

Ook in het arbeidsrecht krijgen we met de AI-verordening te maken

ArbeidsRecht 2025/28

AI is er nu ook op de werkvloer. Het wordt breed ingezet (sollicitaties, belonen, beoordelen, ontslag, etc.). Deze innovatie levert, naast kansen, ook risico's op. Daarom is sinds kort de AI-verordening van kracht. Die deelt AI-systemen in op basis van risicocategorieën: minimaal, beperkt, hoog of onaanvaardbaar. Bij een hoger risico horen strengere regels. Nu veel rondom 'werk' is aangemerkt als 'hoog risico', krijgen werkgevers die AI-systemen inzetten dus met allerlei regels te maken. En werknemers (en hun vertegenwoordigers) krijgen nieuwe subjectieve rechten. Boetes kunnen oplopen tot 3% van de wereldwijde omzet. Genoeg reden dus om kennis te maken met de AI-verordening.

1. Hoe werken AI-systemen?

Als het over artificiële intelligentie (AI) gaat, gaat het vaak over *algoritmes*. Een algoritme is een reeks instructies of regels die een computer volgt om een specifieke taak uit te voeren. Hoewel er vele soorten algoritmes zijn, gaat het – als het over AI gaat – het vaakst over zogeheten *machine learning*-systemen. Dergelijke systemen worden 'getraind' met grote hoeveelheden door de mens ingevoerde data (inputdata) om voorspellingen te doen, bijvoorbeeld om beslissingen voor te bereiden of te ondersteunen.² Een bekend AI-systeem is ChatGPT, een *large language model* dat is ontworpen om taal te analyseren en teksten te genereren. Andere bekende voorbeelden zijn de sollicitatie-robot van Amazon (die vrouwen bleek te discrimineren) of het antifraudesysteem SyRI van de overheid (waarover de rechter oordeelde dat dit inbreuk maakte op de privacy van burgers).³ Hierbij merken wij gelijk op dat er – naast deze voorbeelden die de media hebben gehaald – ook vele AI-systemen zijn die probleemloos en naar ieders tevredenheid in de praktijk worden toegepast.

Omdat er allerlei (technisch onderling verschillende) soorten AI-systemen zijn, heeft de Europese Commissie met de

AI-verordening⁴ – net als met de Algemene verordening gegevensbescherming (de AVG) – gekozen voor een zogeheten 'technologisch neutrale benadering'. In artikel 3 lid 1 AI-verordening wordt de term 'AI-systeem' dus als volgt gedefinieerd:

"AI-systeem: 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;"

2. De Europese AI-verordening

De AI-verordening beoogt de interne markt te verbeteren, mensgerichte en betrouwbare AI te bevorderen en een hoog niveau van bescherming van gezondheid, veiligheid en fundamentele rechten te waarborgen. De AI-verordening streeft naar een evenwicht tussen innovatie en de bescherming van fundamentele rechten, zoals privacy, gelijke behandeling, kinderrechten, consumentenrechten en vrijheid van meningsuiting. De Europese Unie wil hiermee optreden als *Global Regulatory Power* en het voortouw nemen in de ontwikkeling van normen op het gebied van transparantie, risicomanagement en de bescherming van fundamentele rechten. Omdat het gaat om een verordening, werken de daarin vervatte regels rechtstreeks door in de nationale rechtsordes van de lidstaten.

De AI-verordening zal gefaseerd in werking treden. De eerste tranche regels is per 1 augustus 2024 van kracht geworden, maar het zal tot 31 december 2030 duren voordat de AI-verordening volledig werkt. Dit is afhankelijk van de aard van de verplichtingen uit de verordening en het risico van het AI-systeem.⁵ Zo geldt het verbod op systemen met een onaanvaardbaar risico van artikel 5 AI-verordening al sinds 2 februari 2025. De voor het arbeidsrecht meest relevante categorie regels, die over hoogrisicosystemen, zal pas op 2 augustus 2027 gaan gelden.⁶

1 Mrs. M. (Marko) Jovović, J.C. (Jordy) Gelmers en J.A. (Janita) Hofman zijn advocaat en werken bij Stibbe NV in Amsterdam.

2 Zie voor een meer uitgebreide introductie op het onderwerp M. Jovović & W. Roozendaal, 'Algoritmes en personeelselectie', in: Houwerzijl, Montebovi & Zekić (red.), *Platformisering, algoritmisering en sociale bescherming* (MSR nr. 78), Deventer: Wolters Kluwer 2021/8.

3 Rb. Den Haag 5 februari 2020, ECLI:NL:RBDHA:2020:865, JB 2020/63, m.nt. R.H.T. Jansen & M.D. Reijneveld, TRA 2020/49, m.nt. W.L. Roozendaal, AB 2020/236, m.nt. H.B. van Kolfschooten, O&A 2020/28, m.nt. T. Barkhuysen, N. Jak & S.R.P. Bastiaans, NJ 2020/386, m.nt. E.J. Dommering en JBP 2023/197, m.nt. E.N. Üstünlalp, D.J.B. Hereijgers & A.M. Arnbak.

4 Verordening (EU) 2024/1689 van het Europees Parlement en de Raad van 13 juni 2024 tot vaststelling van geharmoniseerde regels betreffende artificiële intelligentie en tot wijziging van de Verordeningen (EG) 300/2008, (EU) 167/2013, (EU) 168/2013, (EU) 2018/858, (EU) 2018/1139 en (EU) 2019/2144 en de Richtlijnen 2014/90/EU, (EU) 2016/797 en (EU) 2020/1828 (verordening artificiële intelligentie), PbEU 2024, L 1689.

5 Art. 113 AI-verordening.

6 Zie echter art. 111 AI-verordening voor systemen die al in de handel zijn.

3. Risicogebaseerde aanpak

Kort gezegd houdt de risicogebaseerde aanpak van de AI-verordening in: hoe hoger het risico, des te strenger de regels.

3.1 Relevante risico's en het karakter van de AI-verordening

De risicoclassificatie moet worden gezien tegen de achtergrond van de doelen van de AI-verordening. Die dient (mede) ter waarborging van een hoog niveau van bescherming van de gezondheid, de veiligheid en de in het Handvest verankerde grondrechten. Deze grondrechtengerichte risicobenadering past bij de overige wetgeving van de Uniewetgever: ook de AVG hanteert deze benadering.

Met een risicogebaseerde benadering kan worden beoordeeld of een grondrecht moet worden gewaarborgd en, zo ja, in hoeverre en hoe dit dient te gebeuren. Grondrechten kunnen immers met elkaar botsen. In de context van AI zou het bijvoorbeeld nodig kunnen zijn om bijzondere persoonsgegevens (bijvoorbeeld over gender, etniciteit of religie) te verwerken om een systeem minder discriminatoir te maken – privacy versus gelijke behandeling dus.⁷ Omdat niet alle grondrechten absoluut zijn, is een botsing van grondrechten niet per se problematisch. Niet-absolute grondrechten kunnen immers tegen elkaar worden afgewogen met behulp van de evenredigheidstoets.⁸

De evenredigheidstoets komt deels neer op een belangenafweging. De risicobenadering uit de AI-verordening is een manier om deze belangenafweging te concretiseren. De Uniewetgever biedt hiervoor de kaders door duidelijk te maken dat, hoe hoger de risico's voor bepaalde rechten en vrijheden van burgers zijn, des te meer verplichtingen gelden. Dit is niet nieuw. Ook de risicogebaseerde benadering van de AVG werkt met een afwegingskader van belangen. Waar de AVG de verwerkingsverantwoordelijke en de verwerker echter relatief vrijlaat in het afwegen van de risico's, biedt de AI-verordening met de risicocategorieën meer houvast.

In de AI-verordening heeft de Uniewetgever de afwegingen die moeten worden gemaakt bij een risicogebaseerde benadering dus in verdergaande mate ingevuld dan met de AVG. Dit maakt de AI-verordening mogelijk iets minder flexibel, maar waarschijnlijk ook gemakkelijker toepasbaar. Daarmee kan de AI-verordening bijdragen aan de bescherming van grondrechten.

Het bovenstaande verklaart de verschillende risicocategorieën: hoe meer risico's voor rechten en vrijheden, des te meer waarborgen moeten worden gesteld in het kader van de evenredigheidstoets.

⁷ De AI-verordening lijkt dit onder strikte voorwaarden mogelijk te maken, zie art. 10 lid 5 AI-verordening.

⁸ Deze evenredigheidstoets ligt bijv. ten grondslag aan art. 52 Handvest.

Wij laten de categorieën 'beperkt risico' en 'minimaal risico' in deze bijdrage voor wat ze zijn, omdat veel systemen die in werksituaties worden gebruikt nu eenmaal als 'hoog risico' zullen worden gekwalificeerd. In de categorie 'beperkt risico' vallen AI-systemen die bedoeld zijn om met personen in contact te komen, zoals chatbots of andere generatieve AI-systemen.⁹ Hiervoor gelden met name transparantieplichtingen, om te voorkomen dat mensen door dergelijke systemen worden misleid. In de categorie 'minimaal risico' vallen AI-systemen die weinig tot geen risico vormen, zoals games of spamfilters. Hiervoor gelden in beginsel geen bijzondere regels.

3.2 Onaanvaardbaar risico

Artikel 5 AI-verordening verbiedt specifieke AI-systemen die een onaanvaardbaar risico meebrengen. Denk in de context van werk bijvoorbeeld aan AI-systemen die emoties kunnen herkennen op de werkplek, mensen categoriseren op basis van biometrische gegevens of systemen die werknemers tijdens hun werkzaamheden monitoren op oplettendheid en signalen afgeven om de werknemer scherp te houden (art. 5 lid 1 sub f AI-verordening). Ook systemen die vakbondslidmaatschap kunnen voorspellen, zijn verboden (art. 5 lid 1 sub g AI-verordening). Verder kan gedacht worden aan het verbod op *social credit systems*. Dit verbiedt echter niet alle beoordelingssystemen in arbeidsrelaties, enkel systemen die leiden tot een *"nadelige of ongunstige behandeling van bepaalde natuurlijke personen of groepen personen in een sociale context die geen verband houdt met de context waarin de data oorspronkelijk werden gegenereerd of verzameld"* (art. 5 lid 1 sub c onder i AI-verordening). Toch lijkt de begrenzing van contexten relevant te zijn voor de mogelijkheid om data die buiten het werk zijn verzameld op het werk te gebruiken.

3.3 Hoogrisico-AI-systemen: algemeen en in werkcontext

Welke AI-systemen een hoog risico hebben, volgt uit artikel 6 AI-verordening in samenhang met Bijlage III AI-verordening. Hieronder vallen ook AI-systemen die door werkgevers worden gebruikt in de context van werkgelegenheid, personeelsbeheer en toegang tot zelfstandige arbeid.¹⁰ De Europese Commissie geeft dan ook als voorbeeld *"software voor het sorteren van cv's voor aanwervingsprocedures"*.¹¹

⁹ Art. 50 AI-verordening.

¹⁰ Art. 4 sub a en c van Bijlage III AI-verordening noemt systemen voor *"het werven of selecteren van natuurlijke personen, met name voor het plaatsen van gerichte vacatures, het analyseren en filteren van sollicitaties, en het beoordelen van kandidaten"* en *"het nemen van besluiten die van invloed zijn op de voorwaarden van arbeidsgelateerde betrekkingen, de bevordering of beëindiging van arbeidsgelateerde contractuele betrekkingen, voor het toewijzen van taken op basis van individueel gedrag of persoonlijke eigenschappen of kenmerken, of voor het monitoren en evalueren van prestaties en gedrag van personen in dergelijke betrekkingen"*.

¹¹ Zie hierover M. Jovović & W. Roozendaal, 'Algoritmes en personeelsselectie', in: Houwerzijl, Montebovi & Zekić (red.), *Platformisering, algoritmisering en sociale bescherming* (MSR nr. 78), Deventer: Wolters Kluwer 2021/8.

Omdat de meeste AI-systemen die in een werkcontext zullen worden gebruikt in de categorie ‘hoog risico’ zullen vallen, gaan wij nader in op de eisen die de AI-verordening stelt aan dergelijke AI-systemen. Deze eisen leveren voornamelijk verplichtingen op voor de aanbieder van AI-systemen,¹² oftewel diegene die het systeem voor algemene doeleinden ontwikkelt, laat ontwikkelen of in de handel brengt.¹³ Meestal zal dat niet de werkgever zijn, maar als een grote werkgever zelf een dergelijk systeem ontwikkelt en gebruikt, moet hij mogelijk alsnog met al deze verplichtingen rekening houden. Meestal zal de werkgever echter de regulatoire positie van *gebruiksverantwoordelijke* hebben (dat is degene die een AI-systeem gebruikt onder zijn eigen verantwoordelijkheid).¹⁴ De AI-verordening schept ook verplichtingen voor *gebruiksverantwoordelijken*, die voortbouwen op de eisen die de AI-verordening aan AI-systemen stelt en op de verplichtingen voor aanbieders. Kortom: ook de algemene eisen die de AI-verordening aan AI-systemen stelt – en de onderliggende risico’s – zijn relevant voor de positie van de werkgever.

3.3.1 Verplichtingen met betrekking tot hoogrisico-AI-systemen

Ieder hoogrisico-AI-systeem moet, naast aan eventueel toepasselijke sector- of toepassings specifieke regels, aan vele AI-specifieke algemene eisen voldoen op grond van afdeling 2 AI-verordening.¹⁵ De AI-verordening stelt veel gedetailleerde eisen aan AI-systemen. Wij beperken ons hier tot de kern.

(A) Risicobeheersingssysteem (art. 9 AI-verordening)

Artikel 9 AI-verordening schrijft voor dat ieder AI-systeem een *risicobeheersingssysteem* moet hebben. In elk geval moet de aanbieder ‘aan de voorkant’ de bekende en de redelijkerwijs voorzienbare risico’s voor de gezondheid, veiligheid of grondrechten identificeren (zowel bij gebruik in overeenstemming met het beoogde doel als in situaties van redelijkerwijs te voorzien misbruik),¹⁶ met bijzondere aandacht voor mogelijke negatieve gevolgen voor jongeren onder achttien jaar en andere kwetsbare groepen. Dat is echter niet genoeg. Een risicobeheersingssysteem is immers een “*tijdens de gehele levensduur van het AI-systeem doorlopend en gepland iteratief proces, dat periodieke systematische toetsing en actualisering vereist*”. De aanbieder moet dus niet alleen het AI-systeem continu in de gaten houden, maar ook het proces en de criteria waarmee hij dat doet kritisch bekijken en updaten. Als bijvoorbeeld tijdens het gebruik van een fraudedetectiesysteem blijkt dat dit altijd Montenegrijnse collega’s aanwijst, ook als die eigenlijk niets verkeerd hebben gedaan, moet dat in het risicobeheersingssysteem worden verwerkt.

Risico’s moeten ook worden gemitigeerd door gepaste en gerichte risicobeheersmaatregelen te treffen.¹⁷ Risico’s moeten zoveel mogelijk technisch worden uitgesloten in het ontwerp en de ontwikkeling van het AI-systeem. Kan dat niet, dan moeten adequate maatregelen voor beperking en controle worden getroffen.¹⁸ Hierbij moet rekening worden gehouden met de gebruikscontext en te verwachten kennis en kunde van de gebruiksverantwoordelijke. Als een AI-systeem dus bedoeld is voor HR-adviseurs, kan de aanbieder niet verwachten dat de HR-adviseurs diepgaande statistische analyses gaan uitvoeren om te kijken of het systeem misschien discrimineert. Bij systemen die bedoeld zijn voor econometristen ligt dat misschien anders.

Kortom: een aanbieder die een algoritme ontwikkelt om sollicitanten te screenen, moet dus de risico’s voor grondrechten in kaart brengen en mitigeren – zowel voordat het algoritme op de markt wordt gebracht als daarna. De werkgever die het gebruikt, zal zich als gebruiksverantwoordelijke hiervan bewust moeten zijn.

(B) Kwaliteitseisen aan datasets (art. 10 AI-verordening)

Waar artikel 9 AI-verordening het gehele AI-systeem normmeert, stelt artikel 10 AI-verordening specifieke *kwaliteitseisen aan de datasets* waarmee de aanbieder AI-systemen ontwikkelt en AI-modellen traint. Hoewel er nogal wat meer komt kijken bij de ontwikkeling van een AI-systeem dan alleen de data, vormen de data wel een essentieel ingrediënt van AI-systemen. Daarop ziet de gevleugelde uitspraak ‘*garbage in, garbage out*’. Als een AI-systeem is ontwikkeld en getraind met goede datasets, draagt dat immers bij aan het beperken van de risico’s (en andersom).

Datasets moeten daarom worden onderworpen aan ‘praktijken’ op het gebied van databeheer die stroken met het beoogde doel van het AI-systeem,¹⁹ om discriminatie en biases te voorkomen.²⁰ Deze ‘praktijken’ moeten bijvoorbeeld betrekking hebben op: (1) ontwerpkeuzes; (2) de wijze van dataverzameling en (bij persoonsgegevens) het doel daarvan; (3) de onderliggende aannames over de informatie die de data meten en vertegenwoordigen; (4) de beschikbaarheid, kwantiteit en geschiktheid van de datasets; en (5) de eventuele aanwezigheid van vooringenomenheid die waarschijnlijk gevolgen heeft voor de gezondheid en de veiligheid van personen, nadelige effecten heeft op de grondrechten of leidt tot verboden discriminatie. Als van biases sprake is, moeten al in de fase van de samenstelling van de dataset passende maatregelen worden genomen om dergelijke vertekeningen op te sporen, te voorkomen en te beperken.²¹

12 Art. 16 sub a AI-verordening.

13 Art. 3 lid 3 AI-verordening.

14 Art. 3 lid 4 AI-verordening.

15 Art. 8 lid 1 jo. Bijlage I AI-verordening.

16 Art. 9 lid 2 sub a en b AI-verordening.

17 Art. 9 lid 2 sub d AI-verordening.

18 Art. 9 lid 5 sub a en b AI-verordening.

19 Art. 10 lid 2 AI-verordening.

20 Blijkens overweging 67 van de preambule moeten datasets “*de passende statistische kenmerken*” hebben, met bijzondere aandacht voor het tegengaan van vertekeningen in de datasets, die mogelijk leiden tot discriminatie die op grond van het Unierecht is verboden.

21 Art. 10 lid 2, slot, AI-verordening.

Deze eisen zijn echt belangrijk. Als de ontwerpers van de systemen van de Belastingdienst hiermee rekening hadden gehouden, had de toeslagenaffaire mogelijk kunnen worden voorkomen, omdat duidelijk zou zijn geworden dat gegevens over dubbele nationaliteit geen plaats mochten hebben in de AI-systemen voor risico-identificatie en fraudebestrijding.²²

Om het risico op garbage in, garbage out te beperken, moeten datasets voldoende representatief, zoveel mogelijk foutenvrij en volledig doelmatig zijn.²³ Ook moeten de datasets passende statistische kenmerken hebben, waaronder met betrekking tot de (groepen) personen ten aanzien van wie een hoogrisicosysteem zal worden gebruikt. De achtergrond van deze eis is dat een dataset waarin bepaalde groepen onder- of oververtegenwoordigd zijn, vaak zal leiden tot vertekende resultaten.²⁴ Een dataset die minder informatie bevat over vrouwen kan leiden tot een systeem dat veel minder nauwkeurig is voor vrouwen. Dan worden 'goede' vrouwen ten onrechte aangemerkt als 'slecht' en andersom. Als het goed is, leiden de door artikel 10 AI-verordening voorgeschreven praktijken ertoe dat dit soort risico's in de ontwikkelfase worden voorkomen.

Als dit nodig is gelet op het doel van het AI-systeem, moet de dataset ook rekening houden met de eigenschappen van elementen die specifiek zijn voor de geografische, contextuele, functionele of gedragsomgeving waarin het AI-systeem wordt gebruikt. Stel dus dat een Nederlandse salesmedewerker een slechte score krijgt van een algoritmisch beoordelingssysteem dat door een Amerikaanse aanbieder is ontwikkeld op basis van Amerikaanse salesmedewerkers. Kan hij dan zeggen dat de beoordeling niet deugt omdat de data voor hem niet representatief zijn nu ze uit een andere geografische, contextuele, functionele of gedragsomgeving komen?

(C) Technische documentatie en registratie van logs (art. 11 en 12 AI-verordening)

Uit de *technische documentatie* van het AI-systeem moet blijken dat en hoe is voldaan aan de eisen die voor hoogrisicosystemen gelden.²⁵ De documentatie moet de noodzakelijke informatie bevatten en moet helder en begrijpelijk zijn. Daarnaast geldt een *registratieplicht voor de 'logs'*: gedurende de levenscyclus van het systeem moeten gebeurtenissen automatisch worden geregistreerd.²⁶ Het AI-systeem moet een passend niveau van traceerbaarheid hebben.

22 Rapport Autoriteit Persoonsgegevens, 'Belastingdienst/Toeslagen De verwerking van de nationaliteit van aanvragers van kinderopvangtoeslag', Onderzoeksrapport z2018-22445, online via www.autoriteitpersoonsgegevens.nl/uploads/imported/onderzoek_belastingdienst_kinderopvangtoeslag.pdf (laatstelijk geraadpleegd op 3 april 2024).

23 Art. 10 lid 3 AI-verordening.

24 Voor een prima algemene introductie tot verschillende manieren waarop bias kan intreden, zie E. Ferrara, 'Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies', *Sci* 2024, 6 (1), 3; www.doi.org/10.3390/sci6010003 (laatstelijk geraadpleegd op 3 april 2025).

25 Art. 11 AI-verordening.

26 Art. 12 AI-verordening.

Het moet gebeurtenissen "*registreren die relevant zijn*" ter identificatie van situaties met mogelijke risico's voor de gezondheid, de veiligheid of de grondrechten van personen.²⁷ In elk geval moeten worden geregistreerd: de duur van elk gebruik van het systeem, de referentiedatabank die is gebruikt om de inputdata te controleren, de inputdata zelf en de natuurlijke persoon die het menselijk toezicht heeft uitgevoerd (zie hierna).

Zowel de technische documentatie als de logs zullen relevant kunnen zijn voor werkgevers en werknemers. In een arbeidsdiscriminatiezaak zou een rol kunnen gaan spelen welke risico's in de technische documentatie stonden omschreven en wat daarover uit de logs valt af te leiden.

(D) Transparantie en informatieverstrekking (art. 13 AI-verordening)

Aanbieders moeten ervoor zorgen dat hun systemen voldoen aan eisen met betrekking tot *transparantie en informatieverstrekking* aan gebruiksverantwoordelijken van het systeem.²⁸ De gebruiksverantwoordelijken moeten met deze informatie in staat zijn de output van het systeem te interpreteren en op passende wijze te gebruiken.²⁹ De informatieverstrekking gebeurt in ieder geval in gebruiksinstructies. Deze moeten beknopte én volledige,³⁰ juiste en duidelijke informatie bevatten die relevant, toegankelijk en begrijpelijk is voor gebruiksverantwoordelijken. De informatie gaat in elk geval over het doel van het systeem en de mate van nauwkeurigheid, waaronder de beoordelingsmaatstaven, robuustheid en cyberbeveiliging (zie voor materiële eisen aan nauwkeurigheid, robuustheid en beveiliging ook art. 15 AI-verordening).³¹ Gebruiksverantwoordelijken moeten worden geïnformeerd over eventuele bekende of te voorzien omstandigheden in verband met het gebruik die kunnen leiden tot risico's voor de gezondheid en veiligheid of de grondrechten (opnieuw zowel bij gebruik in overeenstemming met het beoogde doel als in situaties van redelijkerwijs te voorzien misbruik).

In de werkcontext zou het dus kunnen voorkomen dat een werknemer die zich ziet geconfronteerd met een nadelige of onwelgevallige beslissing, vraagt naar dergelijke risico's en begrenzingen die aan het AI-systeem kleven. Mogelijk zullen werknemers hierom gaan vragen al dan niet in samenhang

27 Art. 12 lid 2 sub a jo. art. 79 lid 1 AI-verordening.

28 Art. 13 AI-verordening.

29 Art. 13 lid 1 AI-verordening.

30 Dit zijn begrippen met een bepaalde onderlinge spanning.

31 Met name nauwkeurigheid zal naar verwachting interessant zijn, omdat hieronder zowel wordt verstaan het risico op foutpositieve als op foutnegatieve uitkomsten. Een foutnegatieve uitkomst is (bijv. in de context van een sollicitatieprogramma) dat het systeem iemand ten onrechte aanmerkt als een 'slechte kandidaat'. Iemand is dan eigenlijk wel geschikt, maar wordt niet als zodanig herkend. Een foutpositieve uitkomst daarentegen is dat iemand eigenlijk helemaal geen goede kandidaat is, maar toch als zodanig wordt aangemerkt. Je kunt je voorstellen dat de kans op foutpositieve uitkomsten in de context van de beoordeling van een beslissing om iemand aan te nemen minder relevant is dan de kans op foutnegatieve uitkomsten. Bij een systeem met een ander doel zal dat weer andersom kunnen zijn. Denk aan een beoordelingssysteem dat beoordeelt of iemand 'onderpresteert'.

met de technische documentatie (art. 11 AI-verordening) en de logs (art. 12 AI-verordening). Misschien zullen werkgevers moeten gaan uitleggen waarom de beslissing toch in stand kan blijven, ondanks de risico's (bijvoorbeeld voor het recht op gelijke behandeling of de privacy) en begrenzingen van het systeem (bijvoorbeeld voor de (on)nauwkeurigheid daarvan of wat het wel of niet kan zeggen). Wij verwachten discussie over de hiermee samenhangende vragen, de (on-)mogelijkheid van het geven van bevredigende antwoorden op de vragen en de (on)toereikendheid van die antwoorden.³²

(E) Menselijk toezicht (art. 14 AI-verordening)

AI-systemen moeten doeltreffend *menselijk toezicht* door de gebruiksverantwoordelijke mogelijk maken, om zo risico's voor de gezondheid, veiligheid of grondrechten te voorkomen (zowel bij normaal gebruik als bij voorzienbaar misbruik). Een AI-systeem moet zo in elkaar zitten, dat de uiteindelijke menselijke gebruiker de capaciteiten en beperkingen van het systeem begrijpt en kan monitoren. Gebruikers moeten in staat zijn onregelmatigheden en storingen te herkennen, op te sporen en te verhelpen en zij moeten de output juist kunnen interpreteren.³³ Zij moeten daarom worden geschoold en bewust gemaakt van de op de loer liggende neiging om automatisch of te veel te vertrouwen op de output van de computer (*automation bias*). Een mens moet altijd kunnen ingrijpen, door het hoogrisicosysteem niet te gebruiken, de output te negeren, een beslissing te vervangen of een beslissing terug te draaien. En er moet een stopknop zijn.

Een gebruiksverantwoordelijke blijft dus altijd verantwoordelijk voor keuzes die met behulp van een hoogrisicosysteem worden gemaakt of voorbereid. Het zal waarschijnlijk geen valide excuus zijn dat 'het algoritme dit nu eenmaal heeft gedaan'.

(F) Andere grondslagen voor materiële eisen aan AI-systemen?

De AI-verordening zal niet de enige bron zijn van materiële eisen aan de werking en resultaten van de AI-systemen. Die kunnen mogelijk ook uit grondrechten voortvloeien. Zo zou uit het recht op gelijke behandeling de eis kunnen voortvloeien dat systemen zoveel mogelijk niet-discriminerend zijn.³⁴

3.3.2 Verplichtingen voor gebruiksverantwoordelijken van hoogrisico-AI-systemen

Zoals gezegd, hangen de verplichtingen die gelden voor de werkgever/gebruiksverantwoordelijke samen met de verplichtingen voor aanbieders. Dat is ook niet gek: om de risico's van AI-systemen te beheersen, is het nodig dat deze systemen goed worden ontwikkeld en vervolgens in lijn daarmee worden gebruikt.

Werkgevers die hoogrisicosystemen gebruiken, moeten die op grond van het systeem gebruiken in overeenstemming met de gebruiksaanwijzing (art. 13 AI-verordening).³⁵ Zij moeten hiertoe ook passende maatregelen treffen, bijvoorbeeld door de gebruiksaanwijzing uit te werken met interne trainingen of beleid waarmee hun werknemers de output van het systeem kunnen interpreteren. De gebruiksverantwoordelijke moet het menselijk toezicht (art. 14 AI-verordening) opdragen aan natuurlijke personen die over de nodige bekwaamheid, opleiding en autoriteit beschikken en de nodige ondersteuning krijgen.³⁶ De gebruiksverantwoordelijke moet ervoor zorgen dat de inputdata (voor zover hij daar controle over heeft) relevant en voldoende representatief zijn.³⁷ De gebruiksverantwoordelijke moet het systeem doorlopend monitoren.³⁸ De logs (zie art. 12 AI-verordening) moeten ten minste een halfjaar worden bewaard.³⁹ Werknemers moeten, al dan niet via hun vertegenwoordigers (OR, PVT of eventueel vakbonden), worden geïnformeerd dat zij worden onderworpen aan een AI-systeem.⁴⁰ De gebruiker moet ervoor zorgen dat hij het AI-systeem op elk moment kan stoppen.⁴¹

Tot slot geldt op grond van artikel 86 AI-verordening een recht op uitleg bij individuele besluitvorming. Een persoon over wie op basis van de output van een hoogrisicosysteem een besluit wordt genomen, heeft op grond van artikel 86 AI-verordening recht op een duidelijke, inhoudelijke uitleg over de rol van het AI-systeem in de besluitvormingsprocedure en de voornaamste elementen van het genomen besluit. Dat is het geval als het besluit: óf (i) rechtsgevolgen heeft voor die persoon; óf (ii) op vergelijkbare wijze aanzienlijke invloed heeft die *de betrokkene beschouwt* als invloed met nadelige gevolgen voor de gezondheid, veiligheid of grondrechten. Niet duidelijk is hoever dit recht op een uitleg zich uitstrekt, en of het bijvoorbeeld ook recht geeft op inzicht in de documentatie, logbestanden, risicoanalyses, enzovoort. Over de precieze grenzen van dit recht op een uitleg verwachten wij, net als over de vereiste mate van

32 Zie voor vindplaatsen van literatuur over uitleg de bronnen die worden genoemd in par. 8.2.2 van W. Roozendaal & M. Jović, 'Algoritmes en personeelselectie', in: M.S. Houwerzijl, S.H.M. Montebovi & N. Zekić (red.), *Platformisering, algoritmisering en sociale bescherming. Sociaalrechtelijke uitdagingen in tijden van digitale transformatie* (MSR nr. 78), Deventer: Wolters Kluwer 2021. Zie voor een recent artikel A.P. Engelfriet, 'Een onuitlegbaar recht: Juridische en praktische grenzen van het recht op uitleg in de AI Act', *Tijdschrift voor Internetrecht* 2024/6, p. 223-226.

33 De AI-verordening verwijst hier wat cryptisch naar "*beschikbare instrumenten en methoden voor interpretatie*", art. 14 lid 4 sub e AI-verordening.

34 Overweging 6 van de preambule benoemt bijv. dat AI moet worden ontwikkeld in overeenstemming met art. 2 Verdrag betreffende de Europese Unie, waarin onder andere het non-discriminatiebeginsel is neergelegd, alsmede met de in de Verdragen vastgelegde grondrechten en fundamentele vrijheden. Overweging 31 waarschuwt expliciet voor het feit dat AI-systemen die worden gebruikt voor de beoordeling van natuurlijke personen, kunnen leiden tot een schending van fundamentele rechten, waaronder het recht op waardigheid en non-discriminatie.

35 Art. 26 lid 1 AI-verordening.

36 Art. 26 lid 2 AI-verordening. Zie echter ook lid 3, dat bepaalt dat dit geen afbreuk doet aan de overige juridische plichten (denk aan privacy of gelijke behandeling) van de gebruiker, maar ook niet aan zijn vrijheid om zijn organisatie in te richten zoals hij wil.

37 Art. 26 lid 4 AI-verordening.

38 Art. 26 lid 5 AI-verordening.

39 Art. 26 lid 6 AI-verordening.

40 Art. 26 lid 7 AI-verordening.

41 Aldus snijdt de Uniewetgever AI-systemen die zouden willen streven naar wereldheerschappij de pas af.

detail en diepgang van de uitleg, de komende tijd de nodige discussie.⁴²

3.3.3 Toezicht en boetes

Overtredingen van de AI-verordening kunnen in de hoog-risicocategorie tot zeer hoge boetes leiden: maximaal € 15.000.000 of (als dat hoger is) 3% van de totale wereldwijde omzet.

4. Situaties uit de praktijk

Om een idee te geven van arbeidsrechtelijk relevante situaties waarin de AI-verordening van toepassing kan zijn, geven wij hierna twee voorbeelden.

4.1 Sollicitatieprocedures

AI-systemen kunnen worden ingezet bij sollicitatieprocedures. Stel dat een werkgever dit doet met behulp van een extern ingekocht AI-systeem, aan welke eisen moet hij dan voldoen? Hij zal verplicht zijn het systeem te gebruiken in overeenstemming met de gebruiksaanwijzing van de aanbieder (art. 26 lid 1 AI-verordening). Die gebruiksaanwijzing is, als het goed is, in begrijpelijke taal opgesteld en geeft onder meer informatie over de werking van het systeem, over de nauwkeurigheid, over de beoordelingsmaatstaven en over de te verwachten risico's voor de veiligheid, gezondheid of grondrechten (art. 13 AI-verordening). Als de werkgever daarin bijvoorbeeld leest dat het systeem voor een bepaalde, door het gelijkebehandelingsrecht beschermde, groep minder nauwkeurig is of tot nadelige uitkomsten kan leiden, zal hij hiermee dus rekening moeten houden – door het systeem niet te gebruiken óf door daar bij het menselijk toezicht scherp op te zijn. OR of vakbond zouden hier scherp op kunnen zijn.

Een sollicitant die is afgewezen op basis van een beslissing die is voorbereid met behulp van een AI-systeem, zal daarover mogelijk moeten worden geïnformeerd op grond van artikel 86 AI-verordening. Als de afwijzing moet worden gekwalificeerd als beslissing met rechtsgevolg of als beslissing die op *“vergelijkbare wijze aanzienlijke invloed heeft die hij of zij als nadelige gevolgen voor zijn of haar gezondheid, veiligheid of grondrechten beschouwt”*, heeft de sollicitant recht op een uitleg. Mogelijk zal de sollicitant willen weten wat voor datasets (art. 10 AI-verordening) zijn gebruikt om het AI-systeem te ontwikkelen of te trainen en wat de kwaliteit daarvan is – een vrouw van kleur die is afgewezen terwijl er geen vrouwen van kleur in de dataset zaten, zal dat relevant kunnen vinden. Mogelijk zal de sollicitant logs opvragen (art. 12 AI-verordening), inzicht willen hebben in de technische documentatie (art. 13 AI-verordening), willen weten of het systeem risico's voor de grondrechten kent voor mensen met zijn of haar kenmerken (art. 13 jo. 9 AI-verordening), hoe het menselijk toezicht is uitgeoefend

(art. 14 AI-verordening) en of de inputdata relevant en voldoende representatief zijn voor het beoogde doel van het AI-systeem (art. 26 lid 4 AI-verordening). Weliswaar verplicht de AI-verordening niet direct tot het verstrekken van al deze informatie, maar een werkgever die zonder goede verklaring weigert relevante informatie te verschaffen die in zijn domein ligt, zou het wel eens moeilijk kunnen krijgen. Dat geldt eens te meer in gelijkebehandelingssituaties, gezien de bewijsvermoedens die daar gelden (zie art. 10 Awgb). Kortom: de rechten en plichten uit de AI-verordening kunnen invloed hebben op discussies over sollicitatiebeslissingen die algoritmisch zijn voorbereid of genomen.

4.2 Functioneringstrajecten

AI-systemen kunnen ook worden ingezet bij functioneringstrajecten. Denk daarbij aan het beoordelen van de prestaties van werknemers: wie is 'de beste callcentermedewerker' of 'de beste verkoper'? Aan welke eisen moet de werkgever dan voldoen?

Net als bij sollicitatieprocedures is de werkgever verplicht het AI-systeem te gebruiken in overeenstemming met de gebruiksaanwijzing (art. 13 AI-verordening). Ook voor het functioneren zal relevant zijn hoe het systeem werkt, hoe nauwkeurig het is, welke beoordelingsmaatstaven het hanteert en welke risico's voor de veiligheid, gezondheid of grondrechten worden verwacht (art. 13 AI-verordening).

Net als een sollicitant die is afgewezen op basis van een algoritmisch voorbereide of genomen beslissing, zal een algoritmisch slecht beoordeelde werknemer daarover moeten worden geïnformeerd – in elk geval achteraf (art. 86 AI-verordening) en mogelijk ook vooraf (art. 27 lid 6 AI-verordening). Deze werknemer zal mogelijk recht op een uitleg kunnen hebben (art. 86 AI-verordening). Een werknemer die het niet eens is met de beoordeling, zou wel eens de kwaliteit of betrouwbaarheid van het AI-systeem in twijfel kunnen trekken. In dat verband zou de werknemer interesse kunnen hebben in de datasets waarmee het AI-systeem is getraind (art. 10 AI-verordening). Hoewel deze verplichting strikt genomen bij de aanbieder van het AI-systeem ligt,⁴³ is immers goed voorstelbaar dat de werknemer hierover vragen zal hebben en dat die vragen in eerste instantie bij de werkgever zullen belanden. Bij wijze van (een extreem) voorbeeld: wat als het systeem door de aanbieder is getraind op vooral callcentermedewerkers, terwijl de werknemer advocaat is? Opnieuw zal de werknemer ter beoordeling van de beslissing interesse kunnen hebben in de logs (art. 12 AI-verordening), de technische documentatie (art. 13 AI-verordening), het menselijk toezicht (art. 14 AI-verordening) en de relevantie en representativiteit van de inputdata (art. 26 lid 4 AI-verordening).

Algoritmische beoordelingen kunnen allerlei nieuwe discussies opleveren. Wat geldt bijvoorbeeld als de werkgever een algoritmisch gegenereerde score niet goed kan

⁴² Zie ter illustratie A.P. Engelfriet, 'Een onuitlegbaar recht: Juridische en praktische grenzen van het recht op uitleg in de AI Act', *Tijdschrift voor Internetrecht* 2024/6, p. 223-224.

⁴³ Art. 16 sub a AI-verordening.

uitleggen? Kan hij er dan mee volstaan dat hij erop wijst dat de aanbieder het systeem zo heeft moeten opstellen dat de risico's voor onder andere de grondrechten en veiligheid van werknemers zijn beheerst? Of zal de rechter daar kritisch naar kijken in een ontbindingsprocedure op de d-grond?

4.3 *Overleg met de werknemersvertegenwoordiging*

Als werkgevers (hoogrisico-)AI-systemen willen implementeren, kan overleg met de OR, de PVT of de vakbonden zijn vereist. Op basis van artikel 25 lid 1 sub k WOR zou de OR in dergelijke gevallen een adviesrecht kunnen hebben.⁴⁴ Denkbaar is dat een plan om een AI-systeem te implementeren, kan worden gekwalificeerd als een voorgenomen besluit met betrekking tot de invoering of wijziging van een belangrijke technologische voorziening. Ook zou sprake kunnen zijn van een instemmingsplichtig besluit in de zin van artikel 27 WOR. Om enkele voorbeelden te noemen: een sollicitatiesysteem zal kunnen worden gekwalificeerd als een regeling op het gebied van het aanstellings-, ontslag- of bevorderingsbeleid (art. 27 lid 1 sub e WOR) en een systeem dat werknemers beoordeelt, zal kunnen worden gekwalificeerd als een regeling op het gebied van de personeelsbeoordeling (art. 27 lid 1 sub g WOR). En er is altijd de 'restcategorie' van artikel 27 lid 1 sub k WOR, dat ziet op regelingen omtrent het verwerken alsmede de bescherming van de persoonsgegevens van de in de onderneming werkzame personen. Daarvan zou immers sprake kunnen zijn als het AI-systeem persoonsgegevens van werknemers verwerkt en dit leidt tot een wijziging in het beleid omtrent persoonsgegevens. Ook voor het medezeggenschapsrecht kunnen AI-systemen dus relevant zijn.

5. **De AI-verordening zal dus ook relevant zijn voor het arbeidsrecht**

Omdat het gebruik van AI-systemen toeneemt, zowel binnen als buiten werksituaties, verwachten wij dat de AI-verordening in toenemende mate impact zal hebben op de arbeidsrechtspraktijk. Dat levert allerlei nieuwe rechten op (voor werknemers en hun vertegenwoordigers) en leidt tot nieuwe verantwoordelijkheden en verplichtingen (voor werkgevers), want de kans is groot dat een AI-systeem dat in arbeidssituaties wordt toegepast als hoogrisicosysteem zal worden aangemerkt. Als een werkgever zijn verplichtingen veronachtzaamt, loopt hij niet alleen het risico dat de toezichthouder hem een boete zal opleggen. De AI-verordening zal immers ook in 'klassieke' arbeidsrechtelijke discussies tussen werkgever en werknemer, zoals over sollicitaties of beoordelingen, kunnen doorwerken.

⁴⁴ Uit de parlementaire geschiedenis volgt dat de OR vroegtijdig moet worden betrokken. Zie *Kamerstukken II 1995/96*, 24615, nr. 3, p. 42.