoon een organisatie ook verzoeken om zijn gegevens over te dragen.
Actiepunten
◊ Besef dat de gegevens die voor de training van een AI-systeem gebruikt worden (bijvoorbeeld koopgedrag) in veel gevallen door de persoon zelf zijn verstrekt. Dit recht zal dan van toepassing zijn indien de verwerking op toestemming of op een overeenkomst gebaseerd is.

[181] Art. 18.1 AVG.
[182] Art. 18.3 AVG.

◊ Weet dat de gegevens op zo een manier getransformeerd kunnen zijn dat ze door het algoritme gemakkelijker geanalyseerd kunnen worden. Indien deze transformatie aanzienlijk is, kan het zijn dat de gegevens niet langer worden beschouwd als zijnde 'verstrekt door de persoon' (afgeleide gegevens'). In dit geval is het niet mogelijk om dit recht uit te oefenen. De andere besproken rechten zoals inzage, en rectificatie blijven echter bestaan.
◊ Hou er rekening mee dat de originele vorm van de gegevens wel aan het recht van overdraagbaarheid onderworpen blijft.

Het recht op de overdraagbaarheid van gegevens geeft aan personen het recht om de door hun verstrekte persoonsgegevens in een **gestructureerde, gangbare en machine-leesbare vorm van de organisatie te verkrijgen**. Bovendien hebben ze het recht om die gegevens aan een andere organisatie over te dragen zonder daarbij te worden gehinderd door de organisatie aan wie de persoonsgegevens oorspronkelijk waren verstrekt.[183] Dit recht is echter **enkel van toepassing** indien:

- de toestemming van de persoon of de uitvoering van een overeenkomst de rechtsgrond voor de verwerking is; en
- de verwerking via geautomatiseerde processen verloopt.[184]

De persoon heeft eveneens het recht dat de persoonsgegevens rechtstreeks van de ene organisatie naar de andere worden **overgedragen**, indien dit technisch mogelijk is.[185] Het recht om gegevens over te dragen is enkel **van toepassing** indien deze gegevens:

- persoonsgegevens van de persoon in kwestie zijn; en
- door de persoon aan een organisatie werden verstrekt.

In veel gevallen zijn deze persoonsgegevens relatief eenvoudig te identificeren (bijvoorbeeld hun naam, email adres, telefoonnummer of leeftijd). De notie 'verstrekt' is echter ruimer dan enkel deze gevallen. Het heeft ook betrekking op gegevens die voortvloeien uit het **observeren van de activiteiten** van de persoon.

Volgens de Werkgroep 29 wordt met 'verstrekt' verwezen naar de persoonsgegevens die uit de **activiteiten van gebruikers kunnen worden opgemaakt**, zoals ruwe gegevens die door een slimme meter of andere soorten verbonden apparaten worden verwerkt, logbestanden van activiteiten, een historiek van internetgebruik of zoekopdrachten. Wat hier niet onder valt zijn gegevens die door de organisatie zijn gecreëerd.[186]

Kort samengevat kunnen de **volgende categorieën** gekwalificeerd worden als 'door de persoon verstrekt':

- gegevens die actief en bewust door de persoon zijn verstrekt (zoals een e-mailadres, gebruikersnaam en leeftijd);
- observatiegegevens die door de persoon zijn verstrekt door het gebruik van een dienst of een apparaat. Daarbij kan het bijvoorbeeld gaan om iemands zoekhistoriek, internetverkeer, gedrag op een website en locatiegegevens. Dit kan relevant zijn voor de gebruikte AI-systemen in de toepassingen besproken in deze gids.

[183] Art. 20.1 AVG.
[184] Art. 20.1 AVG.
[185] Art. 20.2 AVG.
[186] Dit kan op basis van de geobserveerde gegevens of direct ontvangen via invoer, bijvoorbeeld een gebruikersprofiel dat tot stand kwam op basis van een analyse van de verzamelde ruwe gegevens over historisch koopgedrag of bekeken vacatures.

Gegevens die de organisatie daarentegen **afleidt en reduceert** op basis van deze verstrekte gegevens, vallen hier niet onder.

Toepassing
Zo kan bijvoorbeeld het profiel dat een AI-systeem in het kader van sollicitaties voor de organisatie aanmaakt (bijvoorbeeld om te kunnen beoordelen of een persoon in aanmerking komt voor een bepaalde job) op zich niet als 'door de persoon verstrekt' worden beschouwd. Een ander voorbeeld is wanneer een webshop toelaat aan klanten om hun verkoopgeschiedenis te downloaden (observatiegegevens), maar niet de aanbevelingen die een AI-systeem op basis hiervan afleidt zodat klanten producten te zien krijgen die ze mogelijk interessant vinden (afgeleide gegevens).

Tenslotte mag het recht op overdraagbaarheid **geen afbreuk doen aan de rechten en vrijheden van derden** (bijvoorbeeld in het kader van het beroepsgeheim of private levenssfeer).[187]

5.3.G. Recht van bezwaar

Essentie
In bepaalde omstandigheden kan een persoon een organisatie vragen om persoonsgegevens niet langer te verwerken vanwege de specifieke situatie of in het geval van direct marketingdoeleinden, met inbegrip van profilering.
Actiepunten
◊ Hou er rekening mee dat de betrokkene bij direct marketing altijd het recht heeft om zonder motivering bezwaar aan te tekenen. Daardoor moet de organisatie automatisch overgaan tot de stopzetting van de verwerking voor dit doeleinde. ◊ Breng de mogelijkheid tot het uitoefenen van het recht op bezwaar duidelijk en apart van ander informatie onder de aandacht van de betrokkene.

Deze bepaling is gericht op verwerkingen die een **geldige rechtsgrond** hebben, maar die **ingaan tegen de wil** van een persoon. Het recht kan in **drie situaties** ingeroepen worden:

[187] Zijnde andere personen en gegevens die onder intellectuele eigendom en bedrijfsgeheimen vallen.

1. DE SPECIFIEKE SITUATIE VAN EEN PERSOON

Een persoon heeft te alle tijde het recht om in specifieke omstandigheden bezwaar te maken tegen de verwerking van zijn persoonsgegevens die gebaseerd is op één van de volgende rechtsgronden:

- de behartiging van de gerechtvaardigde belangen van de organisatie;
- de noodzakelijkheid voor de vervulling van een taak van algemeen belang of van een taak in het kader van de uitoefening van het openbaar gezag dat aan de organisatie is opgedragen, met inbegrip van profilering op basis van die bepalingen.[188]

Als gevolg van die specifieke situatie van een persoon primeren zijn belangen boven dewelke die als rechtsgrond dienen voor de verwerking. Daardoor kan de persoon dus bezwaar aantekenen tegen de verwerking in kwestie. De specifieke situatie van de persoon kan gebaseerd zijn op zijn rechten of vrijheden zoals bijvoorbeeld familiale omstandigheden of professionele vertrouwelijkheid, zoals advocaten.

De organisatie moet de verwerking van de persoonsgegevens dan ook staken tenzij de organisatie kan aantonen dat:

- er dwingende gerechtvaardigde gronden zijn voor de verwerking die zwaarder wegen dan de belangen, rechten en vrijheden van de persoon; of
- er gronden zijn die verband houden met de instelling, uitoefening of onderbouwing van een rechtsvordering.[189]

Het is aan de organisatie om aan te tonen dat de dwingende gerechtvaardigde belangen **zwaarder wegen** dan de belangen of de grondrechten en de fundamentele vrijheden van de persoon.[190]

2. DIRECT MARKETING

Een persoon heeft te allen tijde het recht om zonder motivering bezwaar te maken tegen de verwerking van zijn persoonsgegevens die verwerkt worden ten behoeve van direct marketing, met inbegrip van profilering die daarvoor gebruikt wordt.[191] Denk bijvoorbeeld aan een webshop die gebruikt maakt van een AI-systeem dat gepersonaliseerde reclame verstuurt op grond van het koopgedrag van klanten.

In tegenstelling tot de vorige grond, kan de organisatie hier geen tegenargumenten aanvoeren. Het is een absoluut recht en de verwerking van de persoonsgegevens moet gestaakt worden.

3. WETENSCHAPPELIJK OF HISTORISCH ONDERZOEK OF STATISTISCHE DOELEINDEN

Wanneer persoonsgegevens met het oog op wetenschappelijk of historisch onderzoek of statistische doeleinden worden verwerkt, heeft een persoon het recht om met zijn specifieke situatie verband houdende redenen bezwaar te maken tegen de verwerking van hem betreffende persoonsgegevens, tenzij de verwerking noodzakelijk is voor de uitvoering van een taak van algemeen belang.[192]

[188] Art. 21.1 AVG.
[189] Art. 21.1 AVG.
[190] Overweging 69 AVG.
[191] Art. 21.2 AVG.
[192] Art. 21.6 AVG. Zie ook deel 4.5 over de verwerking van persoonsgegevens voor wetenschappelijke of statistische

5.4. Wat zegt de AVG over geautomatiseerde individuele besluitvorming en profilering en wat is de impact hiervan op AI-systemen?

Essentie

Het is slechts in bepaalde gevallen toegestaan om AI-systemen te gebruiken die volledig geautomatiseerde beslissingen nemen, dus zonder menselijke tussenkomst, waarbij die beslissingen juridische of soortelijke gevolgen hebben voor personen. AI-systemen die louter menselijke beslissingen ondersteunen of verbeteren vallen niet onder deze beperking. De menselijke tussenkomst moet wel zinvol zijn. Situaties waarin een mens slechts pro forma tussenkomt in het systeem vallen nog steeds onder de strengere voorwaarden.[193] De mate en de kwaliteit van de menselijke beoordeling en tussenkomst voordat een definitieve beslissing wordt genomen over een persoon, is een belangrijke factor bij het bepalen of een AI-systeem al dan niet uitsluitend geautomatiseerd beslissingen neemt.[194]

Profilering is elke vorm van geautomatiseerde verwerking van persoonsgegevens waarbij aan de hand van persoonsgegevens bepaalde persoonlijke aspecten van een natuurlijke persoon worden geëvalueerd met de bedoeling om zijn/haar beroepsprestaties, economische situatie, gezondheid, persoonlijke voorkeuren, interesses, betrouwbaarheid, gedrag, locatie en/of verplaatsingen te analyseren of te voorspellen, en waaraan rechtsgevolgen zijn verbonden of die een persoon op vergelijkbare wijze aanmerkelijk treft.

AI-systemen maken vaak gebruik van de verwerking van persoonsgegevens om een beslissing te nemen of te ondersteunen. Dit kan bijvoorbeeld door HR-teams te ondersteunen bij het identificeren van kandidaten voor sollicitatiegesprekken door middel van het rangschikken van sollicitaties.

Geautomatiseerde besluitvorming heeft een ander toepassingsgebied en kan profilering gedeeltelijk overlappen of het resultaat zijn van profilering. Uitsluitend geautomatiseerde besluitvorming is het nemen van besluiten met technologische middelen en zonder menselijke tussenkomst. In principe is er een algemeen verbod op uitsluitend geautomatiseerde besluitvorming.

Hier bestaan drie uitzonderingen op, namelijk indien het besluit:
- noodzakelijk is voor de uitvoering of totstandkoming van een overeenkomst;
- toegestaan is door een wet;
- berust op de uitdrukkelijke toestemming van de betrokkene.[195]

doeleinden
[193] Zie ook: R. Binns en V. Gallo, "Automated Decision Making: the role of meaningful human reviews", https://ico.org.uk/about-the-ico/news-and-events/ai-blog-automated-decision-making-the-role-of-meaningful-human-reviews/.
[194] Zie ook: R. Binns & V. Gallo, "Automated Decision Making: the role of meaningful human reviews".
[195] Art. 22.2 AVG.

Geautomatiseerde besluitvorming waarbij bijzondere categorieën van persoonsgegevens[196] worden verwerkt is alleen toegestaan indien een van bovenstaande uitzonderingen voldaan is én de betrokken persoon zijn uitdrukkelijke toestemming heeft gegeven of de verwerking noodzakelijk is om redenen van zwaarwegend algemeen belang.

Actiepunten

- ◊ Voer een DPIA[197] uit.
- ◊ Informeer personen dat hun gegevens worden gebruikt voor dit soort verwerkingen. Dit is van toepassing voor gegevens die rechtstreeks bij de persoon zelf of via andere bronnen werden verzameld.
- ◊ Geef informatie over de onderliggende logica van het AI-systeem en de mogelijke gevolgen hiervan op personen.
- ◊ Informeer personen waarom geautomatiseerde verwerkingen gebruikt worden en wat de mogelijke resultaten en aanzienlijke gevolgen zijn, te meer omdat dergelijke verwerkingen niet zichtbaar zijn voor personen.[198]
- ◊ Stel een beleid op zodat personen een beslissing kunnen betwisten en wat de voorwaarden hiervoor zijn.
- ◊ Zorg ervoor dat een evaluatie gedaan wordt door iemand die gekwalificeerd is om de beslissing eventueel te herzien.
- ◊ Zorg dat de procedure voor personen die hun rechten willen uitoefenen simpel en gebruiksvriendelijk is.
- ◊ Onderneem volgende stappen om rekening te kunnen houden met de rechten van personen:
  - ◊ Ga na welke systeemvereisten nodig zijn om vanaf de ontwerpfase een zinvolle menselijke beoordeling mogelijk te maken.
  - ◊ Voorzie de noodzakelijke en aangepaste training/opleiding voor de medewerkers die het systeem overzien.
  - ◊ Geef personeel de nodige mogelijkheden en steun om rekening te houden met de bezorgdheden van personen, en indien nodig de beslissing van het AI-systeem aan de kant te schuiven.
- ◊ Neem volgende passende maatregelen:
  - ◊ Voer regelmatig kwaliteitscontroles van de systemen door om te waarborgen dat personen eerlijk worden behandeld en niet worden gediscrimineerd op grond van bijzondere categorieën van persoonsgegevens of op een andere manier.
  - ◊ Evalueer algoritmen en test de gebruikte AI-systemen om na te gaan of ze daadwerkelijk werken zoals bedoeld en geen discriminerende, onjuiste of ongerechtvaardigde resultaten genereren.
  - ◊ Laat evaluaties door een onafhankelijke derde partij uitvoeren, zeker wanneer de besluitvorming op basis van profilering personen in sterke mate kan treffen. Geef deze onafhankelijke derde partij alle vereiste informatie over hoe het algoritme of het ML systeem werkt.

[196] Zie ook "3.2. Wat zijn bijzondere categorieën van persoonsgegevens?"
[197] Zie ook "4.4. Wanneer moet een DPIA of GEB worden uitgevoerd voor een verwerking van persoonsgegevens door AI-systemen?"
[198] Zie ook "5.1. Welke transparantieverplichtingen legt de AVG op en wat zijn de specifieke aandachtspunten in een AI-context?"

- ◊ Verkrijg contractuele waarborgen voor algoritmen ontwikkeld door derden die aantonen dat de nodige evaluaties en tests zijn uitgevoerd en het algoritme aan de overeengekomen en wettelijke vereisten voldoet.
- ◊ Neem specifieke maatregelen ten behoeve van minimale gegevensverwerking, hanteer duidelijke bewaartermijnen voor profielen en voor persoonsgegevens die voor de samenstelling of toepassing van de profielen zijn gebruikt.
- ◊ Gebruik technieken voor anonimisering en pseudonimisering in het kader van profilering. maken en besluiten aan te vechten.
- ◊ Zorg voor manieren om de betrokkene de mogelijkheid te bieden zijn standpunt kenbaar te maken en besluiten aan te vechten.
- ◊ Voorzie een mechanisme voor menselijke tussenkomst in specifieke gevallen, bijvoorbeeld het verstrekken van een link naar een bezwaarprocedure op het moment waarop het geautomatiseerde besluit aan de betrokkene wordt medegedeeld, met overeengekomen termijnen voor herziening en een aangewezen aanspreekpunt voor vragen.

5.4.A. Voordelen en nadelen profilering en geautomatiseerde besluitvorming

Profilering en geautomatiseerde besluitvorming kunnen **zeer nuttig** zijn voor organisaties en ook voordelen hebben voor personen op veel vlakken, zeker in sectoren zoals gezondheidszorg, onderwijs, financiële dienstverlening en marketing. Ze kunnen leiden tot snellere en meer consistente beslissingen, met name beslissingen waarvoor een grote hoeveelheid gegevens moet worden geanalyseerd en beslissingen snel moeten worden genomen.

Hoewel deze technieken nuttig kunnen zijn, zijn er toch een **aantal risico's** aan verbonden:

- Personen hebben vaak geen weet van profilering en verwachten waarschijnlijk niet dat hun persoonsgegevens op deze manier gebruikt zullen worden.
- Personen begrijpen niet hoe deze processen werken en welke invloed ze kunnen hebben.
- De genomen beslissingen kunnen voor sommige personen grote nadelen met zich meebrengen. Er zijn verschillende voorbeelden die de nadelen van het gebruik van AI in bijvoorbeeld rekrutering aantonen. Het algoritme dat door Amazon werd ontwikkeld voor het scannen van sollicitatiebrieven bleek vrouwen bijvoorbeeld te benadelen.

Het is niet omdat uit de analyse van gegevens mogelijks een correlatie volgt, dat deze ook significant of zelfs relevant is. Gezien het proces enkel assumpties kan maken over iemand zijn gedrag of kenmerken, is er altijd een foutenmarge en is er een afweging nodig wat het risico is om de resultaten ook daadwerkelijk te gebruiken.

5.4.B. Profilering

Profilering is de **geautomatiseerde verwerking van persoonsgegevens ter beoordeling van persoonlijke aspecten van een natuurlijke persoon**, met name om kenmerken over beroepsprestaties, economische situatie, gezondheid, persoonlijke voorkeuren of interesses, betrouwbaarheid of gedrag, locatie of verplaatsingen van de betrokkene te analyseren of te voorspellen, waarbij aan dergelijke beoordeling rechtsgevolgen zijn verbonden of dat de persoon op vergelijkbare wijze aanmerkelijk wordt getroffen.[199]

Profilering kan hierbij gebruik maken van algoritmes en ML om correlaties te vinden tussen verschillende

[199] Overweging 71 AVG.

datasets. Organisaties gebruiken profilering doorgaans om iets te weten te komen over persoonlijke voorkeuren, gedrag te voorspellen en beslissingen over personen te nemen.

Profilering bestaat dus uit **drie elementen**:
1. Het moet een geautomatiseerde vorm van verwerking zijn.
2. Het moet betrekking hebben op persoonsgegevens.
3. Het doel van de profilering moet het evalueren van persoonlijke aspecten van een natuurlijk persoon zijn.

Een **organisatie doet aan profilering wanneer** ze:
- persoonsgegevens op grote schaal verzamelt en analyseert en daarbij algoritmes of ML gebruikt;
- associaties identificeert om tussen verschillende gedragingen en kenmerken verbanden te leggen;
- profielen aanmaakt om toe te passen op personen;
- het gedrag van personen voorspelt op basis van hun toegewezen profiel.

Organisaties verkrijgen persoonsgegevens uit **verschillende bronnen** zoals zoekopdrachten, koopgewoontes of gegevens over gedrag via smartphones, sociale media en *wearable devices*. Deze informatie wordt vervolgens geanalyseerd om personen in te delen in verschillende groepen of sectoren. Deze analyse kan tevens verbanden identificeren tussen verschillende gedragingen en kenmerken om zo profielen voor personen op te stellen. Dit profiel bestaat mogelijks vervolgens uit nieuwe (afgeleide) persoonsgegevens over deze persoon.

Een voorbeeld maakt dit duidelijk. Een webshop kan gegevens verzamelen over klanten zoals aankopen, producten die bijna gekocht werden, locatiegegevens of voorkeur voor bepaalde kleur. Een AI-systeem wordt ingezet om profielen op te stellen en klanten in bepaald segmenten in te delen, gebaseerd op hun kenmerken (zoals leeftijd, geslacht, uitgaven, en voorkeuren). Op basis hiervan kan de inhoud van de webshop gepersonaliseerd worden, bijvoorbeeld door welk type producten de klanten eerst te zien zal krijgen.

Er zijn **een aantal mogelijke manieren** waarop profilering kan worden gebruikt:
- algemene profilering;
- besluitvorming op basis van profilering; en
- uitsluitend geautomatiseerde besluitvorming, waaronder profilering, waaraan rechtsgevolgen verbonden zijn of die de betrokkene anderszins in aanmerkelijke mate treft.

Toepassing
Het verschil tussen besluitvorming op basis van profilering en uitsluitend geautomatiseerde besluitvorming kan worden toegelicht aan de hand van de volgende twee voorbeelden, waarbij een persoon online solliciteert. Enerzijds kan een mens besluiten of een positie wordt toegekend op basis van een profiel dat enkel met geautomatiseerde middelen is aangemaakt (besluitvorming op basis van profilering). Anderzijds kan een algoritme besluiten of de positie wordt toegekend en het besluit wordt automatisch aan de persoon bekendgemaakt, zonder voorafgaande en zinvolle menselijke beoordeling (uitsluitend geautomatiseerde besluitvorming, waaronder profilering).

Organisaties kunnen profilering en geautomatiseerde besluitvorming toepassen zolang zij aan **alle beginselen uit de AVG voldoen en een rechtsgrond** voor de verwerking hebben. Aanvullende beschermingsmaatregelen en beperkingen zijn van toepassing in geval van uitsluitend geautomatiseerde besluitvorming, waaronder profilering.

5.4.C. Geautomatiseerde besluitvorming: de mens beslist

Geautomatiseerde besluitvorming heeft een ander toepassingsgebied en kan profilering gedeeltelijk overlappen of het resultaat zijn van profilering.

Geautomatiseerde besluiten kunnen **op elk type gegevens gebaseerd zijn**, bijvoorbeeld:
- gegevens die **rechtstreeks** door de betrokken personen zijn verstrekt zoals antwoorden op een vragenlijst;
- gegevens over **de betrokkenen die zijn geregistreerd** zoals locatiegegevens die via een applicatie zijn verzameld;
- **afgeleide gegevens**, zoals een profiel van de persoon dat reeds is aangemaakt (bijvoorbeeld een profiel op een webshop.

Geautomatiseerde besluiten **kunnen met of zonder gebruik van profilering worden genomen**. Omgekeerd kan profilering plaatsvinden zonder dat geautomatiseerde besluiten worden genomen. Profilering en geautomatiseerde besluitvorming zijn **niet noodzakelijkerwijs gescheiden activiteiten**. Iets wat begint als een eenvoudig geautomatiseerd besluitvormingsproces, kan zich tot automatische besluitvorming ontwikkelen op basis van profilering, afhankelijk van hoe de gegevens worden gebruikt.

Een geheel geautomatiseerd systeem kan aanbevelingen over personen opleveren. Als er toch nog menselijke tussenkomst is, waarbij andere factoren in overweging worden genomen vooraleer er een definitieve beslissing wordt genomen, is dit niet uitsluitend gebaseerd op een geautomatiseerde verwerking.

5.4.D. Uitsluitend geautomatiseerde besluitvorming: het AI-systeem beslist

ALGEMEEN VERBOD

Uitsluitend geautomatiseerde besluitvorming is het nemen van besluiten met technologische middelen en **zonder enige menselijke tussenkomst**. De AVG voorziet in een **algemeen verbod** op uitsluitend op geautomatiseerde verwerking gebaseerde besluitvorming. Dit verbod is van toepassing ongeacht of de betrokkene al dan niet actie onderneemt met betrekking tot de verwerking van zijn persoonsgegevens.[200]

Het verbod is **enkel van toepassing** in het geval van:
- uitsluitend geautomatiseerde beslissingen;
- die rechtsgevolgen hebben voor iemand of de persoon anderszins in aanmerkelijke mate treffen.

Bij uitsluitend geautomatiseerde beslissingen is er geen enkele **menselijke tussenkomst** in het besluitvormingsproces. Als een dergelijk proces een aanbeveling geeft met betrekking tot een persoon, maar een mens deze aanbeveling beoordeelt en bij het nemen van het definitieve besluit rekening houdt met andere factoren is dit geen besluit 'uitsluitend gebaseerd op geautomatiseerde verwerking'.

Dit verbod kan **niet omzeild worden door een beperkte menselijke tussenkomst**. Als iemand bijvoorbeeld routinematig automatisch gegenereerde profielen toepast op personen zonder het resultaat daadwerkelijk te beïnvloeden, is dit nog steeds een uitsluitend op geautomatiseerde verwerking gebaseerd besluit.

Zoals aangehaald moet het besluit ook **rechtsgevolgen hebben** of de betrokkene **anderszins in**

[200] Artikel 22 AVG.

aanmerkelijke mate treffen. De AVG definieert deze termen niet, maar er kan wel uit afgeleid worden dat enkel ernstige, aanzienlijke effecten bedoeld worden.

Een beslissing die een rechtsgevolg heeft kan enerzijds **invloed hebben op iemands (grond)wettelijke rechten** (zoals vrijheid van vereniging, het stemrecht en het recht om rechtsmiddelen in te stellen) en anderzijds op **iemands juridische status**. Voorbeelden van een dergelijk gevolg zijn geautomatiseerde besluiten over een persoon die leiden tot bijvoorbeeld de beëindiging van een overeenkomst of het recht op of de weigering van een bepaalde wettelijk toegekende sociale uitkering waaronder kinderbijslag of huurtoeslag.

Toepassing
Een uitsluitend geautomatiseerde besluit om iemand tijdens een rekruteringsproces niet langer in het proces te betrekken, zal in principe onder het verbod vallen (zie evenwel uitzonderingen).

Een beslissing die een persoon in aanmerkelijke mate treft, heeft **gelijkaardige gevolgen** voor de situatie, gedrag of keuzes van een persoon.

Toepassing
Typische voorbeelden hiervan zijn een automatische weigering van online rekruteringsprocessen en/of verwerking van sollicitaties via internet zonder menselijke tussenkomst.[201] Een ander relevant voorbeeld is wanneer een bedrijf beslist bepaalde personen te interviewen op basis van de resultaten van een online test die naar bekwaamheid peilt. Deze beslissing heeft een belangrijk effect, aangezien ze bepaalt of iemand al dan niet voor de functie in aanmerking komt.

Indien er **onzekerheid** is of een beslissing een persoon in anderszins in aanmerkelijke mate treft, is het belangrijk om na te gaan in welke mate er **gevolgen** zijn voor:

- financiële omstandigheden;
- gezondheid;
- reputatie;
- gedrag; of
- keuzevrijheid.

[201] Overweging 71 AVG.

Toepassing

Hoewel **online reclame** die gebruik maakt geautomatiseerde instrumenten en besluitvorming op het eerste zicht niet onder het toepassingsgebied van artikel 22 AVG valt, is enige nuancering toch nodig.

Een advertentie voor een gangbare online mode-outlet op basis van een eenvoudig demografisch profiel zoals bijvoorbeeld 'vrouwen in de leeftijd van 25 tot 35 in de regio Brussel die waarschijnlijk geïnteresseerd zijn in mode en bepaalde kleding' zal een persoon niet in aanmerkelijke mate treffen.

Het kan echter voorkomen dat het besluit personen wel anderszins in aanmerkelijke mate treft, afhankelijk van de specifieke kenmerken van het geval waaronder:

- het indringende karakter van het profileringsproces, zoals opsporing via verschillende websites, apparaten en diensten;
- de verwachtingen en wensen van de betrokken personen;
- de manier waarop de advertentie wordt gepresenteerd; of
- het gebruik van kennis over de kwetsbaarheden van de benaderde betrokkenen.

UITZONDERINGEN

Er zijn **drie uitzonderingen** op het algemeen verbod op uitsluitend geautomatiseerde individuele besluitvorming, namelijk indien het besluit:

- noodzakelijk is voor de uitvoering of totstandkoming van een overeenkomst tussen de persoon en de organisatie;
- toegestaan is door een wet;
- berust op de uitdrukkelijke toestemming van de betrokkene.[202]

1. NOODZAKELIJK VOOR DE UITVOERING OF TOTSTANDKOMING VAN EEN OVEREENKOMST TUSSEN DE PERSOON EN DE ORGANISATIE

De organisatie moet kunnen aantonen of de uitsluitend geautomatiseerde verwerking noodzakelijk is voor het bereiken van haar doel, namelijk het sluiten van een overeenkomst met de natuurlijke persoon. De organisatie moet nagaan of er geen andere methode is om het doel te bereiken die voor de betrokken personen minder nadelig is. Indien er andere middelen zijn die even doeltreffend zijn en minder nadelig voor de betrokken persoon is de verwerking niet noodzakelijk.

Neem het voorbeeld van een onderneming die een vacature online plaatst. Gelet op de populariteit van de werkgever ontvangt het bedrijf duizenden sollicitaties. Vanwege het uitzonderlijk hoge aantal kandidaturen, vindt het bedrijf het praktisch onmogelijk om geschikte kandidaten te selecteren zonder eerst volledig geautomatiseerde middelen te gebruiken om ongeschikte kandidaten uit de werving te filteren. In dit geval kan geautomatiseerde besluitvorming noodzakelijk zijn om een voorselectie van mogelijke kandidaten te maken, met de bedoeling om uiteindelijk een overeenkomst met een betrokkene te sluiten.

[202] Art. 22.2 AVG.

2. TOEGESTAAN BIJ WET
Geautomatiseerde verwerkingen, waaronder profilering, kunnen in principe plaatsvinden indien een wet de toepassing ervan toelaat. De wet moet ook voorzien in passende maatregelen ter bescherming van de rechten en vrijheden en gerechtvaardigde belangen van de betrokkene.[203]
3. UITDRUKKELIJKE TOESTEMMING
De uitdrukkelijke toestemming wordt niet gedefinieerd in de AVG. Voor 'gewone' toestemming is 'een verklaring of een ondubbelzinnige actieve handeling' noodzakelijk.[204] Het is dan ook aangewezen om te verduidelijken welke extra inspanningen een organisatie heeft genomen om deze uitdrukkelijke toestemming te bekomen. De notie uitdrukkelijk verwijst naar de manier waarop toestemming is uitgebracht door een persoon en betekent dus dat een persoon een expliciete verklaring van toestemming voor die bepaalde verwerking moet geven. Een voor de hand liggende methode is een schriftelijke verklaring, die dan best ook ondertekend wordt door de betrokken persoon. In een digitale context kan dit bijvoorbeeld gebeuren door het invullen van een online formulier, het versturen van een e-mail, het scannen van een document met een handtekening of het gebruikmaken van een elektronische handtekening. Een zogenaamde tweestapsverificatie van toestemming is ook een mogelijkheid om deze expliciet te maken. Het gaat bijvoorbeeld om een e-mail waarin de persoon wordt geïnformeerd over de intentie van een webshop om bepaalde persoonsgegevens te verwerken. De webshop legt in de e-mail uit dat het toestemming vraagt voor het gebruik van een specifieke set gegevens voor een specifiek doel. Indien de persoon instemt met het gebruik van deze gegevens, vraagt de webshop een antwoord per e-mail met de verklaring 'Ik ga akkoord'. Nadat het antwoord is verzonden, ontvangt de persoon een verificatielink die moet worden aangeklikt of een SMS bericht met een verificatiecode om de instemming te bevestigen.

5.4.E. Bijzondere categorieën van gegevens en kinderen

BIJZONDERE CATEGORIEËN

Geautomatiseerde besluitvorming waarbij persoonsgegevens van bijzondere categorieën worden verwerkt, is alleen toegestaan indien aan één van bovenstaande uitzonderingen voldaan is en de persoon zijn uitdrukkelijke toestemming heeft gegeven of de verwerking noodzakelijk is om redenen van zwaarwegend algemeen belang.[205] De organisatie moet tevens passende maatregelen treffen om de rechten en vrijheden en het gerechtvaardigde belang van personen te waarborgen.

Het is belangrijk om te beseffen dat het samenbrengen van verschillende soorten persoonsgegevens gevoelige informatie over personen kan onthullen. Zo combineerde een studie 'likes' van Facebook met een simpele enquête en kon de seksuele geaardheid van mannen voorspeld worden met een nauwkeurigheid van 88%. Bovendien voorspelden ze etniciteit met 95% nauwkeurigheid en of een gebruiker christen dan wel moslim was met 82% nauwkeurigheid. Voor een dergelijk onderzoek gelden dezelfde wettelijke verplichtingen op grond van de AVG als wanneer gevoelige persoonsgegevens van

[203] Overweging 71 AVG.
[204] Art. 4, 11) AVG.
[205] Art. 22.4 AVG.

meet af aan zouden zijn verwerkt.

KINDEREN

De relevante bepalingen in de AVG maken **geen onderscheid tussen volwassenen en kinderen**. Overweging 71 vermeldt echter dat uitsluitend geautomatiseerde besluitvorming met inbegrip van profilering waaraan rechtsgevolgen verbonden zijn of die de betrokkene anderszins in aanmerkelijke mate treft **niet van toepassing mag zijn op kinderen**.

Omdat deze formulering echter niet in artikel 22 zelf staat, beschouwde de WP29 het dan ook niet als een absoluut verbod met betrekking tot kinderen. In het licht van deze overweging beveelt de WP29 wel aan dat organisaties in beginsel geen gebruik mogen maken van de uitzonderingen om dit soort verwerkingen te rechtvaardigen.

In **uitzonderlijke gevallen** kan het echter noodzakelijk zijn dat uitsluitend geautomatiseerde besluitvorming wordt toegepast, bijvoorbeeld ter bescherming van het welzijn van kinderen. In dergelijke gevallen mag de verwerking worden uitgevoerd op grond van één van de drie besproken uitzonderingen. Er moeten wel **passende beschermingsmaatregelen** worden getroffen die geschikt zijn voor kinderen. De organisatie moet ervoor zorgen dat deze maatregelen de rechten en vrijheden en gerechtvaardigde belangen van de kinderen waarvan zij gegevens verwerken daadwerkelijk beschermen.

De noodzaak van de bijzondere bescherming voor kinderen moet met name gelden voor het gebruik van persoonsgegevens van kinderen voor marketingdoeleinden of voor het opstellen van persoonlijkheids- of gebruikersprofielen en het verzamelen van persoonsgegevens over kinderen bij het gebruik van rechtstreeks aan kinderen verstrekte diensten.[206]

Organisaties moeten dus **afzien van profilering van kinderen voor marketingdoeleinden**. Kinderen zijn bijzonder kwetsbaar in de onlineomgeving en zijn eenvoudiger te beïnvloeden met reclame die afgestemd is op het surfgedrag. Bijvoorbeeld bij onlinespellen kan profilering worden gebruikt om reclame te tonen aan spelers die volgens het algoritme eerder geneigd zijn om geld aan het spel uit te geven en om meer gepersonaliseerde reclame te tonen. Kinderen begrijpen vanwege hun leeftijd en onvolwassenheid niet steeds wat de redenen van dit soort marketing zijn en welke gevolgen het voor hen kan hebben.

5.4.F. Gerelateerde rechten van betrokkenen

Ingevolge het recht van inzage hebben personen het recht om informatie te ontvangen over de gebruikte uitsluitend geautomatiseerde besluitvorming waaronder profilering. Deze informatie heeft betrekking op:[207]

- het bestaan van geautomatiseerde besluitvorming, met inbegrip van profilering;
- nuttige informatie over de onderliggende logica; en
- het belang en de verwachte gevolgen van die verwerking voor de betrokkene.

De organisatie moet in deze context de persoon informatie verstrekken over de **verwachte gevolgen van de verwerking** en een **toelichting van een specifiek besluit**. Elke betrokkene heeft dus een recht van inzage om op de hoogte te kunnen zijn van hoe de automatische besluitvorming werkt, met inbegrip van welke logica eraan ten grondslag ligt en wat de gevolgen van een dergelijke verwerking zijn, ten minste

[206] Overweging 39 AVG.
[207] Art. 15 AVG. Zie ook deel 5.3. over de rechten van de betrokkenen.

wanneer de verwerking op profilering is gebaseerd.[208]

Daarenboven moet een organisatie die geautomatiseerde individuele besluitvorming aanwendt, personen ook volgende mogelijkheden/rechten bieden:

- het recht op menselijke tussenkomst door de organisatie die de geautomatiseerde individuele besluitvorming aanwendt;
- het recht om hun standpunt kenbaar te maken; en
- het recht om het besluit aan te vechten. Uit dit laatste recht vloeit ook voort dat personen uitleg over de na een dergelijke beoordeling genomen besluit moeten kunnen krijgen.[209]

Deze maatregelen omvatten voor een persoon minstens een manier om **menselijke tussenkomst te krijgen**, zijn/haar **standpunt kenbaar te maken** en het **besluit aan te vechten**. Menselijke tussenkomst is hierbij van cruciaal belang. Een beoordeling van dergelijke verzoeken moet worden uitgevoerd door iemand die bevoegd en bekwaam is om het besluit van het AI-systeem te wijzigen. Deze persoon moet alle relevante gegevens grondig (kunnen) analyseren, samen met aanvullende informatie die de betrokkene eventueel bezorgd heeft.[210]

De organisatie moet personen een **eenvoudige manier bieden** om deze rechten uit te oefenen. Dit onderstreept nogmaals de noodzaak om **transparant te zijn over de verwerking**. De betrokkene kan een besluit alleen aanvechten of zijn standpunt kenbaar maken als hij volledig begrijpt hoe en op grond waarvan dat besluit tot stand is gekomen.[211]

[208] Overweging 63 AVG. Zie ook deel 5.1. over transparantie.
[209] Dit wordt bevestigd in overweging 71 AVG. Zie ook "5.1. Welke transparantieverplichtingen legt de AVG op en wat zijn de specifieke aandachtspunten in een AI-context?"
[210] Art. 22.3 en Overweging 71 GDPR.
[211] Zie ook "5.1. Welke transparantieverplichtingen legt de AVG op en wat zijn de specifieke aandachtspunten in een AI-context?"

Hoofdstuk 6: Besluit

6. BESLUIT

Deze verkennende gids bespreekt de toepassing van enkele bepalingen uit de AVG op het ontwerp, de ontwikkeling en het gebruik van AI-systemen. Hoewel het geen exhaustief werk is, probeert het toch op een praktische wijze een aantal fundamentele bepalingen uit de GDPR te verhelderen. Er werd gestreefd naar een meerlagige aanpak. Elk deel startte met een samenvatting en een overzicht van enkele concrete acties die organisaties en gebruikers kunnen doorlopen met betrekking tot AI-systemen.

Het Kenniscentrum Data & Maatschappij zal de komende maanden op grond van de verschillende hoofdstukken in deze gids fiches en andere praktische instrumenten verspreiden. Deze verkennende gids vormt dus zeker geen eindpunt, maar is eerder een 'levend' document dat waar nodig zal worden aangevuld en verfijnd.

Vragen, opmerkingen of feedback op/over deze gids zijn welkom en mogen rechtstreeks aan de betrokken onderzoekers worden overgemaakt.

Hoofdstuk 7: Bibliografie

7. BIBLIOGRAFIE

Hoofdstuk 2

Deskundigengroep op hoog niveau inzake kunstmatige intelligentie, "Een definitie van KI: de belangrijkste capaciteiten en wetenschappelijke disciplines", 7p., https://ec.europa.eu/digital-single-market/en/news/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines

Europese Commissie, "Witboek over kunstmatige intelligentie - een Europese benadering op basis van excellentie en vertrouwen", 19 februari 2020, 30p. https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_nl.pdf.

Europese Commissie, Mededeling van de Commissie aan het Europees Parlement, de Europese Raad, de Raad, het Europees Economisch en Sociaal Comité en het Comité van de Regio's, "Kunstmatige intelligentie voor Europa", 25 april 2018, COM(2018) 237, 23p., https://ec.europa.eu/transparency/regdoc/rep/1/2018/NL/COM-2018-237-F1-NL-MAIN-PART-1.PDF

Kenniscentrum Data & Maatschappij, Brainfood, "Hoe overleef je een gesprek over artificiële intelligentie?", https://data-en-maatschappij.ai/publicaties/brainfood-2-hoe-overleef-je-een-gesprek-over-artifici%C3%ABle-intelligentie.

L. Steels, "Artificiële intelligentie. Naar een vierde industriële revolutie?", Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten, 2017, 49p., https://www.kvab.be/nl/activiteiten/artifici%C3%ABle-intelligentie-naar-een-vierde-industri%C3%ABle-revolutie.

M.J. Vetzo, J.H. Gerards en R. Nehmelman, *Algoritmes en grondrechten*, Boom Juridisch, den Haag, 2018, 244p.

R. Nijman, "Cognitive Computing en IBM Watson – Wat is het en wat biedt het de overheid?", https://www.ibm.com/blogs/think/nl-en/2015/01/12/cognitive-computing-en-ibm-watson-wat-is-het-en-wat-biedt-het-de-overheid/. Zie ook: https://www2.cio.nl/development/85006-wat-is-cognitive-computing.

Verordening 2018/1807 van het Europees Parlement en de Raad van 14 november 2018, inzake een kader voor het vrije verkeer van niet-persoonsgebonden gegevens in de Europese Unie, PE/53/2018/REV/1, OJ L 303, 28.11.2018, p. 59-68.

Verordening 2016/679 van het Europees Parlement en de Raad van 27 april 2016 betreffende de bescherming van natuurlijke personen in verband met de verwerking van persoonsgegevens en betreffende het vrije verkeer van die gegevens en tot intrekking van Richtlijn 95/46/EG (AVG).

Hoofdstuk 3

Commissie voor de bescherming van de persoonlijke levenssfeer, "Een gids om kleine en middelgrote ondernemingen (KMO's) voor te bereiden op de Algemene Verordening Gegevensbescherming", 32p., www.gegevensbeschermingsautoriteit.be/sites/privacycommission/files/documents/KMO_NL_update.pdf.

G. LaFever, "Anonymisation does not work for big data due to lack of protection for direct & indirect identifiers and easy re-identification vs pseudonymization", 2019, gdpr.report/news/2019/08/12/anonymisation-does-not-work-for-big-data-due-to-lack-of-protection-for-direct-indirect-identifiers-and-easy-re-identification-vs-pseudonymisation.

Groep Gegevensbescherming Artikel 29, "Advies 4/2007 over het begrip persoonsgegeven", juni 2007, 28p., https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_nl.pdf.

Groep Gegevensbescherming Artikel 29, "Advies 5/2014 over anonimiseringstechnieken", april 2014, 43p., https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_nl.pdf.

Information Commissioner's Office, "Anonymisation: managing data protection risk code of practice", 2012, 108p., https://ico.org.uk/media/for-organisations/documents/1061/anonymisation-code.pdf.
Information Commissioner's Office, "Guide to the GDPR", 2019, 317 p., https://ico.org.uk/media/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr-1-0.pdf.

L. Rocher, J.M. Hendrickx en Y. de Montjoye, "Estimating the success of re-identifications in incomplete datasets using generative models", *Nature Communication*, 2019, vol. 10, nr. 3069, https://doi.org/10.1038/s41467-019-10933-3.

M. Finck en F. Pallas, "They who must not be identified – Distinguishing Personal from Non-Personal Data under the GDPR", Max Planck Institute for Innovation & Competition, Research Paper No. 19-14, 2019, 48 p., https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3462948.

Mededeling van de Commissie aan het Europese Parlement en de Raad, "Richtsnoeren over de verordening inzake een kader voor het vrije verkeer van niet-persoonsgebonden gegevens in de Europese Unie", COM/2019/250, 29 mei 2019, https://eur-lex.europa.eu/legal-content/NL/TXT/HTML/?uri=CELEX:52019DC0250&from=en.

R.P. Santifor, "Naar een meer genuanceerde benadering van 'pseudonimisering' in het privacyrecht", *P&I*, oktober 2019, nr. 5, 10p.

Verordening 2018/1807 van het Europees Parlement en de Raad van 14 november 2018 inzake een kader voor het vrije verkeer van niet-persoonsgebonden gegevens in de Europese Unie, https://eur-lex.europa.eu/legal-content/NL/TXT/HTML/?uri=CELEX:32018R1807&from=NL.

Hoofdstuk 4

Agencia Española Proteccion Datos (AEPD), "RGPD compliance of processings that embed Artificial Intelligence - An introduction", februari 2020, 49p., https://www.aepd.es/sites/default/files/2020-02/adecuacion-rgpd-ia-en_0.pdf.

Autoriteit Persoonsgegevens, "Data Protection Impact Assessment", https://autoriteitpersoonsgegevens.nl/nl/zelf-doen/data-protection-impact-assessment-dpia.

Autoriteit Persoonsgegevens, "Toezicht op AI en Algoritmes", 17 februari 2020, 11p., https://autoriteitpersoonsgegevens.nl/nl/nieuws/toezicht-op-algoritmes.

Beslissing van het Algemeen secretariaat nr. 1/2019 van 16 januari 2019, B.S. 22 maart 2019, 28512-28514, www.gegevensbeschermingsautoriteit.be/sites/privacycommission/files/documents/01_2019_AS.pdf.

Centrum voor Cybersecurity België (CCB), Cybersecurity-gids voor de KMO, 2017, 17p., https://ccb.belgium.be/nl/document/cybersecurity-gids-voor-de-kmo.

Commissie voor de bescherming van de persoonlijke levenssfeer, "Big Data Rapport", 2017, 54p., https://www.gegevensbeschermingsautoriteit.be/big-data-rapport.

Europees Agentschap voor Netwerk- en Informatiebeveiliging (ENISA), Big Data Security, Good Practices and Recommendations on the Security of Big Data Systems, December 2015, 30p., https://www.enisa.europa.eu/publications/big-data-security.

Europees Agentschap voor Netwerk- en Informatiebeveiliging (ENISA), Big Data Threat Landscape and Good Practice Guide, januari 2016, 62p., https://www.enisa.europa.eu/publications/bigdata-threat-landscape.

Europees Agentschap voor Netwerk- en Informatiebeveiliging (ENISA), Handbook on Security of Personal Data Processing, December 2017, 68p., https://www.enisa.europa.eu/publications/handbook-on-security-of-personal-data-processing.

Europees Agentschap voor Netwerk- en Informatiebeveiliging (ENISA), Privacy by design in big data, An overview of privacy enhancing technologies in the era of big data analytics, december 2015, 80p., https://www.enisa.europa.eu/news/enisa-news/privacy-by-design-in-big-data-an-overview-of-privacy-enhancing-technologies-in-the-era-of-big-data-analytics.

Europees Comité voor gegevensbescherming, Guidelines 4/2019 on Article 25 Data Protection by Design and by Default, 13 november 2019, p. 9, nrs. 32-36.

Europees Comité voor gegevensbescherming, Brief in antwoord op Sophie In't Veld, Ref: OUT2020-0004, 29 januari 2020, 6p., https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_letter_out2020_0004_intveldalgorithms_en.pdf.

Europese Commissie, "Witboek over kunstmatige intelligentie - een Europese benadering op basis van excellentie en vertrouwen", 19 februari 2020, 30p. https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_nl.pdf.

Groep Gegevensbescherming Artikel 29, "Advies 5/2014 over anonimiseringstechnieken", april 2014, 43p., https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_nl.pdf.

Groep Gegevensbescherming Artikel 29, "Richtsnoeren voor gegevensbeschermingseffectbeoordelingen en bepaling of een verwerking "waarschijnlijk een hoog risico inhoudt" in de zin van Verordening 2016/679", 248 rev.01, oktober 2017, 22 p., https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611236.

Groep Gegevensbescherming Artikel 29, "Guidelines on transparency under Regulation 2016/679", 11 april 2018, 40p., https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=622227.

Groep Gegevensbescherming Artikel 29, "Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679", 6 februari 2018, 37p., https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053.

Information Commissioner's Office, "Anonymisation: managing data protection risk code of practice", 2012, 108p., https://ico.org.uk/media/for-organisations/documents/1061/anonymisation-code.pdf.

Information Commissioner's Office, "Big data, artificial intelligence, machine learning and data protection", 4 september 2017, 114p., https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf.

Information Commissioner's Office, "Data minimisation and privacy-preserving techniques in AI systems", https://ico.org.uk/about-the-ico/news-and-events/ai-blog-data-minimisation-and-privacy-preserving-techniques-in-ai-systems.

Information Commissioner's Office, "Guide to the GDPR", 2019, 317p., https://ico.org.uk/media/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr-1-0.pdf.

Information Commissioner's Office, "Guidance on the AI auditing framework. Draft guidance for consultation", 2020, 105p., https://ico.org.uk/media/about-the-ico/consultations/2617219/guidance-on-the-ai-auditing-framework-draft-for-consultation.pdf.

International Conference of Data Protection and Privacy Commissioners (ICDPPC), "Declaration on Ethics and Data Protection in Artificial intelligence", 23 oktober 2018, 6p. https://edps.europa.eu/sites/edp/files/publication/icdppc-40th_ai-declaration_adopted_en_0.pdf.

K. Wuyts, *Privacy Threats in Software Architectures*, Ph.D., 2015.

Norwegian Data Protection Authority, "Artificial intelligence and privacy", 2018, 30 p., https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf.

Hoofdstuk 5

Agencia Española Proteccion Datos (AEPD), "RGPD compliance of processings that embed Artificial Intelligence - An introduction", februari 2020, 49p., https://www.aepd.es/sites/default/files/2020-02/adecuacion-rgpd-ia-en_0.pdf.

Article 29 Working Party, "Guidelines on transparency under Regulation 2016/679", 11 april 2018, 40p., https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=622227.

Article 29 Working Party, "Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679", 6 februari 2018, 37p., https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053.

Autoriteit Persoonsgegevens, "Toezicht op AI en Algoritmes", 17 februari 2020, 11p., https://autoriteitpersoonsgegevens.nl/nl/nieuws/toezicht-op-algoritmes.
Commissie voor de bescherming van de persoonlijke levenssfeer, "Big Data Rapport", 2017, 54p., https://www.gegevensbeschermingsautoriteit.be/big-data-rapport.

Commissie voor de bescherming van de persoonlijke levenssfeer, "Een gids om kleine en middelgrote ondernemingen (KMO's) voor te bereiden op de Algemene Verordening Gegevensbescherming", 32p., www.gegevensbeschermingsautoriteit.be/sites/privacycommission/files/documents/KMO_NL_update.pdf.

Conference of the Independent Federal and State Data Protection Supervisory Authorities of Germany (Datenschutzkonferenz, DSK), "Hambach Declaration on Artificial Intelligence – seven data protection requirements", 3 april 2019, 4p., https://www.datenschutz-berlin.de/fileadmin/user_upload/pdf/publikationen/DSK/2019/2019-DSK-Hambach_Declaration_AI-en.pdf.

Data Ethics Commission (daten ethik kommission), "Opinion of the Data Ethics Commission – executive summary", October 2019, 32p., https://www.bmjv.de/SharedDocs/Downloads/DE/Themen/Fokusthemen/Gutachten_DEK_EN.pdf?__blob=publicationFile&v=2.

Europees Comité voor gegevensbescherming, "Guidelines 05/2020 on consent under Regulation 2016/679", mei 2020, 31 p., https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_202005_consent_en.pdf.

Groep Gegevensbescherming Artikel 29, "Richtlijnen inzake het recht op gegevensoverdraagbaarheid", WP 242 rev.01, april 2017, 24 p., https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611233.

Groep Gegevensbescherming Artikel 29, "Richtsnoeren inzake geautomatiseerde individuele besluitvorming en profilering voor de toepassing van Verordening (EU) 2016/679", WP251rev.01, 2018, 47 p., https://ec.europa.eu/

newsroom/article29/item-detail.cfm?item_id=612053.

Gegevensbeschermingsautoriteit, "Hoe zit dat met de geautomatiseerde beslissingen, waaronder profilering? (Art. 22 AVG)", https://www.gegevensbeschermingsautoriteit.be/hoe-zit-dat-met-de-geautomatiseerde-beslissingen-waaronder-profilering-art-22-avg.

Information Commissioner's Office en The Alan Turing Institute, "Explaining decisions made with AI", 20 mei 2020, 136p., https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/explaining-decisions-made-with-artificial-intelligence/.

Information Commissioner's Office, "Guide to the GDPR", 2019, 317p., https://ico.org.uk/media/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr-1-0.pdf.

Information Commissioner's Office, "Big data, artificial intelligence, machine learning and data protection", 2017, 114p., https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf.

Information Commissioner's Office, "Guidance on the AI auditing framework. Draft guidance for consultation", 2020, 105p., https://ico.org.uk/media/about-the-ico/consultations/2617219/guidance-on-the-ai-auditing-framework-draft-for-consultation.pdf.

Information Commissioner's Office, "Automated individual decision-making and profiling", 2018, 23p., https://ico.org.uk/media/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/automated-decision-making-and-profiling-1-1.pdf.

Information Commissioner's Office, "What does the GDPR say about automated decision-making and profiling?", https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/automated-decision-making-and-profiling/what-does-the-gdpr-say-about-automated-decision-making-and-profiling/.

International Conference of Data Protection and Privacy Commissioners (ICDPPC), "Declaration on Ethics and Data Protection in Artificial intelligence", 23 oktober 2018, 6p. https://edps.europa.eu/sites/edp/files/publication/icdppc-40th_ai-declaration_adopted_en_0.pdf.

Joint Research Centre - European Commission (R. Hamon, H. Junklewitz, I. Sanchez), "Technical Report on Robustness and Explainability of Artificial Intelligence – from technical to policy solutions", 2020, 40p., https://publications.jrc.ec.europa.eu/repository/handle/JRC119336.

M. Mitchell (et al.), "Model Cards for Model Reporting", januari 2019, 10p., https://arxiv.org/abs/1810.03993.

Norwegian Data Protection Authority, "Artificial intelligence and privacy", 2018, 30 p., https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf.

R. Binns en V. Gallo, "Automated Decision Making: the role of meaningful human reviews", 2019, https://ico.org.uk/about-the-ico/news-and-events/ai-blog-automated-decision-making-the-role-of-meaningful-human-reviews/.

T. Gebru (et al.), "Datasheets for Datasets", maart 2020, 24p., https://arxiv.org/abs/1803.09010.

Kenniscentrum
Data & Maatschappij

Kenniscentrum Data & Maatschappij
Pleinlaan 9
1050 Brussels

info@data-en-maatschappij.ai
www.data-en-maatschappij.ai