

Implementatie AI-verordening



Handreiking inventarisatie,
identificatie en classificatie
AI-systemen

Wat is dit document?

Deze handreiking ondersteunt de impactanalyse voor de AI-verordening. Dit document bevat concrete handvatten voor het inventariseren, identificeren en classificeren van AI-systemen. Deze handvatten helpen de gebruiker bij het invullen van de impactanalyse en het bepalen van de toepasselijke vereisten.

Kan ik 100% op dit document vertrouwen?

Hoewel deze informatie zorgvuldig is opgesteld kunnen er bepaalde nuances zijn verloren, gebruik deze handreiking daarom altijd samen met de [officiële tekst van de AI-verordening](#).

De interpretatie van definities en bepalingen in de AI-verordening is nog in ontwikkeling, het is daarom waarschijnlijk dat een deel van de in dit document gegeven informatie in de loop van de tijd niet meer volledig of actueel is. De Europese Commissie heeft een richtsnoer gepubliceerd over de toepassing van het concept AI-systeem.

CIO Rijk zal eventuele hiaten, voortschrijdend inzicht, guidance van toezichthouders verwerken in een geüpdatete versie en deze via dezelfde kanalen verspreiden.

Wie heeft dit document gemaakt?

Dit document is ontwikkeld door CIO Rijk met adviseurs van het RijksICT Gilde en de interdepartementale werkgroep 'Rijksbrede implementatie AI-verordening' in opdracht van de CDO-raad.

- Joas van Ham (RijksICT Gilde)
- Nout van Deijck (RijksICT Gilde)
- Isabel van Vledder (RijksICT Gilde)
- Ouren Kuiper (Digitale Samenleving BZK)
- Lynda la Payne (Privacy team CIO Rijk BZK)
- Martha Romkes (CIO Rijk BZK)
- Esther van Laar (CIO Rijk BZK)

Speciale dank aan de leden van de CDO-raad, de interdepartementale werkgroep, en collega's die hebben bijgedragen aan de totstandkoming van dit document.

Waar kan ik terecht met vragen of feedback?

Voor feedback, vragen of opmerkingen mail: ai-verordening@minbzk.nl

Versie: 1.1 (augustus 2025)

Cover: Jamillah Knowles & Digit / <https://betterimagesofai.org/> / <https://creativecommons.org/licenses/by/4.0/>

Inhoudsopgave

1.	Inleiding en leeswijzer	4
2.	Inventarisatie AI & Algoritmes	5
2.1.	Kies een registratiemethode	5
2.2.	Begin bij bestaande inventarisaties	5
2.3.	Vul bestaande inventarisaties aan	5
3.	Identificatie: 'AI-systemen'	6
3.1.	Uitzonderingen toepassingsgebied	6
3.2.	Het begrip 'AI-systeem'	6
4.	Classificatie	8
4.1	Verboden AI-systemen	9
4.2.	Hoog risico AI-systemen	11
3.	AI-modellen voor algemene doeleinden	14
4.	AI-modellen voor algemene doeleinden met een systeemrisico	15
5.	Bepaalde AI-systemen (transparantieplichtingen)	15

1. Inleiding en leeswijzer

Inzicht in het aantal en type AI-systemen binnen de organisatie vormt een essentieel startpunt om de impact van de AI-verordening te bepalen. De verordening stelt verschillende eisen aan AI-systemen: waar sommige systemen nauwelijks aan nieuwe regels hoeven te voldoen, gelden voor andere juist strenge vereisten. Een zorgvuldige inventarisatie, identificatie en classificatie van AI-systemen binnen de organisatie is daarom cruciaal om te bepalen welke verplichtingen er gelden en welke maatregelen er nodig zijn om aan de verordening te voldoen.

De figuur hieronder beschrijft in hoofdlijnen het proces om AI-systemen in kaart te brengen die binnen de reikwijdte van de AI-verordening vallen.

1 Inventarisatie

- Algoritmes
- AI (systemen)
- Automatische besluitvorming
- Etc.

2 Identificatie

- “AI-systeem” in de zin van de AI verordening

3 Classificatie

- Verboden toepassingen
- Hoog-risico toepassingsgebieden
- Systemen met aanvullende transparantievereisten

2. Inventarisatie AI & Algoritmes

De eerste stap is het inventariseren van algoritmes die de organisatie in gebruik heeft. Het gaat hier om alle algoritmes, AI, automatische besluitvorming en autonome systemen. Dit is een goed startpunt aangezien de definitie van een AI-systeem in de AI-verordening mogelijk niet goed aansluit bij bestaande definities van het begrip "AI-systeem" in de organisatie en het algemene taalgebruik. Deze brede inventarisatie vormt de basis voor het voldoen aan de vereisten uit de AI-verordening. Daarnaast is een bredere inventarisatie van algoritmes handig bij andere inventarisaties, bijvoorbeeld in het kader van onderzoeken, andere wetgeving, of WOO-trajecten.

Deze stap is optioneel, veel organisaties hebben al een actueel overzicht. Dan kan direct gestart worden met identificatie van AI-systemen (hieronder).

Deze stap bestaat uit drie onderdelen:

- 1. Kies een registratiemethode
- 2. Begin bij bestaande inventarisaties
- 3. Vul bestaande inventarisaties aan

2.1. Kies een registratiemethode

Als er geen actueel overzicht is, overweeg dan om een duurzame registratie van algoritmes binnen de organisatie op te zetten (met duidelijke verantwoordelijkheden en registratieprocedures) om toekomstig werk te verminderen.

Het is aan te raden over de systemen tenminste de volgende informatie te registreren:

- Naam AI-systeem
- Registratie algoritme register
- Doel en omschrijving
- Overkoepelend proces
- Risico-classificatie
- Rol organisatie (Ontwikkelaar/ Gebruiksverantwoordelijke)
- Evt. Leverancier
- Status systeem

2.2. Begin bij bestaande inventarisaties

Haak als eerste aan bij bestaande inventarisaties. Zo staan er hoogstwaarschijnlijk al algoritmes geregistreerd in het algoritmeregister. Het is ook aan te raden om het verwerkingsregister van de organisatie te raadplegen. Hierin staan al verwerkingen waarin persoonsgegevens worden verwerkt en die dus mogelijk algoritmes bevatten. Verder kijk je naar gedane inventarisaties voor onderzoeken. Zo is er een onderzoek van de Algemene Rekenkamer (ARK) geweest waar een inventarisatie van AI uitgevraagd werd. Als laatste raadpleeg je reeds uitgevoerde inventarisaties voor automatische besluitvorming in het kader van artikel 22 Algemene Verordening Gegevensbescherming (AVG). Vaak wordt voor geautomatiseerde besluitvorming gebruik gemaakt van algoritmes.

Dus neem bestaande inventarisaties als startpunt om dubbel werk

te voorkomen:

- Algoritmeregister;
- AVG Verwerkingsregister (met name verwerkingen die onder art. 22 AVG vallen);
- Inventarisaties voor (ARK-)onderzoeken of Kamervragen;
- Applicatieregisters, en andere overzichten van (contracten van) software, systemen).

2.3. Vul bestaande inventarisaties aan

De meest efficiënte manier van inventariseren is sterk afhankelijk van de organisatie en de hierboven genoemde al aanwezige informatie. De Handreiking Algoritmeregister geeft een aantal tips voor de inventarisatie van de nog niet geregistreerde algoritmes (p. 18):

- Neem contact op met de IV-afdeling en eventuele data-science-teams. Zij beschikken mogelijk over relevante documentatie over algoritmes.
- Neem contact op met de control-afdeling binnen de organisatie. Zij leggen vaak al veel zaken vast en kunnen helpen bij een eerste start.
- Neem contact op met de afdeling inkoop, zij hebben inzicht in ingekochte toepassingen waarin mogelijk algoritmes worden gebruikt.
- Neem contact op met de (Chief) Privacy Officer, Functionaris Gegevensbescherming (FG), Chief Data Officer (CDO), Chief Information Security Officer (CISO). Zij hebben vaak al een goed beeld van waar zich mogelijk impactvolle algoritmes bevinden.
- Wacht niet tot het laatste moment – tot de wettelijke verplichting. In deze eerste periode is er veel hulp, kun je leren en mee ontwikkelen. Houd er rekening mee dat, afhankelijk van de organisatie, dit hele proces een aantal maanden kan duren.

Situering: overkoepelende processen

Een aanvullende tip is om bij de inventarisatie een overzicht van processen binnen de organisatie op te halen. Door de processen te bekijken kun je algoritmes die in het proces een rol spelen opsporen. Hierbij kan een bevinding uit het rapport van de Algemene Rekenkamer 'Aandacht voor algoritmes' van 2021 als handvat dienen. Hierin werd vastgesteld dat er op hoofdlijnen drie verschillende processen/doelen zijn waarvoor algoritmes worden ingezet:

- bij het automatiseren van administratie en eenvoudige wetgeving;
- voor het faciliteren en verbeteren van bedrijfsvoering en/of dienstverlening;
- voor risicogerichte controle en daarmee ook voor gerichte inzet van beschikbare mensen en middelen.

Andere processen waarin het mogelijk is dat uw organisatie algoritmes inzet:

- Beleidsontwikkeling en besluitvorming;
- Monitoring en handhaving;
- Voorspellende analyses en simulaties;
- Automatische informatieverstrekking en communicatie.

Bovenstaande lijst is niet uitputtend en geeft slechts voorbeelden van processen.

3. Identificatie: 'AI-systemen'

Nu de mogelijke AI-systemen zijn geïnventariseerd zijn moet bepaald worden of een systeem een 'AI-systeem' in de zin van de AI-verordening is. Daarvoor moet eerst worden nagegaan of de AI-verordening van toepassing is op het type systeem, vervolgens kan een korte beslisboom worden doorlopen om na te gaan of het AI-systeem aan alle kenmerken voldoet die de verordening beschrijft.

3.1. Uitzonderingen toepassingsgebied

Voordat een algoritme geïdentificeerd kan worden als 'AI-systeem', is het nuttig om na te gaan of het toepassingsgebied van het algoritme buiten de reikwijdte van de AI-verordening valt.

Er worden in de verordening een aantal soorten AI-systemen beschreven waar geen nieuwe regels voor gelden. Systemen die aan de volgende kenmerken voldoen vallen niet onder de AI-verordening:

- Militaire-, defensie-, en nationale veiligheidsdoeleinden (Art.2(3)). "Dual use" systemen moeten wel voldoen (overweging 24). Ook systemen voor rechtshandaving en openbare veiligheid (Art.3(45)) is de AI-verordening wel van toepassing.
- Internationale organisaties, en overheidsinstanties van derde landen in het kader van internationale samenwerking of overeenkomsten met de Unie of een of meer lidstaten op het gebied van rechtshandaving en justitie (Art.2(4)).
- Systemen of modellen strikt voor wetenschappelijk onderzoek of wetenschappelijke ontwikkeling
- Onderzoeks-, test- of ontwikkelingsactiviteiten met betrekking tot AI-systemen of AI-modellen voor zij in de handel worden gebracht of in gebruik worden gesteld.
- Gebruik van systemen in het kader van een louter persoonlijke niet-professionele activiteit.
- AI-systemen die worden vrijgegeven onder vrije en opensource licenties, tenzij zij in de handel worden gebracht of in gebruik worden gesteld als AI-systemen met een hoog risico of als een AI-systeem dat onder artikel 5 (verboden systemen) of artikel 50 (transparantieverplichtingen) valt.

Wel van toepassing op:

- Systemen voor rechtshandaving en openbare veiligheid (Art.3(45)).

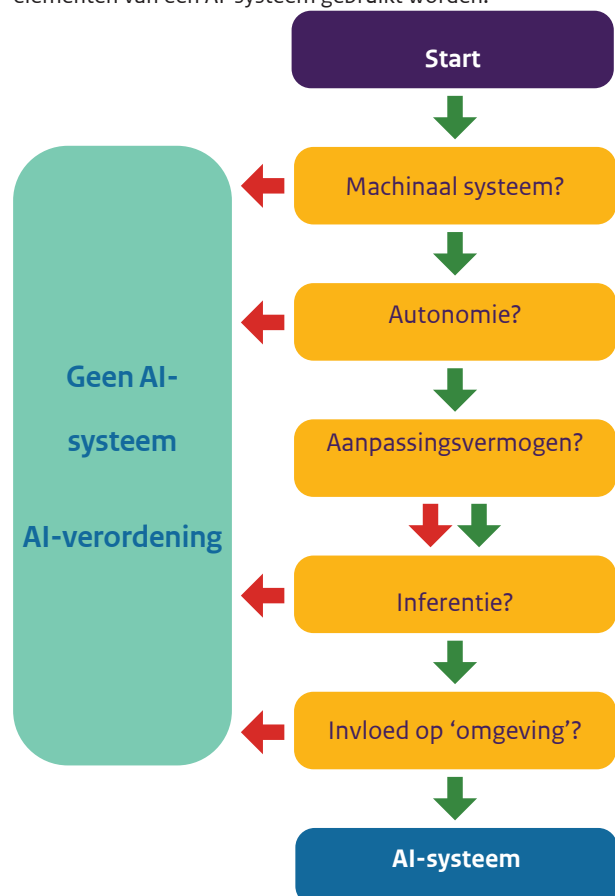
3.2. Het begrip 'AI-systeem'

Disclaimer: hoe de wettekst van de AI-verordening precies geïnterpreteerd moet worden heeft zich nog niet in de praktijk bewezen. Voor onderstaande uitleg is gebruik gemaakt van de wettekst en de richtlijnen van de Europese Commissie over de definitie van een AI-systeem. De richtlijnen zijn gemaakt om te assisteren bij het identificeren van AI-systemen, maar zijn niet bindend.

Uit de geïnventariseerde algoritmes identificeer je AI-systemen die onder de reikwijdte van de AI-verordening vallen. De definitie van 'AI-systeem' wordt als volgt gegeven in de AI-verordening: 'een op een machine gebaseerd systeem dat is ontworpen om met verschillende niveaus van autonomie te werken en dat na het inzetten ervan aanpassingsvermogen kan vertonen, en dat, voor expliciete of impliciete doelstellingen, uit de ontvangen input afleidt hoe output te genereren zoals voorspellingen, inhoud, aanbevelingen of beslissingen die van invloed kunnen zijn op fysieke of virtuele omgevingen'

Deze definitie van 'AI-systeem' in de AI-verordening bevat verschillende elementen. Bij 'aanpassingsvermogen' is het belangrijk op te merken dat een AI-systeem het kan vertonen, het is dus geen harde vereiste. Vooral op 'inferentievermogen' onderscheidt een AI-systeem zich van een niet-AI-systeem (zie uitleg verderop).

Om te bepalen of het om een 'AI-systeem' zoals bedoeld in de AI-verordening gaat, kan de onderstaande beslisboom met de elementen van een AI-systeem gebruikt worden.



Handvatten per element van het begrip 'AI-systeem' staan hieronder.

Een op een machine gebaseerd systeem...

De eerste harde afbakening in de definitie is dat een AI-systeem een op een machine gebaseerd systeem moet zijn. Ze worden dus ontwikkeld en uitgevoerd op een machine. Dit houdt in dat bijvoorbeeld handmatig uitgewerkte beslisbomen niet onder de AI-verordening vallen.

...dat is ontworpen om met verschillende niveaus van autonomie te werken...

- Autonomie gaat over de mate waarin een systeem kan leren of handelen zonder menselijke tussenkomst na het delegeren van autonomie en procesautomatisering door mensen.
- De output van AI-systemen behoort over het algemeen tot een van de algemene categorieën: voorspellingen, aanbevelingen en beslissingen. Deze categorieën komen overeen met verschillende niveaus van menselijke betrokkenheid, waarbij "besluiten" de meest autonome output is (het AI-systeem beïnvloedt zijn omgeving rechtstreeks of stuurt een andere entiteit aan om dat te doen) en "voorspellingen" de minst autonome.
- Enig niveau van autonomie betekent dus dat de output die het systeem genereert in bepaalde mate tot stand moet komen door factoren waar de menselijke gebruiker geen invloed op heeft.

...en dat na het inzetten ervan aanpassingsvermogen kan vertonen...

- Dit is vooral een illustratief onderdeel van de definitie, in de originele Engelse tekst wordt gesproken van 'may exhibit'.
- Aanpassingsvermogen wordt meestal gerelateerd aan AI-systemen op basis van machinaal leren die kunnen blijven evolueren na de initiële ontwikkeling. Het systeem past het gedrag aan door directe interactie met invoer en gegevens voor of na de inzet.
- De definitie van AI-systeem mag geen betrekking hebben op systemen die gebaseerd zijn op regels die uitsluitend door natuurlijke personen zijn vastgesteld om automatisch handelingen uit te voeren (overweging 12).

...en dat voor expliciete of impliciete doelstellingen, uit de ontvangen input afleidt hoe output te genereren...

- Het inferentievermogen van een AI-systeem overstijgt de elementaire verwerking van data door leren, redeneren of modelleren mogelijk te maken (overweging 12).
- Het concept "inferentie" verwijst over het algemeen naar de stap waarin een systeem een output genereert uit zijn inputs, meestal na de implementatie. Wanneer inferentie wordt uitgevoerd tijdens de bouwfase, wordt het in deze zin vaak gebruikt om een versie van een model te evalueren, vooral in de context van machinaal leren.
- In overweging 12 wordt het inferentievermogen beschreven als "proces waarbij output, zoals voorspellingen, content,

aanbevelingen of besluiten, wordt verkregen waarmee fysieke en virtuele omgevingen kunnen worden beïnvloed, en op het vermogen van AI-systemen om modellen of algoritmen, of beide, af te leiden uit input of data."

...die van invloed kunnen zijn op fysieke of virtuele omgevingen.

- Deze formulering dient vooral als illustratie een AI-systeem 'actief' kan zijn. De invloed van de uitkomsten beperkt zich niet alleen tot fysieke, maar ook virtuele omgevingen. Onder omgevingen moeten de contexten worden verstaan waarin de AI-systemen werken.

4. Classificatie

In de volgende fase ga je na of de geïdentificeerde AI-systemen in risicoclassificaties van de verordening vallen. Dit is van belang omdat de vereisten waaraan een AI-systeem onderworpen is, afhangt van de categorie waaronder het systeem valt. AI-systemen die voor verboden doeleinden worden ingezet dienen uitgeschakeld te worden. AI-systemen in hoog risico-toepassingsgebieden zijn onderworpen aan strenge vereisten en voor AI-systemen die kunnen zorgen voor misleidende interactie met de gebruiker gelden transparantieverplichtingen. De AI-verordening classificeert AI-systemen via drie verschillende aspecten: verboden *doeleinden*, hoog-risico *toepassingsgebieden*, en mogelijk misleidende *interactie*.

In dit hoofdstuk worden deze categorieën en eventuele uitzonderingen verder toegelicht. Ook worden de vereisten voor AI-modellen met algemene doeleinden beschreven.

Verboden doeleinden

- Sociale scoring door overheden;
- Exploitatie van kwetsbare mensen;
- Gebruik van subliminale technieken om gedrag te manipuleren;
- Realtime biometrische identificatie door wetshandhavinginstanties in openbare ruimtes;
- Voorspellende politietaken gericht op individuen;
- Emotieherkenning op de werkplek en in onderwijsinstellingen.



Hoog-risico toepassingsgebieden

- Biometrie
- Kritieke infrastructuur
- Onderwijs en beroepsopleiding
- Werkgelegenheid, personeelsbeheer en toegang tot zelfstandige arbeid
- Essentiële particuliere en publieke diensten en uitkeringen
- Rechtshandhaving
- Migratie-, asiel- en grensbeheer
- Rechtspraak en democratische processen



Mogelijk misleidende interactie

- Gebruik van chatbots die interacties met gebruikers sturen zonder dat ze zich ervan bewust zijn
- AI-systemen die content genereren zoals deepfakes of gesimuleerde video's.



4.1 Verboden AI-systemen

Het is aan te raden om als eerste na te gaan of er verboden AI-praktijken zijn. De verboden AI-systemen moeten per 1 februari 2025 zijn uitgeschakeld. Volgens artikel 5 van de AI-verordening zijn de volgende AI-systemen verboden.

4.1.1 Subliminale technieken

AI-systemen die subliminale technieken inzetten of opzettelijk manipulatieve of misleidende technieken gebruiken om het gedrag van een persoon te verstoren, waardoor hun vermogen om geïnformeerde beslissingen te nemen wordt aangetast. Bijvoorbeeld:

- Een AI-systeem dat verborgen boodschappen gebruikt in publieke mededelingen door de overheid om het stemgedrag van burgers te beïnvloeden;
- Het beïnvloeden van gamers door middel van AI om een spel langer te blijven spelen.

Belangrijke begrippen in dit verbod zijn:

- Subliminale technieken (audio-, beeld- en videostimuli die personen niet kunnen waarnemen);
- Doelbewust manipulatieve of misleidende technieken;
- Gedrag wezenlijk verstoren;
- Het vermogen een geïnformeerd besluit te nemen;
- Aanzienlijke schade.

Dit verbod doet geen afbreuk aan rechtmatige praktijken in het kader van medische behandelingen als die praktijken worden uitgevoerd overeenkomstig de toepasselijke medische normen en wetgeving (overweging 29).

4.1.2. Kwetsbaarheden

AI-systemen die kwetsbaarheden van individuen of specifieke groepen uitbuiten vanwege leeftijd, handicap of sociale/economische situatie, waardoor het gedrag van deze groep verstoord wordt of aanzienlijke schade wordt veroorzaakt. Bijvoorbeeld:

- Een AI-systeem dat gericht politieke advertenties of berichten verspreidt die speciaal zijn ontworpen om jongeren te beïnvloeden. Het systeem maakt dan misbruik van hun onervarenheid of leeftijd;
- Een AI-systeem dat manipulatieve advertenties gebruikt gericht op mensen met een verstandelijke beperking om hen aan te zetten tot ongepaste financiële beslissingen;
- Het beïnvloeden van gamende kinderen door middel van AI om een spel langer te blijven spelen.

Belangrijke begrippen in dit verbod zijn:

- Kwetsbaarheden van personen of groepen (bijvoorbeeld vanwege een handicap, sociaaleconomische situatie, of het behoren tot een etnische of religieuze minderheid);
- Gedrag wezenlijk verstoren;
- Aanzienlijke schade.

Dit verbod doet geen afbreuk aan rechtmatige praktijken in het kader van medische behandelingen wanneer die praktijken worden uitgevoerd overeenkomstig de toepasselijke medische normen en wetgeving (overweging 29)

4.1.3. Social scoring

AI-systemen die individuen of groepen evalueren of classificeren op basis van sociaal gedrag of persoonlijke kenmerken, waarbij de sociale score leidt tot schadelijke behandeling die 1) niet gerelateerd is aan de oorspronkelijke context, of 2) onevenredig is met (de ernst van) het sociale gedrag. Bijvoorbeeld:

- Een AI-systeem dat een sociale score toekent aan burgers op basis van hun gedragingen, zoals hun onlineactiviteiten, financiële geschiedenis en verkeersgedrag. Op basis van deze score reguleert de overheid de toegang tot bepaalde diensten of voordelen, zoals het verkrijgen van een lening, het kopen van vastgoed of het reizen naar het buitenland;
- Een AI-systeem dat gegevens van sociale media verzamelt van een persoon om te bepalen of iemand een werkloosheidsuitkering ontvangt.

Belangrijke begrippen in dit verbod zijn:

- Evaluatie of classificatie op basis van gedrag of voorspelde persoonlijkheidskenmerken gedurende een periode;
- Nadelige of ongunstige behandeling:
 - in een context die geen verband houdt met de context van de data (bijvoorbeeld op basis van (slecht) rijgedrag de toegang zorgtoeslag ontzeggen);
 - die ongerechtvaardigd of onevenredig is met het gedrag of de ernst ervan.

! Dit verbod doet geen afbreuk aan wettige praktijken voor de evaluatie van natuurlijke personen die worden verricht voor een specifiek doeleinde in overeenstemming met het Unierecht en het nationale recht (overweging 31)

4.1.4. Evaluatie en classificatie van personen in relatie tot strafbare feiten

AI-systemen die risicobeoordelingen van individuen maken om het plegen van strafbare feiten te voorspellen op basis van profilering of persoonlijkheidstrekken ('predictive policing').

Belangrijke begrippen in dit verbod zijn:

- Het risico dat een natuurlijke persoon een strafbaar feit pleegt te beoordelen of te voorspellen (ook wel 'predictive policing');
- Ondersteuning van de menselijke beoordeling (dit verbod geldt niet als er al een redelijk vermoeden bestaat o.b.v. daadwerkelijk gedrag, niet op basis van profilering of bepaalde kenmerken (overweging 42).

! Uitzondering: wanneer het systeem mensen ondersteunt bij het maken van een beoordeling die gebaseerd is op rechtstreeks met de criminele activiteit verband houdende objectieve en verifieerbare feiten.

4.1.5. Databanken voor gezichtsherkenning

AI-systemen die databanken voor gezichtsherkenning creëren of uitbreiden door ongerichte scraping van gezichtsafbeeldingen. Bijvoorbeeld: databanken voor gezichtsherkenning, zoals die van het Amerikaanse bedrijf Clearview, waarmee op basis van één foto een persoon geïdentificeerd kan worden in een database die 20 miljard foto's bevat.

Belangrijke begrippen in dit verbod zijn:

- Databanken voor gezichtsherkenning;
- Ongelijke scraping.

! Zie overweging 43 van de AI-verordening voor extra toelichting.

4.1.6. Emotieherkenning in werk en onderwijs

AI-systemen die emoties van individuen afleiden in werk- en onderwijsinstellingen, behalve voor medische of veiligheidsredenen. Bijvoorbeeld: het meten van stress, vermoeidheid of frustratie bij medewerkers voor managementdoeleinden zoals het in kaart brengen van productiviteit en van houding medewerkers.

Belangrijke begrippen in dit verbod zijn:

- Emoties van een natuurlijke persoon afleiden;
- Op de werkplek en in het onderwijs¹;
- Behalve wanneer bedoeld om uit medische of veiligheids-overwegingen te worden ingevoerd.

4.1.7. Biometrische categorisatie

AI-systemen die biometrische categorisatiesystemen gebruiken om persoonlijke kenmerken zoals ras, politieke opvattingen of seksuele geaardheid af te leiden en ze op basis van deze kenmerken in te delen in categorieën (behalve voor handhavingdoeleinden). Bijvoorbeeld: een systeem dat op basis van foto's van individuen een categorisering maakt van ras, politieke voorkeur, en/of geaardheid.

Belangrijke begrippen in dit verbod zijn:

- Biometrische categorisering die natuurlijke personen individueel in categorieën indelen;
- Op basis van biometrische gegevens (persoonsgegevens die het resultaat zijn van een specifieke technische verwerking met betrekking tot de fysieke, fysiologische of gedragsgerelateerde kenmerken van een natuurlijk persoon Art.3(34));
- Om sommige bijzondere categorieën van persoonsgegevens af te leiden (vergelijk Art.9 AVG).

Dit verbod geldt niet voor het rechtmatig labelen, filteren of categoriseren van reeksen biometrische gegevens zoals het sorteren van beelden op basis van haar- of oogkleur, die bijvoorbeeld op het gebied van rechtshandhaving kunnen worden gebruikt (overweging 30).

4.1.8. 'Real time' biometrische identificatie op afstand

AI-systemen die 'real-time' biometrische identificatiesystemen of afstand gebruiken in openbare ruimtes voor handhavingdoeleinden, tenzij strikt noodzakelijk voor specifieke doelstellingen zoals het voorkomen van dreigingen of het lokaliseren van criminele verdachten (hier gelden specifieke vereisten voor). Bijvoorbeeld: een AI-gestuurd gezichts-herkenningssysteem dat door de politie wordt ingezet om in real-time iedereen te identificeren en monitoren op openbare locaties voor het opsporen van mensen die afval op de grond gooien.

Belangrijke begrippen in dit verbod zijn:

- Systemen voor biometrische identificatie op afstand in real time in openbare ruimten met het oog op de rechtshandhaving.

Tenzij strikt noodzakelijk voor:

- Zoeken naar bepaalde slachtoffers of vermiste personen;
- Voorkomen van een specifieke aanzienlijke en imminente dreiging voor het leven of fysieke veiligheid, of dreiging voor een terroristische aanslag (deze dreiging kan ook voorkomen uit een ernstige verstoring van kritieke infrastructuur (overweging 33));
- Lokalisatie of identificatie van een verdachte van een strafbaar feit voor strafrechtelijk onderzoek, vervolging, of uitvoer van een straf voor feiten genoemd in Bijlage 2, met een vrijheidsstraf of -benaming van ten minste 4 jaar.

Dit verbod doet geen afbreuk aan de bepalingen voor verwerkingen van bijzondere categorieën persoonsgegevens in de AVG (art. 9).

Het gaat hier louter om het bevestigen van de identiteit van een persoon, en er moet rekening worden gehouden met proportionaliteit en de consequenties voor betrokkenen (Art.5(2)). Ook moet er vooraf een grondrechten impact assessment worden gedaan (Art.27) en moet het systeem geregistreerd zijn in de EU databank (achteraf in gerechtvaardigde gevallen).

Voor elk gebruik van eerdergenoemde AI-systemen moet een gerechtelijke of administratieve instantie (wiens besluit bindend is) toestemming geven. Tenzij er een, behoorlijk gemotiveerde, spoedeisende situatie is. Dan moet binnen 24 uur toestemming worden gevraagd.

Een besluit met nadelige rechtsgevolgen mag niet uitsluitend worden genomen o.b.v. de output van het systeem.

¹ in andere contexten is dit een hoog-risico systeem (Bijlage III(I)(c))

4.2. Hoog risico AI-systemen

De AI-verordening bepaalt dat sommige AI-systemen hoog-risico systemen zijn. Voor deze systemen zijn de meeste vereisten opgesteld. De kwalificatie 'hoog-risico' wordt toegekend aan twee soorten AI-systemen:

- Systemen die een veiligheidscomponent zijn van bepaalde gereguleerde producten, of zelf een gereguleerd product zijn. Deze producten kunnen (ook zonder AI) bepaalde risico's opleveren voor de gezondheid en veiligheid van de gebruikers. Wordt een AI-systeem als veiligheidscomponent in zo'n product gebruikt, dan moet dit AI-systeem tevens voldoen aan AI-specifieke veiligheidsstandaarden.
- Systemen die ingezet worden in een van de door de Europese Commissie aangewezen 8 hoog-risico toepassingsgebieden. Het gebruik van AI-systemen in deze contexten kan een hoog risico vormen voor de gezondheid, veiligheid of grondrechten van betrokken personen. Deze 8 gebieden zijn opgenomen in bijlage III van de AI Act.

4.2.1. Gereguleerde producten

Als een AI-systeem een gereguleerd product is in de EU of een veiligheidscomponent daarvan én onder die productregulering wordt onderworpen aan een conformiteitsbeoordeling door een derde partij dan is er sprake van een hoog risico AI-systeem voor de AI-verordening.

Deze gereguleerde producten zijn opgenomen in Bijlage 1 van de AI-verordening.

Bijlage I, afdeling A Lijst van harmonisatiewetgeving van de Unie op basis van het nieuwe wetgevingskader:

- Machinerichtlijn (2006/42/EG)
- Speelgoedrichtlijn (2009/48/EG)
- Pleziervoertuigenrichtlijn (2013/53/EU)
- Liftenrichtlijn (2014/33/EU)
- ATEX-richtlijn (2014/34/EU)
- Radioapparatuurrichtlijn (2014/53/EU)
- Drukapparatuurrichtlijn (2014/68/EU)
- Kabelbaanverordening (EU 2016/424)
- PBM-verordening (EU 2016/425)
- Gastoestellenverordening (EU 2016/426)
- Medische hulpmiddelenverordening (EU 2017/745)
- In-vitro diagnostica verordening (EU 2017/746)

Bijlage I, afdeling B Lijst van andere harmonisatiewetgeving van de Unie:

- Verordening beveiliging burgerluchtvaart (Verordening (EG) nr. 300/2008)
- Verordening goedkeuring twee- of driewielige voertuigen en vierwielers (Verordening (EU) nr. 168/2013)
- Verordening goedkeuring landbouw- en bosbouwvoertuigen (Verordening (EU) nr. 167/2013)
- Richtlijn uitrusting van zeeschepen (Richtlijn 2014/90/EU)
- Richtlijn interoperabiliteit spoorwegsysteem (Richtlijn (EU) 2016/797)

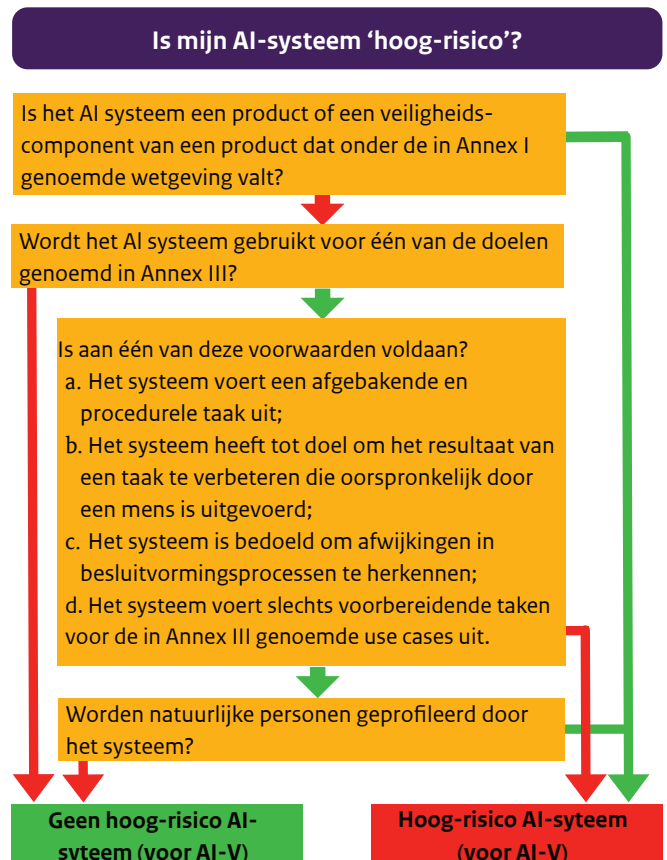
- Verordening goedkeuring motorvoertuigen (Verordening (EU) 2018/858)
- Verordening algemene veiligheid motorvoertuigen (Verordening (EU) 2019/2144)
- Verordening veiligheid burgerluchtvaart (Verordening (EU) 2018/1139)

4.2.2 Hoog-risico toepassingsgebieden

Een AI-systeem valt onder een van de toepassingsgebieden als de aanbieder het gebruik van het AI-systeem voor een toepassing binnen een van deze gebieden heeft bedoeld. De aanbieder moet het doel expliciet opnemen in de documenten van het AI-systeem, waaronder de gebruiksinstructies, reclamematerialen en verdere technische documentatie. Het beoogde doel/gebruik kan ook worden afgeleid uit het normale gebruik zoals bepaald door ontwerp en constructie van het product.

Dit betekent niet dat alle AI-systemen die binnen dat, vaak abstract omschreven, toepassingsgebied vallen hoog-risico zijn. Per gebied worden een aantal specifieke toepassingen genoemd. Als het bedoelde gebruik van het AI-systeem overeenkomt met deze toepassingen dan is er sprake van een hoog risico AI-systeem.

Tip: Ga eerst na of het AI-systeem onder een van de acht gebieden valt, en kijk vervolgens of het AI-systeem een van de beschreven AI-systemen in die categorie is. Alleen in dat geval er sprake van een hoog-risico AI-systeem dat aan eisen moet gaan voldoen.¹



¹ Gids AI-verordening versie 1.0, Ministerie van Economische Zaken, oktober 2024.

4.2.2.1. Biometrie¹

Aangezien biometrische gegevens een bijzondere categorie persoonsgegevens vormen, is het passend verschillende gevallen van kritiek gebruik van biometrische systemen te classificeren als gebruik met een hoog risico. Technische onnauwkeurigheden van deze AI-systemen kunnen tot vertekende resultaten en discriminerende effecten leiden. Het risico op dergelijke vertekende resultaten en discriminerende effecten is met name relevant wanneer zij betrekking hebben op leeftijd, etniciteit, ras, geslacht of handicap.

Uitzondering (dus geen hoog risico): een biometrisch systeem dat alleen bedoeld is om maatregelen op het gebied van cyberbeveiliging en bescherming van persoonsgegevens mogelijk te maken.

Systemen voor biometrische identificatie op afstand

- Systemen bedoeld voor het identificeren van natuurlijke personen, zonder dat zij daar actief bij betrokken zijn en doorgaans van een afstand, door middel van vergelijking van de biometrische gegevens van een persoon met de biometrische gegevens die zijn opgenomen in een referentiedatabank.
- Uitzondering (dus geen hoog risico): een systeem met als enig doel om te bevestigen dat een persoon is wie diegene beweert te zijn. Dit heet biometrische verificatie; dit is geautomatiseerde één-op-één verificatie, met inbegrip van de authenticatie, van de identiteit van natuurlijke personen door hun biometrische gegevens te vergelijken met eerder verstrekte biometrische gegevens.

Systemen voor biometrische categorisering op basis van gevoelige eigenschappen of kenmerken

- Dit is een AI-systeem dat is bedoeld voor het indelen van natuurlijke personen in specifieke categorieën op basis van hun biometrische gegevens.
- Gevoelige eigenschappen of kenmerken: bijzondere persoonsgegevens.
- Uitzondering (dus geen hoog risico): het AI-systeem vormt een aanvulling op een andere commerciële dienst én het is om objectieve technische redenen strikt noodzakelijk.

Systemen voor emotieherkenning

- Dit is een AI-systeem dat is bedoeld voor het vaststellen of afleiden van de emoties of intenties van natuurlijke personen op basis van hun biometrische gegevens.
- Als deze systemen worden ingezet op de werkvloer of in het onderwijs, dan vallen ze zelfs onder verboden AI-praktijken.

4.2.2.2. Kritieke infrastructuur

Met betrekking tot het beheer en de exploitatie van kritieke infrastructuur is het passend om deze categorie AI-systemen te classificeren als systemen met een hoog risico, aangezien het falen of gebrekkig functioneren hiervan een grootschalig risico kan opleveren voor het leven

en de gezondheid van personen en kan leiden tot aanzienlijke verstoringen van de normale beoefening van sociale en economische activiteiten.

- Systemen als veiligheidscomponent voor beheer en exploitatie van kritieke digitale infrastructuur, voor wegverkeer of voor levering van water, gas, verwarming en elektriciteit.
- Veiligheidscomponenten van kritieke infrastructuur, waaronder kritieke digitale infrastructuur, zijn systemen die worden gebruikt om de fysieke integriteit van kritieke infrastructuur of de gezondheid en veiligheid van personen en eigendommen rechtstreeks te beschermen, maar die niet noodzakelijk zijn voor de werking van het systeem.
- Het falen of gebrekkig functioneren van dergelijke componenten kan direct leiden tot risico's voor de fysieke integriteit van kritieke infrastructuur en bijgevolg tot risico's voor de gezondheid en veiligheid van personen en eigendommen.
- Uitzondering: componenten die uitsluitend voor cyberbeveiligingsdoeleinden zijn bestemd, mogen niet als veiligheidscomponenten worden aangemerkt.
- Kritieke digitale infrastructuur omvat onder meer internet exchange points, NDS-systemen en -registers, bepaalde cloud computingdiensten en content delivery netwerken. Het heeft alleen betrekking op fysieke veiligheid, dus niet op cyberveiligheid.

4.2.2.3. Onderwijs en beroepsopleiding

Wanneer systemen in de hieronder genoemde toepassingen op ondeugdelijke wijze zijn ontworpen en worden gebruikt, kunnen zij bijzonder ingrijpend zijn en in strijd zijn met het recht op onderwijs en opleiding, alsook met het recht niet te worden gediscrimineerd, en kunnen historische patronen van discriminatie in stand worden gehouden, bijvoorbeeld ten nadele van vrouwen, bepaalde leeftijdsgroepen, personen met een handicap of personen met een bepaalde raciale of etnische afkomst of seksuele gerichtheid.

- Systemen voor toegang/toelating tot of toewijzing van onderwijs.
- Systemen voor het evalueren van leerresultaten, ook wanneer die resultaten worden gebruikt voor het sturen van het leerproces van leerlingen.
- Systemen voor het beoordelen van het onderwijsniveau.
- Systemen voor het monitoren en detecteren van ongeoorloofd gedrag van leerlingen tijdens examens.

2.2.4. Werkgelegenheid, personeelsbeheer en toegang tot zelfstandige arbeid

AI-systemen die worden gebruikt op onderstaande gebieden, moeten ook als systemen met een hoog risico worden geclassificeerd, aangezien deze systemen aanzienlijke gevolgen kunnen hebben voor de toekomstige carrièrekansen en het levensonderhoud van deze personen, alsook voor de rechten van werknemers. Ook kunnen dergelijke systemen ingezet in bepaalde processen ertoe leiden dat historische patronen van

¹Voor zover het gebruik ervan is toegestaan op grond van het toepasselijke Unierecht of nationale recht.

discriminatie blijven bestaan.

- Systemen voor het werven of selecteren van kandidaten.
- Systemen voor het maken van besluiten die gevolgen hebben voor het arbeidscontract en de arbeidsvoorwaarden, voor het verdelen van taken en voor het monitoren en evalueren van werknemers.
- Dit gaat over prestatie- en talentmanagement: promoties, degradaties, prestatie-evaluatie en taakverdeling.

4.2.2.5. Essentiële particuliere en publieke diensten en uitkeringen

Deze categorie gaat over de toegang tot en het gebruik van bepaalde essentiële particuliere en openbare diensten en uitkeringen die noodzakelijk zijn voor de volledige deelname van personen aan de samenleving of voor het verbeteren van de levensstandaard. Wanneer AI-systemen worden gebruikt om te bepalen of deze uitkeringen en diensten moeten worden verleend, geweigerd, beperkt, ingetrokken of teruggevorderd door de instanties, kunnen deze systemen aanzienlijke effecten hebben op het levensonderhoud van personen en in strijd zijn met hun grondrechten.

- Systemen voor het beoordelen van (de mate van) toegang tot essentiële overheidsuitkeringen en -diensten.
- Systemen voor het beoordelen van de kredietwaardigheid of -score van personen
 - Een kredietscore is een numerieke beoordeling van de kredietwaardigheid van personen, bedrijven of andere entiteiten.
 - Kredietscore is een geautomatiseerd besluit in de zin van de AVG, wanneer de gebruiker van de score zich 'sterk baseert' op de score om een contractuele relatie met een natuurlijke persoon aan te gaan, uit te voeren of te beëindigen.
- Uitzondering (dus geen hoog risico): het systeem wordt ingezet om financiële fraude op te sporen.
- Systemen voor de risicobeoordeling en prijsbepaling voor levens- en ziektekostenverzekeringen.
 - Risicobeoordeling verwijst voornamelijk naar de beoordeling van het risico van het verzekeren van een persoon, organisatie, eigendom of activiteit en het bepalen van de prijs voor de premies en de som(men) die moet(en) worden uitbetaald. Schade-beoordeling wordt over het algemeen als een andere activiteit beschouwd.
- Systemen voor het prioriteren van de inzet van hulpdiensten of voor de triage van patiënten in de zorg.
 - Triage is het proces van patiënten-classificatie waarmee patiënten kunnen worden toegewezen aan de meest geschikte dienst voor een snellere behandeling.

4.2.2.6. Rechtshandhaving¹

Gezien hun rol en verantwoordelijkheid worden maatregelen van rechtshandavings-instanties waarbij bepaalde toepassingen van AI-systemen worden gebruikt, gekenmerkt door een aanzienlijke mate van

machtsverschillen en kunnen deze leiden tot de controle, de arrestatie of de vrijheidsberoving van natuurlijke personen, alsook tot andere negatieve effecten voor de grondrechten.

Een rechtshandavingsinstantie is:

- Iedere overheidsinstantie die bevoegd is voor de voorkoming van, het onderzoek naar, de opsporing en de vervolging van strafbare feiten of de uitvoering van straffen, inclusief de bescherming tegen en de voorkoming van gevaren voor de openbare veiligheid, of
- Ieder ander orgaan dat of iedere andere entiteit die krachtens het recht van de lidstaten is gemachtigd openbaar gezag en openbare bevoegdheden zoals hierboven beschreven, uit te oefenen.

Uitzondering (dus geen hoog risico): AI-systemen die specifiek bedoeld zijn voor gebruik voor administratieve procedures van belasting- en douaneautoriteiten en door financiële-inlichtingeneenheden die administratieve taken uitvoeren en informatie analyseren overeenkomstig de antiwitwaswetgeving van de Unie, mogen niet worden geclassificeerd als AI-systemen met een hoog risico die worden gebruikt door rechtshandavingsinstanties met het oog op het voorkomen, opsporen, onderzoeken en vervolgen van strafbare feiten.

- Systemen voor rechtshandhaving om te bepalen hoe groot de kans is dat iemand slachtoffer van strafbare feiten wordt.
- Systemen voor rechtshandhaving die worden ingezet als leugendetector.
- Systemen voor rechtshandhaving om betrouwbaarheid van bewijsmateriaal te beoordelen.
- Systemen voor rechtshandhaving om te beoordelen of te voorspellen hoe groot de kans is dat iemand een strafbaar feit zal plegen of om eerder crimineel gedrag van (groepen) personen te beoordelen.
 - Verboden indien het systeem uitsluitend beoordeelt op basis van profilering.
- Systemen voor rechtshandhaving voor het profileren van personen tijdens opsporing of vervolging.

4.2.2.7. Migratie-, asiel- en grensbeheer¹

AI-systemen die worden gebruikt op het gebied van migratie, asiel en grenstoezichtbeheer hebben gevolgen voor personen die vaak in een bijzonder kwetsbare positie verkeren en die afhankelijk zijn van de uitkomst van de acties van de bevoegde overheidsinstanties. De nauwkeurigheid, niet-discriminerende aard en transparantie van de AI-systemen die in deze context worden gebruikt, zijn derhalve van bijzonder belang om de eerbiediging van de grondrechten van de betrokken personen te waarborgen.

- Systemen voor overheidsinstanties om te gebruiken als leugendetector.
- Systemen voor overheidsinstanties om risico's voor veiligheid,

¹Voor zover het gebruik ervan is toegestaan op grond van het toepasselijke Unierecht of nationale recht.

op illegale migratie of een gezondheidsrisico bij binnenkomst van een land te beoordelen.

- Systemen om overheidsinstanties bij te staan bij het behandelen van asiel, visa of verblijfsvergunning, inclusief aanverwante klachten.
- Systemen voor overheidsinstanties voor het opsporen, herkennen of identificeren van personen.
- Meest algemene categorie op het gebied van migratie-, asiel- en grensbeheer.
- Uitzondering (dus geen hoog risico): systemen voor de verificatie van reisdocumenten. Dit verwijst meestal naar technieken voor realtime biometrie op afstand in grensgebieden, bijvoorbeeld om onrechtmatige grensoverschrijdingen op te sporen.

4.2.2.8. Rechtspraak en democratische processen

Bepaalde AI-systemen die bedoeld zijn voor de rechtsbedeling en democratische processen moeten als systemen met een hoog risico worden geclassificeerd gezien hun mogelijk aanzienlijke effecten op de democratie, de rechtsstaat, de individuele vrijheden en het recht op een doeltreffende voorziening in rechte en op een onpartijdig gerecht.

- Systemen bedoeld om door of namens een gerechtelijke instantie te worden gebruikt om 1) te ondersteunen bij het onderzoeken en uitleggen van feiten óf om de wet uit te leggen of toe te passen op een set aan feiten, of 2) om te worden gebruikt op soortgelijke wijze in het kader van alternatieve geschillenbeslechting.
- Om te bepalen of een instantie een gerechtelijke instantie is, is een aantal factoren relevant, zoals of het dwingende rechtsbevoegdheid heeft, onafhankelijk is, geen specifieke instructies aanneemt van een ministerie en de rechtsstaat toepast. Dit geldt ook voor organen die beslissen over bestuursrecht.
- Het gebruik van AI-systemen kan de beslissingsbevoegdheid van rechters of de onafhankelijkheid van de rechterlijke macht ondersteunen, maar mag deze niet vervangen: de uiteindelijke besluitvorming moet een door de mens aangestuurde activiteit blijven.
- Uitzondering (dus geen hoog risico): AI-systemen die bedoeld zijn voor louter ondersteunende administratieve activiteiten die geen gevolgen hebben voor de daadwerkelijke rechtsbedeling in afzonderlijke zaken, zoals de anonimisering of pseudonimisering van rechterlijke beslissingen, documenten of data, de communicatie tussen personeelsleden en administratieve taken.
- Systemen voor het beïnvloeden van verkiezingen of referenda of het stemgedrag van personen.
- Uitzondering (dus geen hoog risico): AI-systemen waar personen niet rechtstreeks worden blootgesteld aan de output ervan, zoals instrumenten die worden gebruikt om politieke campagnes te organiseren, te optimaliseren of te structureren vanuit administratief of logistiek oogpunt.

4.2.3. Uitzonderingen op de hoog-risico toepassingsgebieden

Er is een aantal specifieke uitzonderingen waarbij AI-systemen die onder een van de toepassingsgebieden vallen toch niet als hoog-risico AI wordt gezien. Dat is het geval wanneer het AI-systeem geen significant risico op schade voor de gezondheid, veiligheid of de grondrechten van mensen inhoudt. Onder meer wanneer een AI-systeem geen wezenlijke invloed heeft op een besluit, omdat het systeem bedoeld is om:

- Een beperkte procedurele taak uit te voeren;
- Het resultaat van een eerder voltooide menselijke activiteit te verbeteren;
- Om te controleren of menselijke besluiten afwijken van een eerder patroon zonder het besluit te vervangen of te beïnvloeden;
- Een voorbereidende taak uit te voeren voor een beoordeling die relevant is voor een van de hoog-risico toepassingsgebieden.

! Let op, om aan de bovenstaande voorwaarden te voldoen mag er geen sprake zijn van een systeem dat mensen profileert. Dan is er namelijk altijd sprake van een hoog risico AI-systeem.

3. AI-modellen voor algemene doeleinden

Een andere categorie van AI die de AI-verordening reguleert, betreft AI-modellen voor algemene doeleinden. Deze modellen zijn ontworpen om een breed scala aan taken uit te voeren. Ze worden getraind op enorme hoeveelheden data en zijn veelzijdig inzetbaar. Vanwege de brede toepasbaarheid van deze modellen, worden er veel AI-systemen op voortgebouwd. Bekende voorbeelden zijn modellen zoals GPT-4 en BERT. Deze modellen kunnen teksten, audio, video en zelfs computercode genereren. Dat maakt ze waardevol voor verschillende toepassingen.

Belangrijk om te weten is dat deze AI-modellen nog geen zelfstandige systemen zijn. Ze vormen de basis voor AI-systemen, maar zonder extra componenten zoals een gebruikersinterface zijn ze niet meer dan een krachtige tool. Voorbeelden van AI-systemen die gebruikmaken van deze modellen zijn ChatGPT, IBM Watsonx en Microsoft Copilot.

4. AI-modellen voor algemene

doeleinden met een systeemrisico

‘Systeemrisico’: een risico dat specifiek is voor de capaciteiten met een grote impact van AI-modellen voor algemene doeleinden, die aanzienlijke gevolgen hebben voor de markt van de Unie vanwege hun bereik, of vanwege feitelijke of redelijkerwijs te voorziene negatieve gevolgen voor de gezondheid, de veiligheid, de openbare veiligheid, de grondrechten of de samenleving als geheel, en dat op grote schaal in de hele waardeketen kan worden verspreid. ‘Systemisch’ impliceert dat de schade verder gaat dan geïsoleerde incidenten en verwijst naar risico’s die een keten van gevolgen hebben over partijen of markten heen (‘domino-effect’), in plaats van effecten die alleen op een klein aantal spelers van toepassing zijn.

Een AI-model voor algemene doeleinden wordt geclassificeerd als een AI-model voor algemene doeleinden met een systeemrisico als het aan een van de volgende voorwaarden voldoet:

- het beschikt over capaciteiten met een grote impact die worden geëvalueerd op basis van passende technische instrumenten en methoden, met inbegrip van indicatoren en benchmarks.
- ‘capaciteiten met een grote impact’: capaciteiten die overeenkomen met of groter zijn dan de capaciteiten die worden opgetekend bij de meest geavanceerde AI-modellen voor algemene doeleinden.
- Dit wordt verondersteld het geval te zijn wanneer de cumulatieve hoeveelheid berekening (‘compute’) die wordt gebruikt voor de training, gemeten in floating points operations (FLOPS’s) groter is dan 1025, maar kan worden vastgesteld via andere benchmarks of indicatoren.
- Er is een besluit genomen door de Commissie dat het AI-model vergelijkbare capaciteiten of een vergelijkbare impact als beschreven in punt a hierboven. Dit is voor gevallen die niet voldoen aan de algemene criteria, maar waarvan gelijkwaardige capaciteiten/impact wordt vastgesteld. Hierbij neemt de Commissie de volgende criteria in acht:
 - het aantal parameters van het model;
 - de kwaliteit of omvang van de dataset, bijvoorbeeld gemeten aan de hand van tokens;
 - de hoeveelheid rekenhulpmiddelen die is gebruikt om het model te trainen, gemeten in zwevendekomma-bewerkingen, of uitgedrukt door een combinatie van andere variabelen, zoals de geraamde kosten, tijd of energieverbruik voor de opleiding;
 - de input- en outputmodaliteiten van het model, zoals tekst-naar-tekst (grote taalmodellen), tekst-naar-beeld, multimodaliteit, en de geavanceerde drempels voor het bepalen van capaciteiten met een grote impact voor elke modaliteit, en het specifieke type input en output (bv. biologische sequenties);
 - de benchmarks en evaluaties van de capaciteiten van het model, met inbegrip van het aantal taken zonder aanvullende training, het aanpassingsvermogen om nieuwe, afzonderlijke taken te leren, de mate van autonomie en schaalbaarheid en

de instrumenten waartoe het toegang heeft;

- of het een grote impact heeft op de interne markt vanwege het bereik ervan, wat verondersteld wordt wanneer het beschikbaar is gesteld aan ten minste 10.000 geregistreerde zakelijke gebruikers die in de Unie zijn gevestigd;
- het aantal geregistreerde eindgebruikers.

5. Bepaalde AI-systemen

(transparantieverplichtingen)

De AI-verordening introduceert vanaf augustus 2026 ook transparantieverplichtingen voor bepaalde AI-systemen die bedoeld zijn om met natuurlijke personen te interageren of om content te genereren. Dit gaat dus vooral om generatieve AI-systemen. Bij dit type AI bestaat er risico op misleiding van de gebruiker. Daarom moeten de makers van deze AI-systemen ervoor zorgen dat de output duidelijk gemarkeerd wordt als kunstmatig gegenereerde of gemanipuleerde content. Dit moet in een machine-leesbaar formaat zijn, zodat het gemakkelijk te detecteren is. Dit hoeft niet expliciet wanneer dit duidelijk is vanuit het oogpunt van een normaal geïnformeerde en redelijk opmerkzaam persoon (rekening houdend met de omstandigheden en de gebruikscontext).

! Let op, deze transparantieverplichtingen kunnen gelden naast de eisen en verplichtingen voor AI-systemen met een hoog risico.

! Let op, als een AI-systeem dus niet te classificeren is als een van de bovengenoemde categorieën, dan valt het systeem of model niet onder de AI-verordening en is deze dus niet onderworpen aan de vereisten van deze wet. Dit omdat de risico’s laag of nihil zijn. Ontwikkelaars of organisaties kunnen wel vrijwillige gedragscodes opstellen en naleven.

Project
'Rijksbrede implementatie AI-
verordening'
2025