Artificiële intelligentie – Vragen en antwoorden*

Brussels, 1 augustus 2024

Waarom moeten we het gebruik van artificiële intelligentie reguleren?

De AI-verordening van de EU is de eerste uitgebreide AI-wetgeving ter wereld, en is bedoeld om het hoofd te bieden aan risico's voor de gezondheid, de veiligheid en de grondrechten. De verordening beschermt ook de democratie, de rechtsstaat en het milieu.

De invoering van AI-systemen bezit het potentieel om aanzienlijke maatschappelijke voordelen en economische groei te genereren en de innovatie en het mondiale concurrentievermogen van de EU te versterken. In bepaalde gevallen kunnen de specifieke kenmerken van bepaalde AI-systemen echter nieuwe risico's inhouden voor de (fysieke) veiligheid van de gebruiker en de grondrechten. Sommige krachtige AI-modellen die op grote schaal worden gebruikt, kunnen zelfs systeemrisico's met zich meebrengen.

Dit leidt tot rechtsonzekerheid en kan de invoering van AI-technologieën door overheden, bedrijven en burgers vertragen omdat ze er niet genoeg vertrouwen in hebben. Als nationale autoriteiten op regelgevingsgebied uiteenlopende keuzes maken, kan dit tot versnippering van de interne markt leiden.

Om die uitdagingen aan te pakken was wetgeving nodig die de goede werking van de AI-markt waarborgt door de voordelen te benutten en de risico's op gepaste wijze aan te pakken.

Op wie is de AI-verordening van toepassing?

Het rechtskader is van toepassing op zowel publieke als private actoren binnen en buiten de EU die in de Unie een **AI-systeem** in de handel brengen of die een AI-systeem gebruiken dat een impact heeft op mensen in de Unie.

De verplichtingen kunnen zowel betrekking hebben op aanbieders (bv. een ontwikkelaar van een tool om cv's te screenen) als op gebruiksverantwoordelijken van AI-systemen (bv. een bank die een dergelijke screeningtool koopt). Er zijn bepaalde uitzonderingen op de verordening. Activiteiten op het gebied van onderzoek, ontwikkeling en prototyping die plaatsvinden voordat een AI-systeem in de handel wordt gebracht, vallen niet onder deze voorschriften. Daarnaast zijn AI-systemen die uitsluitend zijn ontworpen voor militaire, defensie- of nationale veiligheidsdoeleinden ook vrijgesteld, ongeacht het type entiteit dat deze activiteiten uitvoert.

Welke risicocategorieën zijn er?

De AI-verordening voert een uniform kader voor alle EU-lidstaten in, dat is gebaseerd op een toekomstgerichte definitie van AI en een risicogebaseerde aanpak:

- **Onaanvaardbaar risico**: Een zeer beperkt aantal bijzonder schadelijke toepassingen van AI die in strijd zijn met de waarden van de EU omdat die de grondrechten schenden en daarom verboden worden:
  - **uitbuiting van kwetsbaarheden van personen, manipulatie en gebruik van subliminale technieken;**
  - **sociale scoring** voor de overheid en de particuliere sector;
  - **individueel voorspellend politiewerk** dat uitsluitend gebaseerd is op profilering van mensen;
  - het **niet-gericht scrapen** van internet of CCTV-beelden voor gezichtsopnamen om databanken op te bouwen of uit te breiden;
  - **emotieherkenning op de werkplek en in onderwijsinstellingen**, tenzij om medische of veiligheidsredenen (zoals monitoring van de vermoeidheid van een piloot);
  - **biometrische categorisering** van natuurlijke personen om hun ras, politieke, religieuze of levensbeschouwelijke overtuiging, vakbondslidmaatschap of seksuele geaardheid af te

leiden. Het labelen of filteren van datasets en de categorisering van gegevens op het gebied van rechtshandhaving blijven mogelijk;

- **biometrische identificatie op afstand in realtime in openbare ruimten door rechtshandhavingsinstanties**, met zeer beperkte uitzonderingen (zie hieronder).

- De Commissie zal richtsnoeren over de verbodsbepalingen uitvaardigen voordat die op 2 februari 2025 in werking treden.
- **Hoog risico**: Een beperkt aantal AI-systemen die in het voorstel worden gedefinieerd en die de veiligheid van mensen of hun grondrechten mogelijk schaden (zoals beschermd door het Handvest van de grondrechten van de EU), worden als systemen met een hoog risico beschouwd. Bij de verordening horen lijsten van AI-systemen met een hoog risico, die kunnen worden herzien op basis van de ontwikkeling van AI-toepassingen in de praktijk.
- Het gaat ook om veiligheidscomponenten van producten die onder sectorale wetgeving van de Unie vallen. Zij worden per definitie als een hoog risico beschouwd als ze in het kader van de sectorale wetgeving een conformiteitsbeoordeling door derden moeten ondergaan.
- AI-systemen met een hoog risico zijn bijvoorbeeld AI-systemen die beoordelen of iemand in aanmerking komt voor een bepaalde medische behandeling, een bepaalde baan of een lening om een appartement te kopen. Een ander voorbeeld zijn AI-systemen die de politie gebruikt om mensen te profileren of hun risico op het plegen van een misdaad te beoordelen (tenzij dit verboden is op grond van artikel 5). Ook AI-systemen die robots, drones of medische hulpmiddelen bedienen, kunnen systemen met een hoog risico zijn.
- **Specifiek transparantierisico**: Transparantie rond het gebruik van AI kan het vertrouwen erin bevorderen. Voor specifieke AI-systemen legt de AI-verordening een aantal transparantieverplichtingen op, bijvoorbeeld als er een duidelijk risico op manipulatie bestaat (zoals door het gebruik van chatbots of deepfakes). Gebruikers moeten zich ervan bewust zijn dat ze met een machine in contact staan.
- **Minimaal risico**: De meeste AI-systemen kunnen zonder bijkomende wettelijke verplichtingen conform de bestaande wetgeving worden ontwikkeld en gebruikt. Aanbieders van die systemen kunnen ervoor kiezen de vereisten voor betrouwbare AI toe te passen en niet-bindende gedragscodes na te leven.

Daarnaast wordt in de AI-verordening rekening gehouden met **systeemrisico's** die kunnen voortvloeien uit **AI-modellen voor algemene doeleinden**, met inbegrip van **grote generatieve AI-modellen**. Deze modellen kunnen voor verschillende taken worden gebruikt en vormen intussen de basis voor veel AI-systemen in de EU. Sommige van deze modellen kunnen systeemrisico's met zich meebrengen als ze zeer krachtig zijn of op grote schaal worden gebruikt. Krachtige modellen kunnen bijvoorbeeld ernstige ongevallen veroorzaken of worden misbruikt voor verregaande cyberaanvallen. Als een model schadelijke vooroordelen in veel toepassingen verspreidt, kan dat gevolgen hebben voor heel veel mensen.

Hoe weet ik of een AI-systeem een hoog risico inhoudt?

De AI-verordening voorziet in een solide methode om AI-systemen aan te merken als systemen met een hoog risico. Dit moet bedrijven en andere marktdeelnemers rechtszekerheid bieden.

De risicoclassificatie is gebaseerd op het beoogde doel van het AI-systeem, conform de bestaande EU-wetgeving over productveiligheid. Dit betekent dat de classificatie afhangt van de functie die het AI-systeem vervult en van het specifieke doel en de specifieke modaliteiten waarvoor dat systeem wordt gebruikt.

AI-systemen kunnen in twee gevallen als systemen met een hoog risico worden beschouwd:

- als het AI-systeem is ingebed als veiligheidscomponent in producten die onder de bestaande productwetgeving vallen (bijlage I) of zelf dergelijke producten vormen. Op AI gebaseerde medische software is daar een voorbeeld van;
- als het AI-systeem bedoeld is om te worden gebruikt voor een gebruiksgeval met een hoog risico, opgenomen in bijlage III bij de AI-verordening. De lijst bevat gebruiksgevallen op gebieden als onderwijs, werkgelegenheid, rechtshandhaving of migratie.

De Commissie werkt momenteel aan richtsnoeren voor de indeling van systemen met een hoog risico, die vóór de datum van toepassing van deze regels zullen worden gepubliceerd.

Wat zijn voorbeelden van toepassingen met een hoog risico zoals gedefinieerd in bijlage III?

Bijlage III omvat acht gebieden waarop het gebruik van AI bijzonder gevoelig kan liggen en bevat

concrete gebruiksgevallen voor elk gebied. Een AI-systeem wordt ingedeeld als een AI-systeem met een hoog risico als het bedoeld is om voor een van deze gebruiksgevallen te worden gebruikt.

Voorbeelden zijn:

- AI-systemen die worden gebruikt als veiligheidscomponenten in bepaalde **kritieke infrastructuur**, bijvoorbeeld bij wegverkeer en bij de levering van water, gas, verwarming en elektriciteit;
- **AI-systemen die worden gebruikt in onderwijs en beroepsopleiding**, bijvoorbeeld om leerresultaten te evalueren, het leerproces te sturen en toezicht te houden op bedrog;
- **AI-systemen die worden gebruikt op het gebied van werkgelegenheid, personeelsbeheer** en toegang tot zelfstandige arbeid, bijvoorbeeld om gerichte vacatures te plaatsen, sollicitaties te analyseren en te filteren, en kandidaten te evalueren;
- **AI-systemen die worden gebruikt bij toegang tot essentiële particuliere en openbare diensten** en uitkeringen (bv. gezondheidszorg), **kredietwaardigheidsbeoordeling** van natuurlijke personen en risicobeoordeling en prijsstelling met betrekking tot **levens- en ziektekostenverzekeringen**;
- AI-systemen die worden gebruikt op het gebied van **rechtshandhaving**, migratie en **grenscontrole** – voor zover niet al verboden – en bij **rechtsbedeling en democratische processen**;
- AI-systemen die worden gebruikt voor **biometrische identificatie, biometrische categorisering en emotieherkenning**, voor zover niet verboden.

Aan welke verplichtingen moeten aanbieders van AI-systemen met een hoog risico voldoen?

Alvorens een **AI-systeem met een hoog risico in de EU in de handel te brengen** of anderszins in gebruik te nemen, moeten de aanbieders dat systeem aan een **conformiteitsbeoordeling** onderwerpen. Zo kunnen ze aantonen dat hun systeem voldoet aan de bindende eisen voor betrouwbare AI (bv. gegevenskwaliteit, documentatie en traceerbaarheid, transparantie, menselijk toezicht, nauwkeurigheid, cyberbeveiliging en robuustheid). Deze beoordeling moet worden herhaald als het systeem of het doel ervan ingrijpend wordt gewijzigd.

AI-systemen die fungeren als veiligheidscomponenten van producten die onder sectorale wetgeving van de Unie vallen, vormen een hoog risico als er op grond van de sectorale wetgeving voor de conformiteitsbeoordeling een beroep moet worden gedaan op een derde partij. Bovendien is voor alle biometrische systemen, ongeacht de toepassing ervan, een conformiteitsbeoordeling door derden verplicht.

Aanbieders van AI-systemen met een hoog risico moeten ook **kwaliteits- en risicobeheersystemen toepassen** om ervoor te zorgen dat ze aan de nieuwe eisen voldoen en om de risico's voor gebruikers en betrokkenen zoveel mogelijk te beperken, ook nadat een product in de handel is gebracht.

AI-systemen met een hoog risico die door overheidsinstanties of namens hen optredende entiteiten worden ingezet, moeten worden **geregistreerd in een openbare EU-databank**, tenzij die systemen worden gebruikt voor rechtshandhaving en migratie. Deze gegevens moeten worden geregistreerd in een niet-openbaar deel van de databank dat alleen toegankelijk is voor de betrokken toezichthoudende autoriteiten.

Om de naleving gedurende de hele levenscyclus van het AI-systeem te waarborgen, voeren de markttoezichtautoriteiten regelmatig audits uit, vergemakkelijken ze het toezicht na het in de handel brengen en geven ze aanbieders de kans vrijwillig ernstige incidenten of inbreuken te melden op de verplichtingen op het gebied van de grondrechten die ze opmerken. In uitzonderlijke gevallen kunnen de autoriteiten vrijstellingen verlenen voor specifieke AI-systemen met een hoog risico die in de handel worden gebracht.

In geval van een inbreuk kunnen de nationale autoriteiten op grond van de eisen toegang krijgen tot de informatie die ze nodig hebben om te onderzoeken of het gebruik van het AI-systeem in overeenstemming was met de wet.

Wat zou de rol van normalisatie in de AI-verordening zijn?

Op grond van de AI-verordening zijn AI-systemen met een hoog risico aan specifieke vereisten onderworpen. Europese geharmoniseerde normen zullen een cruciale rol spelen bij de toepassing van deze eisen.

In mei 2023 heeft de Europese Commissie de Europese normalisatieorganisaties CEN en CENELEC opdracht gegeven normen te ontwikkelen voor deze eisen voor systemen met een hoog risico. Deze opdracht wordt nu gewijzigd om in overeenstemming te zijn met de definitieve tekst van de AI-verordening.

De Europese normalisatieorganisaties hebben tot eind april 2025 de tijd om normen te ontwikkelen en te publiceren. Vervolgens zal de Commissie deze normen, die in het Publicatieblad van de EU worden bekendgemaakt, evalueren en eventueel goedkeuren. Zodra die normen zijn gepubliceerd, verlenen ze een "vermoeden van conformiteit" voor AI-systemen die in overeenstemming met die normen zijn ontwikkeld.

Hoe worden AI-modellen voor algemene doeleinden gereguleerd?

AI-modellen voor algemene doeleinden, waaronder **grote generatieve AI-modellen**, kunnen voor uiteenlopende taken worden gebruikt. Individuele modellen kunnen in een groot aantal AI-systemen worden geïntegreerd.

Het is belangrijk dat een aanbieder van een AI-systeem waarin een AI-model voor algemene doeleinden is geïntegreerd, over alle nodige informatie beschikt om ervoor te zorgen dat zijn systeem veilig is en in overeenstemming is met de AI-verordening.

Daarom verplicht de AI-verordening aanbieders van dergelijke modellen om **bepaalde informatie bekend te maken aan systeemaanbieders verder in de AI-waardeketen**. Deze **transparantie** maakt een beter begrip van deze modellen mogelijk.

Aanbieders van modellen moeten bovendien over beleid beschikken om ervoor te zorgen dat ze bij het trainen van hun modellen het **auteursrecht eerbiedigen**.

Daarnaast kunnen sommige van deze modellen **systeemrisico's** met zich meebrengen als ze zeer krachtig zijn of op grote schaal worden gebruikt.

Momenteel wordt van AI-modellen voor algemene doeleinden geoordeeld dat ze systeemrisico's inhouden als ze zijn getraind met een **totale rekenkracht van meer dan 10^{25} flops**. De Commissie kan deze drempel aanpassen of aanvullen in het licht van de technologische vooruitgang en kan ook van andere modellen stellen dat die systeemrisico's inhouden op basis van andere criteria (bv. aantal gebruikers of de mate van autonomie van het model).

Aanbieders van modellen met systeemrisico's worden verplicht om **risico's te beoordelen en te beperken, ernstige incidenten te melden, geavanceerde tests en modelevaluaties uit te voeren en de cyberbeveiliging** van hun modellen te waarborgen.

De aanbieders worden verzocht samen te werken met het AI-bureau en andere belanghebbenden om een praktijkcode te ontwikkelen waarin de regels worden beschreven en waarmee de veilige en verantwoorde ontwikkeling van hun modellen wordt gewaarborgd. Deze code moet een centraal instrument zijn voor aanbieders van AI-modellen voor algemene doeleinden om de naleving aan te tonen.

Waarom is 10^{25} flops een passende drempel voor AI-modellen voor algemene doeleinden met systeemrisico's?

Flop is een maatstaf voor modelcapaciteiten en de Commissie kan de exacte flop-drempel naar boven of naar beneden bijstellen, bijvoorbeeld in het licht van de vooruitgang bij het objectief meten van modelcapaciteiten, en van ontwikkelingen in de rekenkracht die nodig is voor een bepaald prestatieniveau.

Er is namelijk nog niet genoeg inzicht in de capaciteiten van de modellen boven deze drempel. Die kunnen systeemrisico's met zich meebrengen en daarom is het redelijk om hun aanbieders aan de aanvullende reeks verplichtingen te onderwerpen.

Welke verplichtingen bevat de AI-verordening met betrekking tot watermerken en etikettering van de AI-output?

De AI-verordening stelt transparantieregels vast voor inhoud die door generatieve AI wordt geproduceerd om het risico op manipulatie, misleiding en onjuiste informatie aan te pakken.

Zo worden aanbieders van generatieve AI-systemen verplicht de AI-output in een machineleesbaar formaat te markeren en ervoor te zorgen dat die output detecteerbaar is als kunstmatig gegenereerd of gemanipuleerd. De technische oplossingen moeten doeltreffend, interoperabel, robuust en betrouwbaar zijn voor zover dat technisch haalbaar is. Daarbij moet rekening worden gehouden met de specifieke kenmerken en beperkingen van de verschillende soorten content, de uitvoeringskosten

en de algemeen erkende stand van de techniek, zoals tot uiting kan komen in relevante technische normen.

Gebruiksverantwoordelijken van een generatief AI-systeem dat beeld-, audio- of videocontent genereert of bewerkt die deepfakes vormen, moeten bovendien zichtbaar aangeven dat de content kunstmatig is gegenereerd of gemanipuleerd. Gebruiksverantwoordelijken van een AI-systeem dat tekst genereert of bewerkt die wordt gepubliceerd om het publiek te informeren over aangelegenheden van algemeen belang, moeten ook aangeven dat de tekst kunstmatig is gegenereerd of bewerkt. Deze verplichting is niet van toepassing als de door AI gegenereerde content een proces van menselijke toetsing of redactionele controle heeft ondergaan en als een natuurlijke of rechtspersoon redactionele verantwoordelijkheid draagt voor de bekendmaking van de content.

Het AI-bureau zal richtsnoeren uitvaardigen om aanbieders en gebruiksverantwoordelijken verdere richtsnoeren te verstrekken over de verplichtingen in artikel 50, die twee jaar na de inwerkingtreding van de AI-verordening, op 2 augustus 2026, van toepassing zullen worden.

Het AI-bureau stimuleert en vergemakkelijkt de opstelling van praktijkcodes op het niveau van de Unie om de doeltreffende uitvoering van de verplichtingen met betrekking tot het opsporen en het labelen van kunstmatig gegenereerde of bewerkte content te vergemakkelijken.

Is de AI-verordening toekomstbestendig?

De AI-verordening stelt een rechtskader vast dat inspeelt op nieuwe ontwikkelingen, gemakkelijk en snel kan worden aangepast en frequente evaluaties mogelijk maakt.

De AI-verordening stelt resultaatgerichte eisen en plichten vast, maar laat de concrete technische oplossingen en operationalisering over aan door de industrie gestuurde normen en praktijkcodes die flexibel zijn en kunnen worden aangepast aan verschillende gebruikssituaties en nieuwe technologische oplossingen mogelijk kunnen maken.

Daarnaast kan de wetgeving zelf worden gewijzigd door middel van gedelegeerde en uitvoeringshandelingen, bijvoorbeeld om de lijst van gevallen van gebruik met een hoog risico in bijlage III te herzien.

Tot slot zullen bepaalde delen van de AI-verordening en uiteindelijk de gehele verordening regelmatig worden geëvalueerd, zodat kan worden vastgesteld of er behoefte is aan herziening en wijziging.

Hoe reguleert de AI-wet biometrische identificatie?

Het gebruik van **biometrische identificatie op afstand in realtime in openbare ruimten** (d.w.z. gezichtsherkenning met behulp van CCTV) voor rechtshandhavingsdoeleinden is verboden. De lidstaten kunnen bij wet uitzonderingen invoeren die het gebruik van biometrische identificatie op afstand in realtime toelaten in de volgende gevallen:

- rechtshandhavingsactiviteiten in verband met 16 specifieke zeer ernstige misdaden;
- gerichte opsporing van specifieke slachtoffers, ontvoering, mensenhandel en seksuele uitbuiting van mensen, en vermiste personen; of
- het voorkomen van bedreigingen voor het leven of de fysieke veiligheid van personen of het reageren op de huidige of voorzienbare dreiging van een terreuraanslag.

Uitzonderlijk gebruik is onderworpen aan **voorafgaande toestemming van een gerechtelijke of onafhankelijke administratieve autoriteit** waarvan de beslissing bindend is. In spoedeisende gevallen kan de goedkeuring binnen 24 uur worden verleend; als de vergunning wordt geweigerd, moeten alle gegevens en de output worden gewist.

De vergunning moet worden voorafgegaan door een **effectbeoordeling voor de grondrechten** en moet worden **gemeld aan de betrokken markttoezichtautoriteit en de gegevensbeschermingsautoriteit**. In spoedeisende gevallen kan met het gebruik van het systeem worden begonnen zonder registratie.

Voor het gebruik van AI-systemen voor **biometrische identificatie op afstand achteraf** (identificatie van personen in eerder verzameld videomateriaal) van personen tegen wie een onderzoek loopt, is **voorafgaande toestemming** vereist van een gerechtelijke of onafhankelijke administratieve autoriteit, evenals melding aan de gegevensbeschermings- en markttoezichtautoriteit.

Waarom zijn er specifieke regels nodig voor biometrische identificatie op

afstand?

Biometrische identificatie kan verschillende vormen aannemen. Biometrische authenticatie en verificatie, d.w.z. om een smartphone te ontgrendelen of voor verificatie/authenticatie aan grensovergangen om te controleren of de identiteit van een persoon overeenkomt met die op de reisdocumenten (een-op-eenmatching), blijven ongereguleerd, omdat die geen significant risico vormen voor de grondrechten.

Biometrische identificatie kan daarentegen ook op afstand worden gebruikt, bijvoorbeeld om mensen in een menigte te identificeren. Dat kan aanzienlijke gevolgen hebben voor de privacy op openbare plaatsen.

De nauwkeurigheid van systemen voor gezichtsherkenning hangt sterk af van talrijke factoren, zoals de kwaliteit van de camera, het licht, de afstand, de databank, het algoritme en de etniciteit, de leeftijd of het geslacht van de persoon. Hetzelfde geldt voor de herkenning van de manier van stappen, spraakherkenning en andere biometrische systemen. Bij zeer geavanceerde systemen ligt het percentage valse positieven steeds lager.

Hoewel een nauwkeurigheidsgraad van 99% over het algemeen goed lijkt, is het bijzonder riskant als een onschuldig persoon verdacht zou worden. Zelfs een foutenpercentage van 0,1% kan een aanzienlijk effect hebben bij grote populaties, bijvoorbeeld in treinstations.

Hoe beschermen deze nieuwe regels de grondrechten?

In de EU en haar lidstaten bestaat al een sterke bescherming van de grondrechten en tegen discriminatie, maar de complexiteit en ondoorzichtigheid van bepaalde AI-toepassingen (“black boxes”) kunnen een probleem vormen.

Een mensgerichte benadering van AI betekent dat AI-toepassingen in overeenstemming moeten zijn met de wetgeving inzake grondrechten. Door verantwoordings- en transparantievereisten te integreren in de ontwikkeling van AI-systemen met een hoog risico en door de handhavingscapaciteit te verbeteren, kunnen we ervoor zorgen dat deze systemen van meet af aan met inachtneming van de wettelijke voorschriften worden ontworpen. Als een inbreuk wordt vastgesteld, kunnen de nationale autoriteiten op grond van die eisen toegang krijgen tot de informatie die ze nodig hebben om te onderzoeken of het gebruik van AI in overeenstemming was met de EU-regelgeving.

Bovendien vereist de AI-verordening dat bepaalde gebruiksverantwoordelijken van AI-systemen met een hoog risico een effectbeoordeling voor de grondrechten uitvoeren.

Wat is een effectbeoordeling voor de grondrechten? Wie moet die beoordeling uitvoeren, en wanneer?

Aanbieders van AI-systemen met een hoog risico moeten een risicobeoordeling uitvoeren en het systeem zodanig ontwerpen dat de risico's voor de gezondheid, de veiligheid en de grondrechten tot een minimum worden beperkt.

Bepaalde risico's voor de grondrechten kunnen echter alleen ten volle worden vastgesteld als duidelijk is in welke context het AI-systeem met een hoog risico wordt gebruikt. Als AI-systemen met een hoog risico worden gebruikt op bijzonder gevoelige gebieden met mogelijke machtsongelijkheid, zijn aanvullende overwegingen met betrekking tot dergelijke risico's noodzakelijk.

Gebruiksverantwoordelijken die publiekrechtelijke lichamen zijn of particuliere exploitanten die openbare diensten verlenen, en gebruiksverantwoordelijken die AI-systemen met een hoog risico aanbieden die kredietwaardigheidsbeoordelingen of prijsstelling of risicobeoordelingen in levens- en gezondheidsverzekeringen uitvoeren, moeten daarom een beoordeling uitvoeren van de gevolgen voor de grondrechten en de nationale autoriteit in kennis stellen van de resultaten.

In de praktijk zullen veel gebruiksverantwoordelijken ook een gegevensbeschermingseffectbeoordeling moeten uitvoeren. Om substantiële overlappingen in dergelijke gevallen te voorkomen, wordt de effectbeoordeling voor de grondrechten uitgevoerd in samenhang met die gegevensbeschermingseffectbeoordeling.

Hoe biedt deze verordening een antwoord op de vooroordelen op het gebied van ras en geslacht in AI?

Het is zeer belangrijk om te benadrukken dat AI-systemen **geen vooroordelen creëren of reproduceren**. Als AI-systemen correct worden ontworpen en gebruikt, kunnen die net **bijdragen tot het verminderen van vooroordelen en bestaande structurele discriminatie**, en zo tot

eerlijkere en niet-discriminerende beslissingen leiden (bv. bij een aanwerving).

Dat is de **doelstelling van de nieuwe bindende eisen voor alle AI-systemen met een hoog risico**. AI-systemen moeten **technisch robuust** zijn om ervoor te zorgen dat ze geschikt zijn voor het beoogde doel en geen vertekende resultaten opleveren, zoals valse positieve of negatieve resultaten, die gemarginaliseerde groepen onevenredig treffen, onder meer op basis van ras of etnische afkomst, geslacht, leeftijd en andere beschermde kenmerken.

Systemen met een hoog risico moeten ook worden **getraind en getest met voldoende representatieve datasets** om het **risico op oneerlijke vooroordelen in het model tot een minimum te beperken** en ervoor te zorgen dat er passende maatregelen zijn om die vooroordelen op te sporen en te corrigeren en de risico's ervan te beperken.

Ze moeten ook **traceerbaar en controleerbaar** zijn en passende **documentatie bijhouden**, waaronder de gegevens die worden gebruikt om het algoritme te trainen, die cruciaal kunnen zijn voor onderzoeken achteraf.

Het **systeem voor de bewaking van de conformiteit vóór en na het in de handel brengen**, moet ervoor zorgen dat de systemen **regelmatig worden gecontroleerd** en dat **potentiële risico's onmiddellijk worden aangepakt**.

Wanneer is de AI-verordening volledig van toepassing?

De AI-verordening wordt twee jaar na de inwerkingtreding van toepassing, op 2 augustus 2026, met uitzondering van de volgende specifieke bepalingen:

- de verbodsbepalingen, definities en bepalingen met betrekking tot AI-geletterdheid worden zes maanden na de inwerkingtreding van toepassing, op 2 februari 2025;
- de regels inzake governance en de verplichtingen voor AI voor algemene doeleinden worden een jaar na de inwerkingtreding van toepassing, op 2 augustus 2025;
- de verplichtingen voor AI-systemen met een hoog risico die als systemen met een hoog risico worden geclassificeerd omdat zij zijn ingebed in gereglementeerde producten die zijn opgenomen in bijlage II (lijst van harmonisatiewetgeving van de Unie), worden drie jaar na de inwerkingtreding van toepassing, op 2 augustus 2027.

Hoe zal de AI-wet worden gehandhaafd?

De AI-verordening voorziet in een tweeledig governancesysteem, waarbij **nationale autoriteiten** verantwoordelijk zijn voor het toezicht op en de handhaving van de regels voor AI-systemen, terwijl de EU verantwoordelijk is voor het beheer van AI-modellen voor algemene doeleinden.

Om de EU-brede samenhang en samenwerking te waarborgen, zal een **Europese raad voor artificiële intelligentie** (de AI-board) worden opgericht, bestaande uit vertegenwoordigers van de lidstaten, met gespecialiseerde subgroepen voor nationale regelgevende instanties en andere bevoegde autoriteiten.

Het **AI-bureau**, het uitvoeringsorgaan van de Commissie voor de AI-verordening, zal de AI-board strategische richtsnoeren verstrekken.

Daarnaast worden bij de AI-verordening twee adviesorganen opgericht om deskundige input te leveren: het **wetenschappelijk panel** en het **adviesforum**. Deze organen zullen zorgen voor waardevolle inzichten van belanghebbenden en interdisciplinaire wetenschappelijke gemeenschappen, die kunnen dienen als input voor de besluitvorming en zorgen voor een evenwichtige aanpak van de ontwikkeling van AI.

Waarom is een AI-board nodig en wat gaat die doen?

De AI-board bestaat uit **hooggeplaatste vertegenwoordigers van de lidstaten** en de Europese Toezichthouder voor gegevensbescherming (EDPS). Als belangrijke adviseur verstrekt de AI-board richtsnoeren over alle kwesties die verband houden met AI-beleid, met name AI-regelgeving, innovatie- en excellentiebeleid en internationale samenwerking op het gebied van AI.

De AI-board speelt een cruciale rol bij het waarborgen van de soepele, doeltreffende en geharmoniseerde uitvoering van de AI-verordening. De AI-board zal fungeren als forum waar de AI-regelgevers, namelijk het AI-bureau, de nationale autoriteiten en de EPDS, de consistente toepassing van de AI-verordening kunnen coördineren.

Hoe worden inbreuken bestraft?

De lidstaten moeten doeltreffende, evenredige en afschrikwekkende sancties vaststellen

voor inbreuken op de regels voor AI-systemen.

In de verordening worden de volgende drempels vastgesteld:

- **tot €35 miljoen of 7%** van de totale wereldwijde jaaromzet in het voorgaande boekjaar (als dat bedrag hoger is) in geval van **verboden praktijken of inbreuken** op de wettelijke eisen inzake gegevens;
- **tot €15 miljoen of 3%** van de totale wereldwijde jaaromzet in het voorgaande boekjaar voor **inbreuken op andere vereisten** of verplichtingen van de verordening;
- **tot €7,5 miljoen of 1,5%** van de totale wereldwijde jaaromzet in het voorgaande boekjaar voor het **verstrekken van onjuiste, onvolledige of misleidende informatie** aan aangemelde instanties en nationale bevoegde autoriteiten in antwoord op een verzoek.
- Voor elke categorie inbreuken is de drempel voor kmo's het laagste van de twee bedragen en voor andere ondernemingen het hoogste.

De Commissie kan de regels voor aanbieders van AI-modellen voor algemene doeleinden ook handhaven door middel van boetes, rekening houdend met de volgende drempel:

- **tot €15 miljoen of 3%** van de totale wereldwijde jaaromzet in het voorgaande boekjaar voor **inbreuken op verplichtingen** of maatregelen die de Commissie vereist in het kader van de verordening.

Van de EU-instellingen, -agentschappen of -organen wordt verwacht dat ze het goede voorbeeld geven. Daarom zijn ook zij aan de regels en mogelijke sancties onderworpen. De Europese Toezichthouder voor gegevensbescherming zal de bevoegdheid hebben om hen boetes op te leggen in geval van niet-naleving.

Hoe zal de AI-praktijkcode voor algemene doeleinden tot stand komen?

De eerste code wordt opgesteld na een inclusief en transparant proces. Er wordt een plenaire vergadering over de praktijkcode samengesteld om het iteratieve ontwerpproces te vergemakkelijken. Die vergadering bestaat uit alle geïnteresseerde en in aanmerking komende aanbieders van AI-modellen voor algemene doeleinden, downstreamaanbieders die een AI-model voor algemene doeleinden in hun AI-systeem integreren, andere brancheorganisaties, andere organisaties van belanghebbenden zoals maatschappelijke organisaties of organisaties van rechthebbenden, alsook de academische wereld en andere onafhankelijke deskundigen.

Het AI-bureau heeft een oproep tot het indienen van blijken van belangstelling gelanceerd om mee te werken aan het opstellen van de eerste praktijkcode. Parallel aan deze oproep tot het indienen van blijken van belangstelling wordt een raadpleging van meerdere belanghebbenden gehouden om standpunten en input van alle belanghebbenden over de eerste praktijkcode te verzamelen. De antwoorden en bijdragen zullen de basis vormen voor de eerste versie van de praktijkcode. Van meet af aan is de code dus gebaseerd op een breed scala aan perspectieven en deskundigheid.

De plenaire vergadering wordt verdeeld in vier werkgroepen om gerichte besprekingen mogelijk te maken over specifieke onderwerpen die relevant zijn voor de nadere invulling van de verplichtingen voor aanbieders van AI-modellen voor algemene doeleinden en AI-modellen voor algemene doeleinden met een systeemrisico. De deelnemers aan de plenaire vergadering kunnen zelf een of meer werkgroepen kiezen. De vergaderingen vinden uitsluitend online plaats.

Het AI-bureau benoemt voorzitters en waar nodig vicevoorzitters voor elk van de vier werkgroepen van de plenaire vergadering, die worden geselecteerd uit geïnteresseerde onafhankelijke deskundigen. De voorzitters zullen de bijdragen en opmerkingen van deelnemers aan de plenaire vergadering samenvatten om de eerste praktijkcode op te stellen.

De code is in de eerste plaats gericht aan de aanbieders van AI-modellen voor algemene doeleinden; zij worden daarom – naast hun deelname aan de plenaire vergadering – ook uitgenodigd voor specifieke workshops om voor elke ontwerpronde input te leveren.

Na negen maanden zal de definitieve versie van de eerste praktijkcode in een slotzitting, die naar verwachting in april 2025 zal plaatsvinden, worden gepresenteerd en gepubliceerd. De slotzitting biedt aanbieders van AI-modellen voor algemene doeleinden de mogelijkheid om zich uit te spreken of ze voornemens zijn de code te gebruiken.

Hoe wordt de praktijkcode voor aanbieders van AI-modellen voor algemene doeleinden, als ze wordt goedgekeurd, een centraal instrument voor naleving?

Aan het einde van het ontwerpproces van de praktijkcode beoordelen het AI-bureau en de AI-board de adequaatheid van de code en maken ze hun beoordeling bekend. Na die beoordeling kan de Commissie besluiten de praktijkcode goed te keuren en deze via uitvoeringshandelingen binnen de Unie algemeen geldig te maken. Als het AI-bureau op het tijdstip waarop de verordening van toepassing wordt, de praktijkcode niet toereikend acht, kan de Commissie gemeenschappelijke regels vaststellen voor de uitvoering van de relevante verplichtingen.

Aanbieders van AI-modellen voor algemene doeleinden kunnen zich baseren op de praktijkcode om aan te tonen dat ze voldoen aan de verplichtingen van de AI-verordening.

Overeenkomstig de AI-verordening moet de praktijkcode doelstellingen, maatregelen en waar nodig kernprestatie-indicatoren (KPI's) bevatten.

Aanbieders die zich aan de code houden, moeten regelmatig verslag uitbrengen aan het AI-bureau over de uitvoering van de genomen maatregelen en de resultaten daarvan, in voorkomend geval ook gemeten aan de hand van KPI's.

Dit vergemakkelijkt de handhaving door het AI-bureau, dat wordt gesteund door de bevoegdheden die door de AI-verordening aan de Commissie zijn verleend. Dit omvat de mogelijkheid om evaluaties van AI-modellen voor algemene doeleinden uit te voeren, modelaanbieders om informatie en maatregelen te verzoeken en sancties op te leggen.

Het AI-bureau stimuleert en faciliteert in voorkomend geval ook de evaluatie en aanpassing van de code aan de technologische vooruitgang.

Zodra een geharmoniseerde norm gepubliceerd is en het AI-bureau die in het licht van de relevante verplichtingen geschikt heeft bevonden, moet naleving van een Europese geharmoniseerde norm aanbieders het vermoeden van conformiteit bieden.

Aanbieders van AI-modellen voor algemene doeleinden moeten ook met alternatieve geschikte middelen kunnen aantonen dat ze aan de relevante verplichtingen voldoen als er geen praktijkcodes of geharmoniseerde normen beschikbaar zijn, of als ze ervoor kiezen geen beroep op die codes of normen te doen.

Bevat de AI-verordening bepalingen inzake milieubescherming en duurzaamheid?

De AI-verordening heeft tot doel risico's voor de veiligheid en de grondrechten aan te pakken, zoals het grondrecht op een hoog niveau van milieubescherming. Het milieu is ook een van de uitdrukkelijk genoemde en beschermde rechtsbelangen.

De Commissie wordt verzocht Europese normalisatieorganisaties om een normalisatieproduct te vragen over rapportage- en documentatieprocessen om de hulpbronnenprestaties van AI-systemen te verbeteren, zoals de vermindering van het energieverbruik en het verbruik van andere hulpbronnen van het AI-systeem met een hoog risico tijdens zijn levenscyclus, en over de energie-efficiënte ontwikkeling van AI-modellen voor algemene doeleinden.

Voorts wordt de Commissie verzocht uiterlijk twee jaar na de toepassing van de verordening en vervolgens om de vier jaar een verslag in te dienen over de evaluatie van de vooruitgang bij de ontwikkeling van normalisatieproducten met betrekking tot de energie-efficiënte ontwikkeling van modellen voor algemeen gebruik, en na te gaan of er behoefte is aan verdere maatregelen of acties, met inbegrip van bindende maatregelen of acties.

Daarnaast zijn aanbieders van AI-modellen voor algemene doeleinden, die worden getraind met grote hoeveelheden gegevens en daardoor gevoelig zijn voor een hoog energieverbruik, verplicht het energieverbruik bekend te maken. In het geval van AI-modellen voor algemene doeleinden met systeemrisico's moet bovendien de energie-efficiëntie worden beoordeeld.

De Commissie is bevoegd om een passende en vergelijkbare meetmethode voor deze openbaarmakingsverplichtingen te ontwikkelen.

Hoe kunnen de nieuwe regels innovatie ondersteunen?

Het regelgevingskader kan het gebruik van AI op twee manieren bevorderen. Enerzijds zal een groter vertrouwen onder de gebruikers de vraag naar AI die door bedrijven en overheidsinstanties wordt gebruikt, doen toenemen. Anderzijds zullen AI-aanbieders, dankzij de grotere rechtszekerheid en harmonisering van de regels, toegang krijgen tot grotere markten, met producten die gebruikers en consumenten zullen waarderen en kopen. De regels zullen alleen van toepassing zijn als dat strikt noodzakelijk is en zodanig dat de lasten voor de marktdeelnemers tot een minimum worden beperkt, met een lichte governancestructuur.

De AI-verordening maakt verder de opzet mogelijk van **testomgevingen voor regelgeving** en **tests onder reële omstandigheden**, die een gecontroleerde omgeving bieden om gedurende een beperkte tijd innovatieve technologieën te testen, waardoor innovatie door bedrijven, kmo's en start-ups wordt bevorderd in overeenstemming met de AI-verordening. Samen met andere maatregelen, zoals de aanvullende **netwerken van AI-kenniscentra** en het **publiek-privaat partnerschap voor AI, data en robotica,** en toegang tot **digitale-innovatiehubs** en **test- en experimenteerfaciliteiten** zullen zij helpen om voor bedrijven de gunstige randvoorwaarden te scheppen om AI te ontwikkelen en in te zetten.

Tests onder reële omstandigheden van AI-systemen met een hoog risico kunnen worden uitgevoerd voor maximaal zes maanden (een termijn die met nog eens zes maanden kan worden verlengd). Voorafgaand aan de tests moet een plan worden opgesteld en ingediend bij de markttoezichtautoriteit, die het plan en de specifieke testvoorwaarden moet goedkeuren, met standaard stilzwijgende goedkeuring indien binnen 30 dagen geen antwoord is gegeven. Tests kunnen onaangekondigd door de autoriteit worden geïnspecteerd.

Tests onder reële omstandigheden kunnen alleen worden uitgevoerd met specifieke waarborgen. Gebruikers van de systemen die onder reële omstandigheden worden getest, moeten bijvoorbeeld geïnformeerde toestemming geven, de tests mogen geen negatieve gevolgen voor hen hebben, de resultaten moeten omkeerbaar zijn of buiten beschouwing kunnen worden gelaten, en hun gegevens moeten na afloop van de test worden gewist. Bijzondere bescherming moet worden geboden aan kwetsbare groepen, met name vanwege hun leeftijd of lichamelijke of geestelijke handicap.

Welke rol speelt het AI-pact bij de uitvoering van de AI-verordening?

Het AI-pact, dat in mei 2023 door commissaris Breton is gestart, heeft tot doel de betrokkenheid van het AI-bureau en organisaties te vergroten (pijler I) en de vrijwillige toezegging van de sector aan te moedigen om vóór de wettelijke termijn te beginnen met de uitvoering van de vereisten van de AI-verordening (pijler II).

In het kader van pijler I dragen de deelnemers met name bij tot de oprichting van een samenwerkingsgemeenschap en delen ze hun ervaringen en kennis. Dit omvat workshops die door het AI-bureau worden georganiseerd om de deelnemers een beter inzicht te geven in de AI-verordening, hun verantwoordelijkheden en de wijze waarop ze zich kunnen voorbereiden op de uitvoering ervan. Op zijn beurt kan het AI-bureau inzichten verzamelen in goede praktijken en uitdagingen waarmee de deelnemers worden geconfronteerd.

In het kader van pijler II worden organisaties aangemoedigd om proactief bekend te maken welke processen en praktijken ze toepassen om op naleving te anticiperen, door middel van vrijwillige toezeggingen. Toezeggingen zijn bedoeld als "verbintenisverklaringen" en bevatten (geplande of lopende) acties om aan een aantal van de vereisten van de AI-verordening te voldoen.

De meeste regels van de AI-verordening (bijvoorbeeld sommige voorschriften voor AI-systemen met een hoog risico) worden van toepassing aan het einde van een overgangsperiode (d.w.z. de tijd tussen de inwerkingtreding en de datum van toepassing).

In dit verband en in het kader van het AI-pact roept het AI-bureau alle organisaties op alvast vooruit te lopen op de AI-verordening en een aantal van de belangrijkste bepalingen uit te voeren, om zo de risico's voor de gezondheid, de veiligheid en de grondrechten zo snel mogelijk te beperken.

Naar aanleiding van een oproep in november 2023 hebben meer dan 700 organisaties al belangstelling getoond om toe te treden tot het AI-pact. Op 6 mei vond een eerste online informatiesessie plaats, met 300 deelnemers. De officiële ondertekening van de vrijwillige verbintenissen is gepland voor het najaar van 2024. In de eerste week van september vindt er een workshop over het AI-pact plaats.

Wat is de internationale dimensie van de aanpak van de EU?

De gevolgen en uitdagingen van AI stoppen niet bij grenzen; daarom is internationale samenwerking belangrijk. Het AI-bureau is verantwoordelijk voor de internationale betrokkenheid van de Europese Unie op het gebied van AI, op basis van de AI-verordening en het gecoördineerde plan inzake AI. De EU streeft ernaar het verantwoord beheer en goed bestuur van AI te bevorderen in samenwerking met internationale partners en in overeenstemming met het op regels gebaseerde multilaterale systeem en de waarden die het in stand houdt.

De EU zet zich bilateraal en multilateraal in om betrouwbare, mensgerichte en ethische AI te bevorderen. Daarom is de EU betrokken bij multilaterale fora waar AI wordt besproken – met name de G7, de G20, de OESO, de Raad van Europa, het mondiaal partnerschap inzake AI en de Verenigde Naties – en onderhoudt de EU nauwe bilaterale betrekkingen met bijvoorbeeld Canada, de VS, India,

Japan, Zuid-Korea, Singapore en de Latijns-Amerikaanse en Caribische regio.

**Bijgewerkt op 1.8.2024*

QANDA/21/1683

Contactpersoon voor de pers:

Thomas Regnier (+32 2 29 9 1099)
Patricia Poropat (+32 2 298 04 85)

Voor het publiek: Europe Direct per telefoon 00 800 67 89 10 11 of e-mail