De Europese AI-verordening: complexer dan een piramide

Prof. dr. A.C.M. Meuwese en prof. mr. dr. C.J. Wolswinkel*

Dit artikel biedt een beknopt overzicht van de in de zomer van 2024 eindelijk definitief geworden AI-verordening. Terugkijkend wordt ingegaan op de transformatie die deze veelbesproken 'eerste AI-wet ter wereld' heeft doorgemaakt ten opzichte van het Commissievoorstel uit 2021. Om tegemoet te komen aan geuite kritiek op zaken als de definitie van AI, de bijzondere aard van AI als 'product' en het ontbreken van regels voor generatieve AI, is de verordening steeds complexer geworden. Daarnaast wordt vooruitkijkend ingegaan op de uitdagingen die de AI-verordening voor de nationale wetgever meebrengt en op de positie van de AI-verordening in relatie tot het nieuwe Kaderverdrag over AI van de Raad van Europa.

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) nr. 300/2008, (EU) nr. 167/2013, (EU) nr. 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).

1. Inleiding

In een bespreking van het oorspronkelijke Commissievoorstel[1] voor een Europese AI-verordening uit 2021 waagden wij ons aan de voorspelling dat 'het voorstel de eindstreep gaat bereiken, weliswaar in gewijzigde vorm wat betreft definities en formuleringen, maar ongeschonden ten aanzien van de hoofdlijnen en de "piramidevorm" in het bijzonder'.[2] We hadden destijds eigenlijk moeten kijken of er *bookmakers* waren die ons hier geld op hadden laten inzetten, want deze voorspelling is in het voorjaar van 2024 uitgekomen. In de tussentijd werd er in de media en het publieke debat vaak over de verordening gesproken alsof deze al een feit was. Inderdaad was het witterookmoment in dit dossier waarschijnlijk december 2023 met de 'doorbraak' in de trialoog. Maar daarna volgden nog vele rondes van stemmingen en taalkundige aanpassingen, voordat op 12 juli 2024 Verordening (EU) 2024/1689, de 'AI Act', zoals deze inmiddels ook in de Nederlandse wandelgangen is gaan heten, in het officiële publicatieblad is terechtgekomen.[3] Inmiddels is de AI-verordening per 1 augustus 2024 in werking getreden, maar de meeste materiële bepalingen gaan pas gelden 24 maanden na inwerkingtreding, dus in augustus 2026.[4]

Hoewel aanpassingen aan het oorspronkelijke voorstel kenmerkend zijn voor vrijwel elk wetgevingsproces, is er een extra reden voor een nieuwe bijdrage over de AI-verordening.[5] Juist de ontwikkeling van artificiële intelligentie (AI) als technologie heeft in de tussentijd niet stil gestaan, evenmin als het maatschappelijke debat over

* Prof. dr. A.C.M. (Anne) Meuwese is hoogleraar Public Law & Governance of Artificial Intelligence aan de Universiteit Leiden. Prof. mr. dr. C.J. (Johan) Wolswinkel is hoogleraar Bestuursrecht aan Tilburg University.

1 COM(2021)206 def.

2 A.C.M. Meuwese & C.J. Wolswinkel, 'Een Wet op de Artificiële Intelligentie? De Europese wetgever haalt de nationale in', *NJB* 2022/92, afl. 2, p. 92-100.

3 Verordening (EU) 2024/1689 van het Europees Parlement en de Raad van 13 juni 2024 tot vaststelling van geharmoniseerde regels betreffende de artificiële intelligentie en tot wijziging van de Verordeningen (EG) nr. 300/2008, (EU) nr. 167/2013, (EU) nr. 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).

4 Afwijkende inwerkingtredingsdata zijn er voor: regels over verboden toepassingen (zes maanden na inwerkingtreding = februari 2025), regels over praktijkcodes (negen maanden na inwerkingtreding = lente 2025), bepalingen over de 'aangemelde instanties' en het 'governancekader' (twaalf maanden na inwerkingtreding = zomer 2025), regels voor 'AI voor algemene doeleinden' (twaalf maanden na inwerkingtreding = augustus 2025) en bepalingen rond sancties (twaalf maanden na inwerkingtreding = augustus 2025).

5 Overigens is momenteel de trend om in Nederlandstalige stukken Engelstalige afkortingen te gebruiken voor Europese tech-wetgeving. Zie A.C.M. Meuwese & D.R.P. de Kok, 'Toegankelijkheid van wetgeving: what's in a name?', *RegelMaat* 2024, afl. 1, p. 3-6.

doi: 10.5553/NtER/138241202024030005004

de risico's van AI.[6] In dit nieuwe overzichtsartikel staat niet alleen de opgave die er de komende (twee) jaren ligt voordat de AI-verordening volledig gaat gelden centraal, maar ook de transformatie die de verordening heeft doorgemaakt in de drie jaar tussen voorstel en publicatieblad.

Al snel na publicatie van het Commissievoorstel in 2021 zwol de kritiek hierop aan. De definitie van AI werd te breed en te weinig gangbaar bevonden en er was twijfel over de geschiktheid van 'productveiligheidswetgeving' voor een 'systeemtechnologie'[7] als AI. Daarnaast waren er bij verschillende belanghebbenden zorgen over de balans tussen ruimte voor innovatie en bescherming van kwetsbare belangen. Volgens de AI-industrie stond de voorgestelde AI-verordening ontwikkeling en beschikbaarheid van AI-technologie in Europa in de weg, terwijl belangengroeperingen die aandacht vragen voor digitaliseringsrisico's, juist meenden dat sprake was van te weinig aandacht voor fundamentele rechten en inhoudelijke waarden. Later kwam daar een breed gedragen overtuiging bij dat een Europese AI-verordening veel meer aandacht zou moeten besteden aan de opkomst van nieuwe, met name zogeheten 'generatieve' AI-modellen, zoals het bekende GPT-3.5 van OpenAI, die nog niet sterk aan de orde waren toen het Commissievoorstel in 2021 werd gelanceerd. Het oorspronkelijke Commissievoorstel ging vooral over zogeheten '*narrow*' AI-systemen, die voor een specifiek doel of een specifieke taak ontwikkeld worden, simpelweg omdat niemand destijds nog van ChatGPT gehoord had. Er leek wel consensus te bestaan over de risicogebaseerde aanpak die de verordening voorschreef ten aanzien van deze AI-systemen met een specifiek doel. Die aanpak wordt vaak gevisualiseerd als een 'piramide' met in de top de verboden toepassingen, daaronder de hoogrisicotoepassingen en AI-systemen waarvoor enkel transparantie-eisen gelden en helemaal onderin de laagrisicosystemen. De vraag voor de EU-wetgever was dus hoe generatieve AI, dat veel minder goed is af te bakenen qua mogelijke toepassingen, in te passen viel in deze structuur.

Wij kijken eerst terug met een hoofdstuksgewijze bespreking van de definitieve tekst van de AI-verordening in verhouding tot het oorspronkelijke Commissievoorstel. Uiteraard is het niet mogelijk om binnen het bestek van dit artikel een uitputtend overzicht te geven van de verordening, ook niet samenvattend. Daarom blijven sommige hoofdstukken grotendeels onbesproken en kiezen wij ervoor om binnen elk hoofdstuk bepaalde accenten te leggen die de lezer zouden moeten helpen de kern van de AI-verordening te doorgronden of die illustratief zijn voor de transformatie die de tekst de afgelopen drie jaar heeft doorgemaakt. Vervolgens gaan wij dieper in op de uitdagingen voor de komende jaren, in het bijzonder binnen de Nederlandse rechtsorde: de opdracht aan de nationale wetgever en beleidsmakers en de positie van de nieuwe verordening binnen het mondiale reguleringsveld. In de slotbeschouwing vatten wij de belangrijkste keuzes van de Europese wetgever samen en benoemen wij twee risico's die deze keuzes met zich lijken te gaan brengen.

2. De AI-verordening in vogelvlucht

2.1 Algemene bepalingen (hoofdstuk I)

Aangezien de Verdragen geen specifieke bevoegdheidsbasis voor regelgeving over AI bieden (zoals art. 16 Verdrag betreffende de werking van de Europese Unie (VWEU) wel biedt op het vlak van de bescherming van persoonsgegevens), was artikel 114 VWEU in dit wetgevingsproces onmisbaar. Daarmee was de keuze voor internemarktregulering, om precies te zijn 'productveiligheidswetgeving', er een van *second best* en het kritiekpunt dat AI-systemen eigenlijk geen 'product' in klassieke zin zijn, lastig te adresseren. De Uniewetgever heeft wel geprobeerd de bijzondere aard van dit 'product' specifiek mee te nemen in de reguleringsaanpak. Dit blijkt allereerst uit hoofdstuk I, waar de lijst met te definiëren termen nog flink uitgebreid is. Een term die daar *niet* gedefinieerd wordt, maar die wel cruciaal is voor de pogingen van de Europese wetgever om tegemoet te komen aan dit kritiekpunt, is het begrip 'AI-waardeketen'. Hiermee wordt bedoeld dat een AI-systeem meestal niet op één moment en plaats en door een bepaalde partij 'geproduceerd' wordt, maar dat het een 'keten' betreft, waarbinnen bepaalde modellen worden ontwikkeld, gefinetuned en daarna weer geïntegreerd in software of andere producten. Hoewel de meeste verplichtingen in de AI-verordening nog steeds rusten op aanbieders[8] van AI-systemen is er meer aandacht gekomen voor wie eerst nog 'gebruikers', daarna 'exploitanten' en inmiddels 'gebruiksverantwoordelijken' van AI-systemen heetten/heten.[9]

In de definitiebepaling van artikel 3 gaat de meeste aandacht uit naar de, in de Nederlandstalige versie wat houterig klinkende, definitie van een 'AI-systeem'. Om tegemoet te komen aan de kritiek dat deze verwarrend was en niet aansloot bij bestaande definities, heeft de Europese wetgever uiteindelijk voor een functionele definitie

6 Zie bijvoorbeeld NOS Nieuws, 'Toezichthouder: samenleving moet zich voorbereiden op incidenten met AI', 18 juli 2024.

7 Wetenschappelijke Raad voor het Regeringsbeleid, *Opgave AI. De nieuwe systeemtechnologie*, Den Haag 2021.

8 Art. 3 lid 3 AI-verordening definieert 'aanbieder' als 'een natuurlijke of rechtspersoon, overheidsinstantie, agentschap of ander orgaan die/dat een AI-systeem of een AI-model voor algemene doeleinden ontwikkelt of laat ontwikkelen en dat systeem of model in de handel brengt of het AI-systeem in gebruik stelt onder de eigen naam of merknaam, al dan niet tegen betaling'. In de Engelstalige versie is ook een terminologische wijziging doorgevoerd. Omdat *developer* een te smalle term werd bevonden, wordt nu gesproken van *providers*. 'Gebruikers' zijn *deployers* gaan heten, in plaats van *users*.

9 Art. 3 lid 4 AI-verordening definieert 'exploitant' als 'een natuurlijke of rechtspersoon, overheidsinstantie, agentschap of ander orgaan die/dat een AI-systeem onder eigen verantwoordelijkheid gebruikt, tenzij het AI-systeem wordt gebruikt in het kader van een persoonlijke niet-beroepsactiviteit'.

doi: 10.5553/NtER/138241202024030005004

gekozen, die dichter ligt bij de definities uit de voorbereidende documenten[10] en de definitie die de Organisatie voor Economische Samenwerking en Ontwikkeling (OESO) hanteert.[11] Een AI-systeem wordt nu gedefinieerd als 'een machinaal systeem dat is ontworpen om met verschillende niveaus van autonomie te werken en dat na de uitrol 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'.[12]

In de reacties op de verordening valt op dat er nogal eens van uit wordt gegaan dat dit een zeer brede definitie is die alle algoritmische systemen omvat.[13] Zowel overweging 12 uit de considerans als de tekst van de definitie zelf geeft echter aanleiding om hier ook weer niet in alle gevallen te snel van uit te gaan. In deze overweging staat namelijk dat 'eenvoudigere traditionele softwaresystemen of programmeringsbenaderingen' niet onder de definitie vallen en dat deze 'geen betrekking [mag] hebben op systemen die gebaseerd zijn op regels die uitsluitend door natuurlijke personen zijn vastgesteld om automatisch handelingen uit te voeren'.[14]

2.2 Verboden AI-praktijken (hoofdstuk II)

Artikel 5 bevat een opsomming van acht, in plaats van de oorspronkelijke vier, AI-praktijken die door de Europese wetgever zo schadelijk worden geacht, en zozeer in tegenspraak met de Uniewaarden van respect voor menselijke waardigheid, vrijheid, gelijkheid, democratie, rechtsstaat en fundamentele rechten,[15] dat zij verboden zijn.

Veel categorieën draaien om misbruik gerelateerd aan kwetsbaarheden en emoties. Zo worden AI-systemen verboden die subliminale technieken gebruiken waarvan personen zich niet bewust zijn of die doelbewust manipulatieve of misleidende technieken gebruiken.[16] Hetzelfde geldt voor systemen die zwaktes van een persoon uitbuiten en leiden tot keuzes of handelingen die schade bij een persoon veroorzaken[17] en systemen om de betrouwbaarheid van een person in te schatten (beter bekend onder de term '*social scoring*').[18] De opsomming gaat verder met het verbod op het gebruiken van een AI-systeem voor 'risicobeoordelingen van natuurlijke personen om het risico dat een natuurlijke persoon een strafbaar feit pleegt te beoordelen of te voorspellen, uitsluitend op basis van de profilering van een natuurlijke persoon of op basis van de beoordeling van diens persoonlijkheidseigenschappen en -kenmerken'[19]. Nieuw in de lijst, ten opzichte van het Commissievoorstel en eerder bekend geworden onderhandelingsversies, zijn het verbod op 'AI-systemen die databanken voor gezichtsherkenning aanleggen of aanvullen door ongerichte scraping van gezichtsafbeeldingen van internet of CCTV-beelden',[20] het verbod op 'AI-systemen om emoties van een natuurlijke persoon op de werkplek en in onderwijsinstellingen af te leiden'[21] en het verbod op systemen voor biometrische categorisering.[22] Dit laatste type AI-systeem refereert aan het indelen in categorieën van individuele natuurlijke personen 'op basis van biometrische gegevens om hun ras, politieke opvattingen, lidmaatschap van een vakbond, religieuze of levensbeschouwelijke overtuigingen, seksleven of seksuele gerichtheid af te leiden'. De laatste verboden praktijk is waarschijnlijk de meest besproken en het meest op detailniveau uitonderhandeld, namelijk het gebruik van systemen 'voor biometrische identificatie op afstand in real time in openbare ruimten met het oog op de rechtshandhaving'.[23] Daaraan wordt vervolgens een lijst van doelstellingen toegevoegd die *real time* biometrische identificatie wel mogelijk maken als die techniek strikt noodzakelijk is, maar met inhoudelijke en procedurele randvoorwaarden geclausuleerd wordt.

Overigens gelden soms specifieke uitzonderingen en geldt voor alle toepassingen dat het verbod gekoppeld is aan het (waarschijnlijk) optreden van bepaalde gevolgen of andere randvoorwaarden. Zo moet het bij *social scoring* gaan om 'de nadelige of ongunstige behandeling van bepaalde natuurlijke personen of volledige groepen personen in een sociale context die geen verband houdt met de context waarin de data oorspronkelijk werden gegenereerd of verzameld' en/of 'de nadelige of ongunstige behandeling van bepaalde natuurlijke personen of groepen personen die ongerechtvaardigd of onevenredig met hun sociale gedrag of de ernst hiervan is'.

2.3 AI-systemen met een hoog risico (hoofdstuk III)

De meeste normen binnen de AI-verordening hebben betrekking op AI-systemen die een 'hoog risico' vormen voor de gezondheid en veiligheid van mensen of voor de grondrechten. Inhoudelijk moeten deze systemen voldoen aan een aantal in de AI-verordening gespecificeerde eisen om hun betrouwbaarheid te waarborgen. Met

10 COM(2018)237 def; EU High-Level Expert Group on Artificial Intelligence, Ethical Guidelines for Trustworthy AI, april 2019; COM(2020)65 def.

11 De OESO-definitie luidt: 'AI system means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments'.

12 Art. 3 lid 1 AI-verordening.

13 In de media werd bijvoorbeeld Europarlementariër Van Sparrentak als volgt geciteerd: 'De technische dingen die ervoor zorgden dat het toeslagenschandaal mogelijk was, die zijn nu verboden'. N. Kasteleijn, 'Zwaarbevochten wet kunstmatige intelligentie: wat mag straks wel en niet?', nos.nl, 11 december 2023.

14 Overweging 12 AI-verordening.

15 Overweging 28 AI-verordening.

16 Art. 5 lid 1 onder a AI-verordening.

17 Art. 5 lid 1 onder b AI-verordening.

18 Art. 5 lid 1 onder c AI-verordening.

19 Art. 5 lid 1 onder d AI-verordening. Er was nog even sprake van dat dit niet alleen voor strafbare feiten, maar ook voor administratieve overtredingen zou gaan gelden, maar dit voorstel van het Europees Parlement heeft het niet gehaald.

20 Art. 5 lid 1 onder e AI-verordening.

21 Art. 5 lid 1 onder f AI-verordening.

22 Art. 5 lid 1 onder g AI-verordening.

23 Art. 5 lid 1 onder h AI-verordening.

doi: 10.5553/NtER/138241202024030005004

het oog hierop worden aan verschillende partijen in de AI-waardeketen, waaronder aanbieders en gebruiksverantwoordelijken, verplichtingen opgelegd.

Grofweg zijn er twee manieren waarop AI-systemen als 'hoog risico' worden geclassificeerd onder de AI-verordening. In de eerste plaats kunnen AI-systemen een veiligheidscomponent zijn in producten die als zodanig reeds zijn onderworpen aan een Unierechtelijke verplichting tot conformiteitsbeoordeling door een derde partij, zoals speelgoed of liften.[24] De AI-verordening verwijst in dit verband naar een lijst van bestaande EU-harmonisatiewetgeving. Deze wetgeving is grotendeels gebaseerd op het zogeheten 'nieuwe wetgevingskader',[25] dat het op een veilige wijze in de handel brengen van producten op de Europese markt regelt.[26] De reden om deze 'afhankelijke' AI-systemen als 'hoog risico' aan te merken, is dat juist wanneer AI-systemen fungeren als veiligheidscomponent,[27] zij nadelige gevolgen kunnen hebben voor de gezondheid en de veiligheid van personen.[28] Doordat op de producten waarvan deze AI-systemen onderdeel zijn, ook andere Uniewetgeving van toepassing is die dwingt tot een conformiteitsbeoordeling van dit product als zodanig, geldt specifiek voor deze categorie van AI-systemen met een hoog risico dat beide regelgevingskaders op elkaar moeten worden afgestemd.

In de tweede plaats zijn er AI-systemen die niet verbonden zijn met andere harmonisatiewetgeving, maar zelfstandig als 'hoog risico' worden bestempeld.[29] Het gaat hierbij om *stand alone* of autonome AI-systemen die worden toegepast in een aantal specifieke domeinen, te weten: biometrie, kritieke infrastructuur, onderwijs en beroepsopleiding, werkgeleiden, toegang tot essentiële (private en publieke) diensten, rechtshandhaving, migratie en grenscontrole, en rechtspraak en democratische processen.[30] De 'hoogrisicotoepassingen' (*use cases*) binnen deze domeinen zijn opgenomen in Bijlage III. In de tekst van de verordening wordt de soep echter niet zo heet gegeten als die wordt opgediend: wanneer binnen deze domeinen een AI-systeem geen significant risico op schade voor de gezondheid, veiligheid of de grondrechten van natuurlijke personen inhoudt, onder meer doordat het systeem de uitkomst van de besluitvorming niet wezenlijk beïnvloedt, dan is geen sprake van 'hoog risico' en hoeft het dus niet aan de voorschriften van dit hoofdstuk te voldoen,[31] op registratie in de EU-databank na.[32]

Waarom is het van belang of een AI-systeem al dan niet als een systeem met hoog risico wordt beschouwd? Voor AI-systemen met een hoog risico bevat de AI-verordening een aantal (algemeen geformuleerde) eisen, 'rekening houdend met hun beoogde doel en de algemeen erkende stand van de techniek op het gebied van AI en AI-gerelateerde technologieën'.[33] Deze eisen hebben onder meer betrekking op risicobeheer, datagovernance, technische documentatie, logging, transparantie, menselijk toezicht en robuustheid.[34] Op de aanbieder van AI-systemen met een hoog risico rust vervolgens onder meer de verplichting ervoor te zorgen (1) dat hun systemen in overeenstemming zijn met deze eisen en (2) dat voor deze systemen de desbetreffende conformiteitsbeoordelingsprocedure wordt uitgevoerd voordat dit systeem in de handel wordt gebracht of in gebruik wordt gesteld.[35]

Deze figuur van conformiteitsbeoordeling, die al langer binnen het 'nieuwe wetgevingskader' wordt toegepast in het kader van productveiligheid, speelt een centrale rol bij AI-systemen met een hoog risico. Wanneer deze systemen op grond van een dergelijke beoordeling in overeenstemming blijken te zijn met zogeheten 'geharmoniseerde normen', worden deze systemen tevens geacht in overeenstemming te zijn met de (algemeen geformuleerde) eisen die de AI-verordening aan deze systemen stelt.[36] Deze 'geharmoniseerde normen' worden op verzoek van de Commissie opgesteld door normalisatieorganisaties,[37] waarin allerlei belanghebbenden zijn vertegenwoordigd, inclusief het midden- en kleinbedrijf (mkb), consumentenorganisaties en belanghebbenden op sociaal en milieugebied.[38]

Een conformiteitsbeoordeling kan zowel door de aanbieder zelf als door een externe instantie worden verricht. Ten aanzien van 'afhankelijke' AI-systemen die

24 Art. 6 lid 1 AI-verordening.

25 Verordening (EG) Nr. 765/2008 van het Europees Parlement en de Raad van 9 juli 2008 tot vaststelling van de eisen inzake accreditatie en markttoezicht betreffende het verhandelen van producten en tot intrekking van Verordening (EEG) nr. 339/93 (*PbEG* 2008, L 218/30) en Besluit Nr. 768/2008/EG van het Europees Parlement en de Raad van 9 juli 2008 betreffende een gemeenschappelijk kader voor het verhandelen van producten en tot intrekking van Besluit 93/465/EEG van de Raad (*PbEG* 2008, L 218/82).

26 Zie Bijlage I van de AI-verordening, waarbinnen een onderscheid wordt gemaakt tussen harmonisatiewetgeving van de Unie op basis van het nieuwe wetgevingskader (afdeling A) en andere harmonisatiewetgeving van de Unie (afdeling B), zoals vierwielers. Voor zover producten onder afdeling B vallen, is de AI-verordening slechts zeer beperkt van toepassing: zij vereist enkel dat bij de vaststelling van uitvoeringswetgeving op basis van die andere harmonisatiewetgeving de eisen inzake AI-systemen met een hoog risico 'in acht worden genomen' (art. 2 lid 2 jo. art. 102 tot en met 109 AI-verordening).

27 Art. 3 onder 14 AI-verordening definieert een 'veiligheidscomponent' als 'een component van een product of van een AI-systeem die een veiligheidsfunctie voor dat product of AI-systeem vervult of waarvan het falen of gebrekkig functioneren de gezondheid en veiligheid van personen of eigendom in gevaar brengt'.

28 Zie overweging 47 AI-verordening.

29 Art. 6 lid 2 AI-verordening.

30 Bijlage III van de AI-verordening. Zie specifiek over deze laatste categorie van de rechtspraak: B.J. van Ettekoven & J.E.J. Prins, 'Artificiële intelligentie en de Rechtspraak. Implicaties van de Europese AI Act en de noodzaak van een kompas voor toepassing en beoordeling van AI systemen', *NJB* 2023, afl. 36, p. 3153-3162.

31 Art. 6 lid 3 AI-verordening. Zie ook overweging 53.

32 Art. 71 AI-verordening.

33 Art. 8 lid 1 AI-verordening.

34 Afdeling 2 van Hoofdstuk III van de AI-verordening.

35 Art. 16 onder a en f AI-verordening. Ook buiten hoofdstuk III bevat de AI-verordening een aantal verplichtingen voor aanbieders van AI-systemen met een hoog risico, waaronder de verplichting om een systeem voor monitoring op te zetten (art. 72) en een meldingsplicht voor ernstige incidenten (art. 73).

36 Art. 40 AI-verordening.

37 Wanneer dergelijke geharmoniseerde normen ontbreken, kan de Commissie zelf zogeheten 'gemeenschappelijke specificaties' vaststellen die eenzelfde functie vervullen als geharmoniseerde normen.

38 Art. 5 AI-verordening.

doi: 10.5553/NtER/138241202024030005004

een veiligheidscomponent binnen een gereguleerd product vormen, geldt per definitie dat deze beoordeling door een externe instantie moet worden verricht. Met het oog op de afstemming van beide regelgevingskaders bepaalt de AI-verordening dat de conformiteitsbeoordeling ten aanzien van de eisen die aan het AI-systeem worden gesteld, deel uitmaakt van de conformiteitsbeoordeling van het gereguleerde product als zodanig (zoals het speelgoed of de lift).[39] Voor 'stand-alone' AI-systemen volstaat daarentegen in beginsel een interne controle door de aanbieder zelf.[40]

Hoewel een groot deel van de tekst van de AI-verordening is gewijd aan AI-systemen met een hoog risico, is de vraag hoeveel AI-systemen als zodanig kwalificeren. De verwachting is dat slechts een beperkt deel van de AI-systemen zal moeten worden aangemerkt als AI-systemen met een hoog risico.[41] Tegelijkertijd zal, wanneer sprake is van een systeem binnen deze categorie, uit de regels van de AI-verordening voor deze categorie systemen het meeste werk voor de publieke en de private sector voortvloeien.[42]

In bepaalde omstandigheden worden distributeurs, importeurs, gebruiksverantwoordelijken of derden als aanbieder van een AI-systeem met een hoog risico beschouwd. Zij zijn dan onderworpen aan de verplichtingen die rusten op aanbieders en de aanbieder die het AI-systeem voor het eerst in de handel heeft gebracht niet meer. Dit is het geval als zij hun naam of handelsmerk op een al in de handel gebracht AI-systeem zetten, als zij een substantiële wijziging daarin aanbrengen met behoud van een hoog risico of als zij het beoogde doel van een AI-systeem dat niet als een systeem met een hoog risico was geclassificeerd zodanig wijzigen dat het wel als een hoogrisicosysteem gezien moet worden.[43] Hiermee komt de al genoemde zogeheten AI-waardeketen tot uitdrukking: een AI-product is niet statisch. In aanvulling hierop gelden voor importeurs, distributeurs en gebruiksverantwoordelijken ook eigen verplichtingen, waarbij met name opvallend is dat ten opzichte van het Commissievoorstel de aandacht voor de gebruiksverantwoordelijke aanzienlijk is toegenomen. Zo heeft elke getroffen persoon het recht op uitleg van een gebruiksverantwoordelijke bij het gebruik van AI-systemen ten behoeve van individuele besluitvorming.[44] Verder zijn overheidsinstanties die een AI-systeem in gebruik willen stellen, verplicht om eerst een beoordeling uit te voeren van de gevolgen voor de grondrechten.[45] Hoewel de AI-verordening als zodanig geldt voor zowel de publieke als de private sector,[46] betreffen veel *use cases* binnen de AI-systemen met een hoog risico naar hun aard vooral de publieke sector.[47]

2.4 Transparantieverplichtingen voor aanbieders en gebruiksverantwoordelijken van bepaalde AI-systemen (hoofdstuk IV)

Een categorie die in de oorspronkelijke piramidestructuur ook al bestond, is die voor 'bepaalde AI-systemen' waarvoor transparantieverplichtingen gelden. We moeten hierbij met name denken aan het gebruik van chatbots of deepfakes. Voor deze toepassingen is een eenmalige conformiteitsbeoordeling niet geschikt om de risico's te mitigeren; die treden immers vooral op als consumenten met AI gegenereerde afbeeldingen voor niet-kunstmatig aanzien.

Aanbieders moeten daarom ervoor zorgen dat bij chatbots wordt gemeld dat het een AI-systeem betreft in de interactie met natuurlijke personen. Als het systemen betreft die beeld, audio, video of tekst kunnen genereren, moeten zij ervoor zorgen dat de outputs van het AI-systeem worden gemarkeerd als kunstmatig gegenereerd of gemanipuleerd. Als het echt gaat om *deepfakes*, rust er ook een verplichting op gebruiksverantwoordelijken om bekend te maken dat de content kunstmatig is gegenereerd of gemanipuleerd.[48] Het is nog maar de vraag of in alle gevallen de techniek al ver genoeg gevorderd is om het 'watermerk', waarop het naleven van deze verplichtingen al snel zal neerkomen, naar behoren te laten functioneren.[49] In dit verband dient opgemerkt te worden dat de EU-wetgever de norm ook als inspanningsverplichting geformuleerd heeft: 'aanbieders zorgen ervoor dat hun technische oplossingen doeltreffend, interoperabel, robuust en betrouwbaar zijn voor zover dat technisch haalbaar is'.[50]

2.5 AI-modellen voor algemene doeleinden (hoofdstuk V)

De ontwikkelingen rond generatieve AI, in de volksmond, ChatGPT, zijn in de AI-verordening terechtgekomen in het hoofdstuk 'AI-modellen voor algemene doeleinden' dat in het initiële Commissievoorstel nog niet bestond. Een complicatie in de laatste stadia van het onderhandelingsproces was dat de Europese generatieve AI-industrie net een beetje op gang begon te komen, Frans paradepaardje Mistral voorop. De wetgever moest

39 Art. 43 lid 3 AI-verordening.

40 Art. 43 lid 1 en 2 AI-verordening. De enige uitzondering hierop vormt toepassing van AI-systemen in het kader van biometrie in het geval dat geharmoniseerde normen of gemeenschappelijke specificaties ontbreken.

41 Applied AI, *AI Act: Risk Classification of AI Systems from a Practical Perspective* (maart 2023), identificeert binnen een steekproef slechts 18% van de onderzochte AI-systemen als 'hoog risico' (en 5% als verboden). De Europese Commissie ging zelf uit van 5 tot 15% (Commissie, *Impact Assessment*, 2021).

42 Presentatie AI-verordening, technische briefing voor de Commissie Digitale Zaken van de Tweede Kamer, 30 mei 2024, https://debatdirect.tweedekamer.nl.

43 Art. 25 ('Verantwoordelijkheden in de AI-waardeketen') en art. 16 AI-verordening.

44 Art. 86 AI-verordening.

45 Art. 27 AI-verordening.

46 Het verbod op (bepaalde vormen van) *social scoring* (art. 5 lid 1 onder c AI-verordening) is zelfs uitgebreid naar private toepassingen.

47 Zie A.A.H.M. van der Wijst, '"Computer says no" of "experts say no" – wat is de impact van de AI-verordening op de black box jurisprudentie', *Tijdschrift voor Internetrecht* 2024, afl. 3, p. 122-128 voor een analyse van de effecten van de AI-verordening op SyRI.

48 Art. 50 lid 1, 2 en 4 AI-verordening.

49 Th. Gils, 'A Detailed Analysis of Article 50 of the EU's Artificial Intelligence Act' (14 juni 2024). Te verschijnen in: C.N. Pehlivan, N. Forgó & P. Valcke (red.), *The EU Artificial Intelligence (AI) Act: A Commentary*, Kluwer Law International, beschikbaar via SSRN: https://ssrn.com/abstract=4865427 of http://dx.doi.org/10.2139/ssrn.4865427.

50 Art. 50 lid 2 AI-verordening.

doi: 10.5553/NtER/138241202024030005004

dus uitkijken dat de wens om Europese burgers te beschermen innovatie (of eigenlijk eerder: een broodnodige inhaalslag) binnen de EU-grenzen in de weg zou staan.[51] Uiteindelijk is ervoor gekozen één risicocategorie toe te voegen die samenvalt met een nieuw toegevoegd reguleringsobject, namelijk de al genoemde AI-modellen voor algemene doeleinden (*general purpose AI models* oftewel *GPAI*). Binnen deze categorie gelden vervolgens strengere regels voor modellen die een systeemrisico[52] met zich meedragen, zoals primair maar niet uitsluitend afgemeten aan de rekenkracht die ze bezitten. Deze moeten behalve aan transparantieverplichtingen[53] ook aan veiligheidseisen[54] voldoen. De concessie is vervolgens wel dat er uitzonderingen gelden voor *free and open-source* modellen – nu net waar sommige van de Europese initiatieven rond generatieve AI op inzetten.[55]

De in de vorige paragraaf genoemde watermerkverplichting volstaat niet voor generatieve AI, aangezien output hiervan lang niet altijd door natuurlijke personen verkregen en gebruikt zal worden. In toenemende mate zal het gespecialiseerde software zijn die gebruikmaakt van de generatieve kracht van de modellen. Ook is in de hoofdstukken III en IV altijd sprake van een 'systeem', terwijl bij generatieve AI al op het niveau van de onderliggende modellen risico's spelen. Deze modellen 'zijn weliswaar een essentieel onderdeel van AI-systemen, maar vormen geen AI-systemen op zich'.[56] Om een AI-systeem te worden dienen 'nog andere componenten [te] worden toegevoegd, bijvoorbeeld een gebruikersinterface'.[57] AI-modellen voor algemene doeleinden 'kunnen op verschillende manieren in de handel worden gebracht, onder meer via bibliotheken en applicatie-programma-interfaces (API's), of als rechtstreekse download of fysieke kopie'[58] en eventueel 'nader worden gewijzigd of verfijnd tot nieuwe modellen'.[59]

Om zo'n model te kunnen doorlichten en valideren, op een vergelijkbare manier als dat met de hoogrisicosystemen van hoofdstuk III gaat gebeuren, is een mate van inzicht nodig, die ook voor de ontwikkelaars illusoir is. Hoe zou een *GPAI*-model, waarvan de trainingsdata uit het hele internet bij elkaar zijn geschraapt, bijvoorbeeld betekenisvol kunnen voldoen aan de eis dat deze 'zoveel mogelijk foutenvrij'[60] zijn? Een 'voldoende gedetailleerde samenvatting opstellen en openbaar maken over de voor het trainen van het AI-model voor algemene doeleinden gebruikte content'[61] is wél haalbaar, maar een veel minder strenge norm. Verbieden van deze grote generatie modellen zou echter verstrekkende gevolgen hebben voor de Europese economie, samenleving en wetenschap en was geen reële optie.

Het gegeven dat de Europese wetgever heeft gekozen voor een aparte categorie voor GPAI betekent niet dat er geen enkele verbinding is met andere risicocategorieën. De verordening kent ook 'AI-systemen voor algemene doeleinden', die weer 'gebaseerd [zijn] op een AI-model voor algemene doeleinden en dat verschillende doeleinden kan dienen, zowel voor direct gebruik als voor integratie in andere AI-systemen'.[62] De preambule gaat ook expliciet in op de mogelijkheid dat 'AI-systemen voor algemene doeleinden kunnen worden gebruikt (…) als component van AI-systemen met een hoog risico'.[63] De introductie van de nieuwe GPAI-categorie heeft de verordening er niet overzichtelijker op gemaakt, nu deze als een nieuw sausje over de volledige piramide wordt uitgegoten.

De transparantieverplichtingen en veiligheidseisen uit de afdelingen 2 en 3 van hoofdstuk V, die dus mede van toepassing kunnen zijn op aanbieders van GPAI-systemen, zullen zich uitkristalliseren in geharmoniseerde normen.[64] Daarnaast wil men gaan werken met praktijkcodes op Unieniveau voor AI-modellen voor algemene doeleinden met een systeemrisico en die tot het moment waarop de geharmoniseerde normen zijn vastgesteld, kunnen dienen om de naleving van de verplichtingen die gelden voor GPAI-modellen aan te tonen.[65] Het proces van totstandkoming van deze praktijkcodes wordt gecoördineerd door het AI-bureau en de AI-board. Voor de ontwikkeling van de eerste praktijkcode, die zal gaan over onderwerpen als transparantie, auteursrecht, risico-identificatie en risicobeheersing, is onlangs het consultatieproces gestart.[66]

2.6 Maatregelen ter ondersteuning van innovatie (hoofdstuk VI)

Hoofdstuk VI bevat regels over de zogeheten 'AI-testomgevingen voor regelgeving', beter bekend onder de Engelse benaming '*AI regulatory sandboxes*'. In afwijking van het Commissievoorstel moeten lidstaten ervoor zorgen 'dat hun bevoegde autoriteiten ten minste één AI-testomgeving voor regelgeving op nationaal niveau opzetten, die uiterlijk op 2 augustus 2026 operationeel is'. Aan deze verplichting kan ook in gezamenlijkheid met de bevoegde autoriteiten van een of meer andere lidstaten worden voldaan en het mag ook gaan om deel-

51 Open letter to the representatives of the European Commission, the European Council and the European Parliament, https://drive.google.com/file/d/1wrtxfvcD9FwfNfWGDL37Q6Nd8wBKXCkn/view, juni 2023.
52 Art. 3 lid 63 AI-verordening definieert 'systeemrisico' als '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'.
53 Art. 53 AI-verordening.
54 Art. 55 AI-verordening.
55 Art. 53 lid 2 AI-verordening.
56 Overweging 97 AI-verordening.
57 Overweging 97 AI-verordening.
58 Overweging 97 AI-verordening.
59 Overweging 97 AI-verordening.
60 Art. 10 lid 3 AI-verordening.
61 Art. 53 lid 1 onder d AI-verordening.
62 Art. 3 lid 66 AI-verordening.
63 Overweging 85 AI-verordening.
64 Art. 40 AI-verordening.
65 Art. 56 AI-verordening en art. 53 lid 4 AI-verordening. Deze codes moeten overigens onderscheiden worden van de gedragscodes die in art. 95 genoemd worden en de vrijwillige toepassing van eisen op andere dan hoogrisicosystemen mogelijk moeten maken.
66 Europese Commissie start consultatie over de praktijkgids voor AI, nieuwsbericht 6 augustus 2024, https://ecer.minbuza.nl.

doi: 10.5553/NtER/138241202024030005004

name aan een bestaande testomgeving. Daarnaast is na veel, ook door Nederland aangezwengelde, discussie in het onderhandelingsproces het testen van AI-systemen met een hoog risico onder reële omstandigheden, ook buiten AI-testomgevingen voor regelgeving, onder strikte voorwaarden mogelijk gemaakt.[67]

De Nederlandse vertaling van *regulatory sandbox* als 'testomgeving voor de regelgeving' lijkt wat ongelukkig. Of het aan deze vertaling ligt, is onzeker, maar het valt op dat in Nederland de 'sandbox', zoals die momenteel wordt voorbereid en getest door de Autoriteit Persoonsgegevens (AP), de Rijksinspectie Digitale Infrastructuur (RDI) en het Ministerie van Economische Zaken vooral wordt gezien als een loket om *compliance*-vragen over de AI-verordening op te lossen, in plaats van als een proeftuin waarin innovatieve AI-systemen kunnen worden getest.[68]

2.7 Governance, markttoezicht en sancties (hoofdstuk VII e.v.)

In aanvulling op de eisen die aan verschillende AI-systemen worden gesteld, bevat de AI-verordening ook een uitgebreid hoofdstuk over governance, met een onderscheid tussen Unieniveau en nationaal niveau. Op Unieniveau voorziet de AI-verordening in de oprichting van een aantal nieuwe instanties, waaronder het AI-bureau (als onderdeel van de Commissie),[69] de AI-board (bestaande uit één vertegenwoordiger per lidstaat),[70] een adviesforum[71] en een wetenschappelijk panel van onafhankelijke deskundigen (specifiek ter ondersteuning van de handhavingsactiviteiten in het kader van de AI-verordening).[72] Op nationaal niveau moeten lidstaten ten minste één aanmeldende autoriteit en ten minste één markttoezichtautoriteit aanwijzen.[73] De rol van de aanmeldende autoriteit is specifiek beperkt tot AI-systemen met een hoog risico waarvoor een 'externe' conformiteitsbeoordeling nodig is; deze autoriteit is verantwoordelijk voor het opzetten en uitvoeren van procedures voor de beoordeling, aanwijzing en aanmelding van conformiteitsbeoordelingsinstanties en voor het toezicht hierop.[74] Het is in het bijzonder de taak van deze instantie[75] om conformiteitsbeoordelingsinstanties die voldoen aan de eisen die de AI-verordening hieraan stelt,[76] aan te melden bij de Commissie en bij andere lidstaten. De rol van de markttoezichtautoriteit strekt zich daarentegen uit tot handhaving van de AI-verordening als geheel en wordt nader uitgewerkt in hoofdstuk IX inzake markttoezicht.[77] Dit hoofdstuk maakt niet alleen duidelijk dat het markttoezicht op AI-systemen is ingebed in het algemene stelsel van markttoezicht zoals dat voortvloeit uit het nieuwe wetgevingskader,[78] maar ook dat markttoezicht op AI-systemen specifiek hierop toegesneden bevoegdheden van markttoezichthouders vergt. Zo kan een markttoezichthouder die voldoende reden heeft om van mening te zijn dat een AI-systeem een 'risico' vormt, een evaluatie verrichten van het betrokken AI-systeem ten aanzien van de overeenstemming ervan met alle eisen en verplichtingen van deze verordening. Dit risico kan betrekking hebben op verboden AI-systemen (met een ontoelaatbaar risico), AI-systemen met een hoog risico en AI-systemen waarvoor enkel een transparantieverplichting geldt.[79] Wanneer een AI-systeem niet aan de toepasselijke eisen en verplichtingen van de verordening voldoet, verplicht de markttoezichthouder de betrokken operator om zonder onnodige vertraging passende corrigerende maatregelen te nemen om het AI-systeem binnen een bepaalde termijn hiermee in overeenstemming te brengen.[80] Bovendien heeft elke natuurlijke of rechtspersoon die redenen heeft om van mening te zijn dat er inbreuk is gepleegd op de bepalingen van de AI-verordening, het recht om een klacht in te dienen bij deze markttoezichtautoriteit.[81] De markttoezichtautoriteit zal doorgaans ook een rol vervullen bij de oplegging van sancties, waaronder 'administratieve geldboeten', vanwege inbreuken op de AI-verordening (hoofdstuk XII),[82] maar het is aan de lidstaten om tijdig, dat wil zeggen uiterlijk 2 augustus 2025, de noodzakelijke voorschriften hiervoor vast te stellen.

Een voorbeeld van hoe de waardeketenaanpak ook doorklinkt in de handhaving is gelegen in de toevoeging dat een nieuwe, specifieke categorie aanbieder, namelijk 'aanbieder verder in de AI-waardeketen'[83] een rol krijgt in de handhaving van de AI-verordening doordat deze het recht heeft een klacht in te dienen wegens inbreuk

67 Art. 60 AI-verordening.

68 Autoriteit Persoonsgegevens (Directie Coördinatie Algoritmes (DCA)), Rapportage AI- & Algoritmerisico's Nederland, editie 3, zomer 2024, p. 47, www.autoriteitpersoonsgegevens.nl.

69 Art. 64 AI-verordening.

70 Art. 65 AI-verordening. Taak van de AI-board is om de Commissie en de lidstaten te adviseren en assisteren teneinde de consistente en doeltreffende toepassing van de AI-verordening te vergemakkelijken (art. 66 AI-verordening).

71 Art. 67 AI-verordening. NB De samenstelling van het adviesforum vertegenwoordigt een evenwichtige selectie van belanghebbenden, waaronder het bedrijfsleven, start-ups, kleine en middelgrote ondernemingen (kmo's), het maatschappelijk middenveld en de academische wereld. Bij de samenstelling van het adviesforum wordt een evenwicht in acht genomen tussen commerciële en niet-commerciële belangen en, binnen de categorie commerciële belangen, tussen kmo's en andere ondernemingen.

72 Art. 68 AI-verordening.

73 Art. 69 AI-verordening.

74 Art. 28 AI-verordening e.v.

75 Dit kan een accreditatie-instelling zijn in de zin van Verordening (EG) nr. 765/2008 van het Europees Parlement en de Raad van 9 juli 2008 tot vaststelling van de eisen inzake accreditatie en markttoezicht betreffende het verhandelen van producten (*PbEU* 2008, L 2018/30-47).

76 Zie art. 31 AI-verordening. Zo moeten aangemelde instanties onafhankelijk zijn van de aanbieder van een AI-systeem met een hoog risico met betrekking waartoe de aanbieder conformiteitsbeoordelingsactiviteiten verricht.

77 Art. 74 AI-verordening e.v.

78 Art. 74 lid 1 AI-verordening: 'Verordening (EU) 2019/1020 is van toepassing op AI-systemen die onder deze verordening vallen.'

79 Zie specifiek art. 79 lid 6 AI-verordening.

80 Art. 79 lid 2 AI-verordening.

81 Art. 85 AI-verordening.

82 Art. 99 AI-verordening.

83 Gedefinieerd als 'een aanbieder van een AI-systeem, met inbegrip van een AI-systeem voor algemene doeleinden, waarin een AI-model is geïntegreerd, ongeacht of het AI-model door hemzelf wordt verstrekt en verticaal geïntegreerd is of door een andere entiteit wordt aangeboden op basis van contractuele betrekkingen', art. 3 lid 68 AI-verordening.

op deze verordening.[84] Daarnaast moeten aanbieders van AI-modellen voor algemene doeleinden voor deze doelgroep specifieke technische documentatie opstellen.[85]

3. De AI-verordening in een veellagige rechtsorde

3.1 Opdracht aan nationale wetgever

Hoewel de AI-verordening verbindend is in al haar onderdelen en rechtstreeks toepasselijk in elke lidstaat,[86] betekent de keuze voor een verordening als wetgevingsinstrument niet dat lidstaten volledig buitenspel staan als het gaat om de uitvoering hiervan. Nog los van de te verwachten uitvoeringswetgeving en aanpassingen van bestaande wetgeving, wijst de AI-verordening op onderdelen juist nadrukkelijk naar de lidstaat of naar bevoegde instanties binnen deze lidstaat om een adequate toepassing van de AI-verordening te verzekeren. Zo moeten, zoals hiervoor is besproken, de lidstaten ervoor zorgen dat hun bevoegde autoriteiten ten minste één AI-testomgeving voor regelgeving op nationaal niveau opzetten.[87] Het is daarom van belang om juist in de huidige periode tussen de inwerkingtreding van de AI-verordening (1 augustus 2024) en de toepassing van de meeste inhoudelijke bepalingen (2 augustus 2026)[88] te zien welke gevolgen de AI-verordening heeft voor lidstaten in het algemeen en voor de Nederlandse rechtsorde in het bijzonder.

Voorop staat dat de AI-verordening geharmoniseerde regels stelt voor het in de handel brengen, in gebruik stellen en gebruiken van AI-systemen in de Unie.[89] Daarmee wordt beoogd het vrije verkeer van op AI gebaseerde goederen en diensten te waarborgen, zodat lidstaten geen beperkingen kunnen opleggen aan de ontwikkeling, het in de handel brengen en het gebruik van AI-systemen, tenzij dat door deze verordening uitdrukkelijk wordt toegestaan.[90] In deze overweging klinkt sterk artikel 114 VWEU (verwezenlijking van de interne markt) als Verdragsbasis door. Intussen roept die grondslag wel de vraag op welke ruimte de AI-verordening zelf aan lidstaten biedt om regels te stellen ten aanzien van AI-systemen.

Het algemene beeld is dat lidstaten ten aanzien van de normstelling een tamelijk volgende rol hebben. De AI-verordening schrijft een risicogebaseerde typologie van AI-systemen voor en bepaalt in dat verband welke eisen voor de verschillende AI-systemen gelden. Bij de inhoud van die eisen, zeker als het gaat om AI-systemen met een hoog risico, leunt de AI-verordening sterk op de invulling die hieraan wordt gegeven door normalisatie-instellingen. Daarmee beoogt de AI-verordening de industrie 'mee' te krijgen in haar poging om een ingewikkelde en zich rap ontwikkelende technologie te reguleren. Omgekeerd komen lidstaten als gevolg van die benadering mogelijk nog meer aan de zijlijn van de regulering te staan, aangezien geharmoniseerde normen die binnen normalisatie-instellingen worden ontwikkeld, een vermoeden van overeenstemming bevatten ten aanzien van de eisen die in de AI verordening zijn opgenomen. De eigenlijke harmonisatie vindt dus plaats bij de ontwikkeling van deze geharmoniseerde normen.

Inmiddels heeft de Commissie aan Europese normalisatie-instellingen de opdracht gegeven om de eerste geharmoniseerde standaarden te ontwikkelen.[91] Het is de bedoeling dat die in het voorjaar van 2025 gereed zijn. Lidstaten staan echter niet volledig buitenspel bij dit proces van Europese normalisatie, dat wel is getypeerd als de privatisering van AI regulering.[92] Allereerst kunnen lidstaten indien nodig overheden, met inbegrip van markttoezichtautoriteiten, ertoe aansporen om deel te nemen aan normalisatieactiviteiten die de ontwikkeling beogen van geharmoniseerde normen op Europees niveau, dus om betrokken te zijn bij het proces van totstandkoming van de geharmoniseerde norm.[93] Daarnaast kan een lidstaat – na totstandkoming van een geharmoniseerde norm – indien hij van mening is dat een geharmoniseerde norm niet volledig beantwoordt aan de beoogde eisen die beschreven zijn in de desbetreffende harmonisatiewetgeving van de Unie, de Commissie daarvan op de hoogte brengen met een gedetailleerde toelichting.[94]

De AI-verordening sluit niet uit dat lidstaten op een aantal onderdelen eigen accenten leggen in de regulering van AI. Allereerst zijn er onderdelen die buiten de reikwijdte van de AI-verordening vallen. Zo stelt de AI-verordening buiten twijfel dat zij geen afbreuk doet aan de bevoegdheden van de lidstaten op het gebied van defensie en nationale veiligheid, waarmee dit domein buiten het bereik van de AI-verordening valt.[95] Daarnaast staat de AI-verordening niet eraan in de weg dat lidstaten wettelijke of bestuursrechtelijke bepalingen handhaven of invoeren die gunstiger zijn voor werknemers wat betreft de bescherming van hun rechten met betrekking tot het gebruik van AI-systemen door werkgevers.[96] Voor zover een AI-systeem wel binnen de

84 Art. 89 lid 2 AI-verordening.
85 Art. 53 lid 1 punt b AI-verordening en Bijlage XII.
86 Vgl. art. 288 VWEU.
87 Art. 57 AI-verordening.
88 Hierop zijn enkele uitzonderingen (art. 113 AI-verordening), waaronder aanmeldende autoriteiten, governance en sancties.
89 Art. 1 lid 2 AI-verordening.
90 Overweging 1 AI-verordening.
91 Uitvoeringsbeschikking van de Commissie van 22 mei 'on a standardisation request to the European Committee for Standardisation and the European Committee for Electrotechnical Standardisation in support of Union policy on artificial intelligence', C(2023)3215 final.
92 S. de Vries, O. Kanevskaia & R. de Jager, 'Internal Market 3.0: The Old "New Approach" for Harmonising AI Regulation', www.europeanpapers.eu.
93 Art. 7 Verordening (EU) nr. 1025/2012 van het Europees Parlement en de Raad van 25 oktober 2012 betreffende Europese normalisatie (*PbEU* 2012, L 316). Zie in dit verband Normcommissie, Artificial Intelligence en Big Data, www.nen.nl.
94 Art. 11 lid 1 Verordening (EU) nr. 1025/2012. Hetzelfde geldt bij de vaststelling van gemeenschappelijke specificaties door de Commissie (art. 41 lid 6 AI-verordening).
95 Art. 2 lid 3 AI-verordening.
96 Art. 2 lid 11 AI-verordening.

doi: 10.5553/NtER/138241202024030005004

reikwijdte van de verordening valt, is van belang welke ruimte per categorie van AI-systeem aan lidstaten wordt geboden.[97] Ten aanzien van verboden AI-systemen biedt de AI-verordening enige ruimte aan lidstaten om met het oog op de rechtshandhaving te voorzien in de mogelijkheid om volledig of gedeeltelijk toestemming te verlenen voor het gebruik van systemen voor biometrische identificatie op afstand in *real time* in openbare ruimten.[98] Verder stelt de AI-verordening buiten twijfel dat de verplichtingen die ten aanzien van AI-systemen met een hoog risico op 'gebruiksverantwoordelijken' (dus niet op aanbieders) rusten, bijvoorbeeld met betrekking tot menselijk toezicht,[99] geen afbreuk doen aan andere verplichtingen van gebruiksverantwoordelijken op grond van het Unierecht of nationaal recht.[100] Hier kan het nationale recht dus aanvullende eisen stellen. Verplichtingen ten aanzien van de aanbieders van AI-systemen lijken daarentegen volledig geharmoniseerd, hoewel de exacte strekking van deze verplichtingen afhangt van de wijze waarop geharmoniseerde standaarden worden vormgegeven, 'rekening houdend met de stand van de technologie'. Interessant is met name in hoeverre lidstaten aan andere AI-systemen (met een 'laag risico') ook andere verplichtingen dan transparantieverplichtingen kunnen opleggen. De AI-verordening stelt namelijk buiten twijfel dat transparantieverplichtingen voor deze AI-systemen zijn geharmoniseerd,[101] maar laat daarmee open of ook andere verplichtingen aan deze AI-systemen kunnen worden verbonden.

Al met al lijkt de nationale wetgever ten aanzien van de normstelling in de AI-verordening een grotendeels volgende rol te hebben. In het kader van de handhaving van de AI-verordening is daarentegen een grotere rol voor nationale instanties weggelegd. Allereerst zal moeten worden voorzien in de aanwijzing van een of meer markttoezichthouders, die ook kunnen worden belast met de oplegging van sancties. Momenteel lijkt binnen Nederland de Autoriteit Persoonsgegevens (AP) daarvoor de beste papieren te hebben, aangezien zij zich reeds profileert als 'coördinerend toezichthouder op algoritmes en AI'.[102] Tegelijk blijft ook een rol voor andere markttoezichthouders weggelegd, met name bij de zogeheten afhankelijke AI-systemen met een hoog risico, omdat de gereguleerde producten waarvan het AI-systeem een veiligheidscomponent vormt, al onderworpen zijn aan sectoraal markttoezicht.[103]

In algemenere zin roept de inwerkingtreding van de AI-verordening de vraag op hoe de afbakening tussen AI-systemen en andere algoritmische systemen (kortweg: algoritmen) zal plaatsvinden. Bepalend voor de toepasselijkheid van de AI-verordening is namelijk allereerst of sprake is van een AI-systeem. Met de hiervoor besproken definitie van AI-systemen in de verordening die (inmiddels) de nadruk legt op het inferentievermogen hiervan wordt beoogd AI-systemen te onderscheiden van andere algoritmes. In de Nederlandse rechtspraktijk is echter zichtbaar dat het onderscheid tussen algoritmes en AI-systemen niet altijd scherp wordt gemaakt.[104] Bovendien wordt het onderscheid in risicocategorieën dat de AI-verordening ten aanzien van AI-systemen maakt, inmiddels ook buiten de context van AI-systemen toegepast. Daarmee is niet uitgesloten dat de eisen die de AI-verordening bevat ten aanzien van AI-systemen met hoog risico, niet alleen hun weg vinden naar andere AI-systemen,[105] maar ook dat deze eisen uit de AI-verordening als 'nuttige proeftuin' hun weg vinden naar andere algoritmische toepassingen.[106]

3.2 De AI-verordening binnen de mondiale reguleringspolitiek

Een extra uitdaging voor de nationale wetgever is dat de ontwikkelingen rond de AI-verordening zich niet in een isolement voltrekken. Ten tijde van de publicatie van het Commissievoorstel voor een AI-verordening schreven wij al dat de Europese wetgever voorsorteert op een rol als *global regulatory leader* op het vlak van AI. Als 'bedrijven wereldwijd hun AI-producten op de Europese standaarden zullen afstemmen, gezien de grootte van de Europese interne markt',[107] kunnen we spreken van een *Brussels effect*,[108] waardoor de EU haar gebrek aan technische en economische invloed op de ontwikkeling van AI zou kunnen compenseren. In hoeverre dit *Brussels effect* zich zal gaan voordoen bij digitale producten, wordt betwijfeld en hangt in elk geval af van de aanwezigheid van alternatieve reguleringsstandaarden.[109]

De concreetste ontwikkeling op het gebied van mondiale regulering van AI is dat, vrijwel gelijktijdig met de AI-verordening, binnen de Raad van Europa (RvE) het Kaderverdrag over artificiële intelligentie en de mensenrechten, de democratie en de rechtsstaat tot stand is

97 Zie ook M. Veale & F. Zuiderveen Borgesius, 'Demystifying the Draft EU Artificial Intelligence Act', *Computer Law Review International* 2021, afl. 4, p. 97-112.

98 Art. 5 lid 5 AI-verordening.

99 Op grond van art. 26 lid 2 AI-verordening moeten gebruiksverantwoordelijken het menselijk toezicht opdragen aan natuurlijke personen die over de nodige bekwaamheid, opleiding en autoriteit beschikken en de nodige ondersteuning krijgen.

100 Art. 26 lid 3 AI-verordening. Overweging 63 benadrukt in dit verband dat eventueel nationaal recht uiteraard wel verenigbaar moet zijn met het Unierecht, zoals met betrekking tot de bescherming van de persoonsgegevens.

101 Art. 1 lid 2 onder d AI-verordening.

102 Zie AP, 'Algoritmes & AI', www.autoriteitpersoonsgegevens.nl. In de technische briefing (zie voetnoot 42) kwam wel nadrukkelijk aan de orde dat Nederland zich in de onderhandelingen ervoor heeft ingespannen dat in de tekst 'markttoezichthouders' in het meervoud kwam te staan.

103 Zie uitgebreider hierover: AP, 'AP en RDI: Toezicht op AI-systemen vraagt samenwerking en moet snel geregeld worden', 11 juni 2024, www.autoriteitpersoonsgegevens.nl.

104 Zie bovengenoemd rapport van de AP en het rapport van de Rekenkamer Amsterdam.

105 Zie expliciet hierover art. 95 AI-verordening.

106 Zie voor deze uitdrukking: Raad van State, 'Digitalisering. Wetgeving en bestuursrechtspraak', Den Haag 2021, p. 66.

107 Meuwese & Wolswinkel 2022, p. 93.

108 A. Bradford, *The Brussels Effect: How the European Union Rules the World*, Oxford: Oxford University Press 2020.

109 A. Bradford, *Digital Empires. The Global Battle to Regulate Technology*, Oxford: Oxford University Press 2023; The AI Act in perspective, The Tech Brief, podcast 26 januari 2024.

gekomen (hierna: AI-kaderverdrag).[110] Waar de AI-verordening het moet hebben van voldoende handhavingscapaciteit bij de eigen lidstaten om aanbieders in een derde land te bereiken en een mogelijk maar onzeker *Brussels effect* via vrijwillige naleving door de private sector wereldwijd, is de hoop voor het AI-kaderverdrag gevestigd op een groot aantal ondertekenaars buiten Europa. Op de dag van openstelling van de ondertekening van het AI-kaderverdrag, op 5 september 2024 in Vilnius, waren dat Andorra, Georgië, IJsland, Noorwegen, Moldavië, San Marino, het Verenigd Koninkrijk en de Europese Unie, maar ook Israël en de Verenigde Staten van Amerika.[111]

De Europese Commissie heeft er tijdens de onderhandelingen alles aan gedaan om ervoor te zorgen dat de AI-verordening en het AI-kaderverdrag inhoudelijk op elkaar aansluiten. Zo moest, om een voorbeeld te geven, volgens een conceptversie van de verdragstekst elke partij binnen haar nationale rechtssysteem de mogelijkheid hebben om een verbod in te stellen op bepaalde toepassingen van AI-systemen.[112] Deze benadering sloot niet goed aan op het uitgangspunt van uniforme toepassing van de regels van de AI-verordening, die 'over het algemeen gebaseerd zijn op volledige harmonisatie',[113] waardoor EU-lidstaten bij voorbaat al niet aan deze verdragsverplichting zouden kunnen voldoen. In de definitieve tekst staat daarom alleen dat verdragspartijen de noodzaak van een verbod moeten onderzoeken.[114] Ook heeft de Raad van de Europese Unie onlangs besloten om namens alle lidstaten het AI-kaderverdrag te ondertekenen, aangezien '[h]et verdrag […] in de Unie uitsluitend door middel van Verordening (EU) 2024/1689 en ander toepasselijk acquis van de Unie uitgevoerd [wordt]'.[115]

Door deze positionering van de AI-verordening als hét vehikel voor naleving van het AI-kaderverdrag binnen de Europese Unie, hoopt de EU-wetgever ongetwijfeld dat de slagkracht van deze verordening vergroot wordt. Enerzijds wordt AI hiermee binnen de EU, duidelijker dan in de preambule van de verordening zelf en stelliger dan de belangrijkste verdragsbasis van artikel 114 VWEU rechtvaardigt, neergezet als EU-bevoegdheid. Anderzijds verhoogt deze benadering de kans dat derden-verdragspartijen bij het vervullen van hun verplichtingen op grond van het AI-kaderverdrag, naar de AI-verordening zullen kijken als een blauwdruk voor specifiekere AI-regulering die in dat kaderverdrag zelf ontbreekt.

110 Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, CETS No. 225.

111 https://www.coe.int/en/web/artificial-intelligence/cai#:~:text=With%20the%20signature%20in%20Vilnius%20on%205%20September.

112 Art. 24 lid 3 Revised Zero Draft [Framework] Convention, Straatsburg, 6 januari 2023.

113 Besluit (EU) 2024/2218 van de Raad van 28 augustus 2024 inzake de ondertekening namens de Europese Unie van het Kaderverdrag van de Raad van Europa over artificiële intelligentie en de mensenrechten, de democratie en de rechtsstaat, overweging 3.

114 Art. 16 lid 4 AI-kaderverdrag.

115 Besluit (EU) 2024/2218 van de Raad van 28 augustus 2024, overweging 3.

4. Slotbeschouwing

De sterke druk om een vuist te maken tegen de AI-activiteiten vanuit de hoek van Big Tech en om over de hele linie strengere regels te hanteren, is bepalend geweest in de onderhandelingen over de AI-verordening de afgelopen drie jaar. Wel is er op het laatste moment een tegenbeweging geweest, gedreven door de wens om toch nog een Europese AI-industrie op te tuigen. Dit heeft zich vertaald in een uitbreiding en verfijning van zowel de 'risicopiramide' als bepaalde algemene en procedurele verplichtingen waarbij de grenzen in zicht komen van wat nog in een reguleringskader rond productveiligheid kan worden ingepast. Zo bevat de verordening een nieuwe inspanningsverplichting, voor aanbieders en gebruiksverantwoordelijken van álle AI-systemen, om hun personeel en relevante derden 'AI-geletterdheid' bij te brengen.[116] Ook het vastleggen van een algemeen individueel klachtrecht en een recht op uitleg bij individuele besluitvorming[117] passen meer bij een AVG[118]-achtige reguleringsaanpak, gericht op de rechten van individueel getroffenen, dan bij productveiligheidsregulering.[119] Bovendien verschuift door deze aan het Commissievoorstel toegevoegde uitbreidingen de voor productveiligheidsregulering kenmerkende oriëntatie op het aanbieden van producten naar het gebruiken van deze 'producten', omdat deze nieuwe verplichtingen zich richten op de 'gebruiksverantwoordelijken'. Deze uitbreidingen moeten daarom gezien worden in het licht van de aanscherping van de 'AI-waardeketenaanpak', met als doel een 'eerlijke verdeling van verantwoordelijkheden', waarvan de hele verordening in haar definitieve hoedanigheid doordrenkt is, en met als gevolg een minder eenzijdige reguleringsfocus op het aanbieden en op de markt brengen van AI-systemen.

Ook zien we dat de uitruil van aanscherpingen en uitzonderingen leidt tot complexiteit. De aanpassingen die volgden uit de inhoudelijke discussie over het gewenste beschermingsniveau, ook op het vlak van *narrow AI*, zijn gepaard gegaan met nieuwe uitzonderingen. Het Europees Parlement heeft meerdere aanscherpingen binnengehaald, zowel in de categorie 'verboden toepassingen' als in de lijst met *use cases* voor hoogrisicotoepassingen. Zoals wel vaker gebeurt, zijn moeilijke punten in de onderhandelingen platgeslagen door uitgebreide clausuleringen. Zo mag de Europese Commissie wel door middel van gedelegeerde handelingen de cruciale Annex III inzake *use cases* amenderen, maar alleen als aan een lange lijst procedurele en inhoudelijke criteria is voldaan.[120]

Het resultaat is een complex geheel dat minder goed als een piramide gevisualiseerd kan worden dan het geval

116 Art. 4 AI-verordening.

117 Art. 85 en 86 AI-verordening.

118 Algemene verordening gegevensbescherming.

119 Zie uitgebreider M. Almada & N. Petit, 'The EU AI Act: a medley of product safety and fundamental rights?', Robert Schuman Centre for Advanced Studies Research Paper 2023/59 (2023).

120 Art. 7 AI-verordening.

doi: 10.5553/NtER/138241202024030005004

was bij het Commissievoorstel uit 2021. De inpassing van GPAI draagt bij aan het verder uitdijen (of zelfs ondergraven) van de piramidestructuur in een mate die ervoor zorgt dat een andere metafoor de structuur van de verordening misschien beter verwoordt. Alleen de verboden toepassingen staan immers overeind als strak afgebakende categorie. Voor het overige geldt dat in de praktijk meerdere categorieën (meerdere hoofdstukken) van toepassing kunnen zijn, vanwege het feit dat AI-systemen onderdeel kunnen zijn van een ander product, GPAI-modellen een heel eigen rol hebben in de ontwikkel- en gebruiksketen, er meer nadruk ligt op veranderingen binnen producten door de tijd heen en bij het omschrijven van de hoogrisicotoepassingen veel meer de nuance is opgezocht. Daarmee staat de oorspronkelijke redenering, dat met de (noodgedwongen) keuze voor productveiligheidsregulering in ieder geval bereikt wordt dat de reguleringslasten bij de industrie terechtkomen en dan op een wijze waar men mee uit de voeten kan, op de tocht.

De verordening biedt niet meteen rechtszekerheid voor Nederlandse ontwikkelaars en gebruikers. Naast de hierboven genoemde complexiteit, komt dit doordat veel nog afhangt van de richtsnoeren van de Commissie, de geharmoniseerde normen en de praktijkcodes die allemaal nog in de maak zijn. Ook speelt de onduidelijkheid van de definitie, waarover wat ons betreft wel wat meer discussie zou mogen zijn, een rol. De neiging om deze definitie maar breed te interpreteren doet geen afbreuk aan de verwachting dat het Hof van Justitie vroeg of laat gevraagd zal worden zich hierover uit te spreken. Een '*better safe than sorry*'-benadering kan leiden tot over-*compliance*, een fenomeen dan eveneens is gezien bij de AVG, maar ook tot AI-vermijding, met name als gevolg van de onzekerheid die de rol van gebruiksverantwoordelijke met zich meebrengt. Dat zal niet het *Brussels effect* zijn dat met de AI-verordening is beoogd.

doi: 10.5553/NtER/138241202024030005004