Publicatieblad
van de Europese Unie

NL
C-serie

C/2026/6005

12.11.2026

MEDEDELING VAN DE COMMISSIE

Richtsnoeren voor de uitvoering van de transparantieverplichtingen voor bepaalde AI-systemen op grond van artikel 50 van Verordening (EU) 2024/1689 (AI-verordening)

(C/2026/6005)

Inhoudsopgave

Bladzijde

1. ACHTERGROND EN DOELSTELLINGEN 3
2. OVERZICHT VAN DE TRANSPARANTIEVERPLICHTINGEN EN HORIZONTALE THEMA'S 3
   - 2.1. Overzicht van de transparantieverplichtingen van artikel 50 van de AI-verordening 3
   - 2.2. Motivering en doelstellingen 4
   - 2.3. Verantwoordelijke actoren uit hoofde van artikel 50 van de AI-verordening 5
   - 2.4. Uitsluitingen van het toepassingsgebied van de AI-verordening 7
     - 2.4.1. Gebruiksverantwoordelijken die handelen in het kader van een louter persoonlijke niet-professionele activiteit 7
     - 2.4.2. Onderzoek & ontwikkeling 8
     - 2.4.3. AI die wordt vrijgegeven onder vrije en opensource licenties 8
   - 2.5. Wisselwerking met de verboden praktijken, de vereisten voor AI-systemen met een hoog risico en AI-geletterdheid 9
   - 2.6. Wisselwerking met regels die van toepassing zijn op AI-modellen en/of -systemen voor algemene doeleinden 9
3. ARTIKEL 50, LID 1, VAN DE AI-VERORDENING: TRANSPARANTIE VOOR INTERACTIEVE AI-SYSTEMEN 9
   - 3.1. Belangrijkste componenten, concepten en daarmee samenhangende transparantieverplichting(en) 10
     - 3.1.1. AI-systemen die voor directe interactie met natuurlijke personen bedoeld zijn 10
     - 3.1.2. Informatieverplichting uit hoofde van artikel 50, lid 1, van de AI-verordening 12
   - 3.2. Uitzonderingen op de informatieverplichting uit hoofde van artikel 50, lid 1, van de AI-verordening 14
     - 3.2.1. Uitzondering voor duidelijke interactie met een AI-systeem 14
     - 3.2.2. Uitzondering voor AI-systemen die bij wet zijn toegestaan met het oog op de rechtshandhaving 16
   - 3.3. Wisselwerking met andere rechtshandelingen van de Unie 17
4. ARTIKEL 50, LID 2, VAN DE AI-VERORDENING: MARKERING EN DETECTIE VAN DOOR AI GEGENEREERDE OF GEMANIPULEERDE CONTENT 18
   - 4.1. Belangrijkste componenten en concepten van artikel 50, lid 2, van de AI-verordening 18
     - 4.1.1. AI-systemen die synthetische content genereren of manipuleren 18
     - 4.1.2. Modaliteiten van synthetische content die binnen het toepassingsgebied vallen 19
     - 4.1.3. Content die buiten het toepassingsgebied van artikel 50, lid 2, van de AI-verordening valt 20
   - 4.2. De markerings- en detectieverplichtingen van artikel 50, lid 2, van de AI-verordening 21
     - 4.2.1. De markeringsverplichting 21
     - 4.2.2. De detectieverplichting 22
     - 4.2.3. Naleving van de eisen voor de technische oplossingen: doeltreffend, interoperabel, robuust en betrouwbaar 22
   - 4.3. Uitzonderingen op de verplichtingen op grond van artikel 50, lid 2, van de AI-verordening 24
   - 4.4. Wisselwerking met andere rechtshandelingen van de Unie 25

5. ARTIKEL 50, LID 3, VAN DE AI-VERORDENING: SYSTEMEN VOOR HET HERKENNEN VAN EMOTIES EN SYSTEMEN VOOR BIOMETRISCHE CATEGORISERING 26

5.1. Belangrijkste componenten, begrippen en daarmee samenhangende transparantieverplichting(en) van artikel 50, lid 3, van de AI-verordening 27

5.1.1. Het begrip “systeem voor het herkennen van emoties” 27

5.1.2. Het begrip “systeem voor biometrische categorisering” 27

5.1.3. De informatieverplichting uit hoofde van artikel 50, lid 3, van de AI-verordening 27

5.2. Uitzondering met het oog op de rechtshandhaving 28

5.3. Wisselwerking met andere rechtshandelingen van de Unie 28

6. ARTIKEL 50, LID 4: LABELEN VAN DEEPFAKES EN BEPAALDE TEKSTPUBLICATIES 28

6.1. Belangrijkste componenten, concepten en daarmee samenhangende transparantieverplichting(en) voor deepfakes op grond van artikel 50, lid 4, van de AI-verordening 28

6.1.1. Het begrip “deepfake” 29

6.1.2. De verplichting tot openbaarmaking van artikel 50, lid 4, eerste alinea, van de AI-verordening .. 31

6.1.3. Transparantie van artistieke, creatieve, satirische, fictieve of analoge deepfake-content 32

6.1.4. Uitzondering in verband met rechtshandhaving 34

6.1.5. Wisselwerking met andere rechtshandelingen van de Unie 34

6.2. Belangrijkste componenten, concepten en daarmee samenhangende transparantieverplichting(en) voor door AI gegenereerde of gemanipuleerde tekst uit hoofde van artikel 50, lid 4, van de AI-verordening .. 35

6.2.1. Tekst die wordt gepubliceerd om het publiek te informeren over aangelegenheden van algemeen belang 36

6.2.2. De verplichting tot openbaarmaking uit hoofde van artikel 50, lid 4, tweede alinea, van de AI-verordening 37

6.2.3. Uitzondering op de transparantieverplichting voor tekst die een proces van menselijke toetsing of redactionele controle heeft ondergaan en onder redactionele verantwoordelijkheid valt 37

6.2.4. Uitzondering in verband met rechtshandhaving 38

6.2.5. Wisselwerking met andere rechtshandelingen van de Unie 39

7. HORIZONTALE EISEN DIE VAN TOEPASSING ZIJN OP DE INFORMATIE DIE WORDT VERSTREKT OP GROND VAN ARTIKEL 50, LID 5, VAN DE AI-VERORDENING 39

7.1. Het verstrekken van informatie op duidelijke en te onderscheiden wijze 39

7.2. Eerste interactie of blootstelling 40

7.3. Naleving van de toepasselijke toegankelijkheidsvoorschriften 40

8. HANDHAVING VAN ARTIKEL 50 VAN DE AI-VERORDENING 40

8.1. Effecten van de naleving van een als toereikend beoordeelde praktijkcode 41

8.2. Markttoezichtautoriteiten 42

8.3. Sancties 42

8.4. Inwerkingtreding 42

9. HERZIENING EN ACTUALISERING VAN DE RICHTSNOEREN 43

1. **ACHTERGROND EN DOELSTELLINGEN**

(1) 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 bepaalde verordeningen (de AI-verordening) is op 1 augustus 2024 in werking getreden ([1]). Bij de AI-verordening worden geharmoniseerde regels vastgesteld voor het in de handel brengen, in gebruik stellen en gebruik van artificiële intelligentie (AI) in de Unie ([2]). Het doel ervan is innovatie en introductie van betrouwbare AI te bevorderen en te zorgen voor een hoge mate van bescherming van de gezondheid, de veiligheid en de grondrechten in de Unie, met inbegrip van de democratie en de rechtsstaat.

(2) In de AI-verordening wordt een risicogebaseerde aanpak gevolgd, waarbij AI-systemen worden ingedeeld in vier verschillende risicocategorieën, waarvan er één bestaat uit AI-systemen die transparantierisico's met zich meebrengen en die onderworpen zijn aan de verplichtingen van artikel 50 van de AI-verordening. Deze transparantieverplichtingen worden twee jaar na de inwerkingtreding van de AI-verordening van toepassing, d.w.z. met ingang van 2 augustus 2026.

(3) Overeenkomstig artikel 96, lid 1, punt d), van de AI-verordening worden deze richtsnoeren van de Commissie vastgesteld met als doel praktische richtsnoeren te bieden aan bevoegde autoriteiten, alsook aan aanbieders en gebruiksverantwoordelijken van AI-systemen, om consistent, doeltreffend, evenredig en uniform te waarborgen dat aan de transparantieverplichtingen uit hoofde van artikel 50 van de AI-verordening wordt voldaan.

(4) Bij het opstellen van deze richtsnoeren is gebruikgemaakt van de input van diverse belanghebbenden, verzameld tijdens een door de Commissie georganiseerde brede raadpleging, alsmede van de input van de lidstaten in de AI-board en van de input van belanghebbenden over de ontwerprichtsnoeren die ter raadpleging zijn gepubliceerd.

(5) Deze richtsnoeren zijn niet bindend. De bindende uitlegging van de AI-verordening komt uiteindelijk uitsluitend toe aan het Hof van Justitie van de Europese Unie ("HvJ-EU").

2. **OVERZICHT VAN DE TRANSPARANTIEVERPLICHTINGEN EN HORIZONTALE THEMA'S**

2.1. **Overzicht van de transparantieverplichtingen van artikel 50 van de AI-verordening**

(6) Artikel 50 van de AI-verordening bevat vier transparantieverplichtingen, die elk van toepassing zijn op verschillende soorten AI-systemen of de output ervan.

Bepaling	Soort AI-systeem/ output	Transparantieverplichting	Uitzonderingen of bijzondere regelingen
Artikel 50, lid 1	AI-systemen die direct interageren met natuurlijke personen	Aanbieders moeten het AI-systeem zodanig ontwikkelen en ontwerpen dat de betrokken natuurlijke personen worden geïnformeerd dat zij interageren met een AI-systeem.	Uitzonderingen gelden indien 1) de artificiële oorsprong van de interactie duidelijk is of 2) het systeem bij wet is toegestaan voor het opsporen, voorkomen, onderzoeken of vervolgen van strafbare feiten, tenzij het systeem voor het publiek beschikbaar is om een strafbaar feit te melden.

([1]) Verordening (EU) 2024/1689 van het Europees Parlement en de Raad van 13 juni 2024 tot vaststelling van geharmoniseerde regels betreffende artificiële intelligentie (verordening artificiële intelligentie) (PB L, 2024/1689, 12.7.2024, ELI: http://data.europa.eu/eli/reg/2024/1689/oj).

([2]) Artikel 1 van de AI-verordening.

Bepaling	Soort AI-systeem/ output	Transparantieverplichting	Uitzonderingen of bijzondere regelingen
Artikel 50, lid 2	AI-systemen die synthetische beeld-, video-, audio- of tekstinhoud genereren of manipuleren	Aanbieders moeten ervoor zorgen dat de output van het AI-systeem worden gemarkeerd in een machineleesbaar formaat en detecteerbaar zijn als kunstmatig gegenereerd of gemanipuleerd met behulp van technische oplossingen die doeltreffend, interoperabel, robuust en betrouwbaar zijn.	Uitzonderingen gelden indien 1) het AI-systeem een ondersteunende functie voor standaardbewerking vervult of de inputdata of de semantiek daarvan niet substantieel wijzigt of 2) het AI-systeem bij wet is toegestaan strafbare feiten op te sporen, te voorkomen, te onderzoeken of te vervolgen.
Artikel 50, lid 3	AI-systemen voor het herkennen van emoties of biometrische categorisering	Gebruiksverantwoordelijken moeten de natuurlijke personen die aan de werking van het AI-systeem worden blootgesteld, informeren.	Een uitzondering geldt indien het AI-systeem bij wet is toegestaan om strafbare feiten op te sporen, te voorkomen of te onderzoeken.
Artikel 50, lid 4	AI-systemen die deepfakes of teksten genereren of manipuleren die worden gepubliceerd om het publiek te informeren over aangelegenheden van algemeen belang	Gebruiksverantwoordelijken moeten bekendmaken dat de content kunstmatig is gegenereerd of gemanipuleerd.	Uitzonderingen gelden indien 1) het AI-systeem bij wet is toegestaan om strafbare feiten op te sporen, te voorkomen, te onderzoeken of te vervolgen of 2) de gepubliceerde tekst een proces van menselijke toetsing of redactionele controle heeft ondergaan en onder redactionele verantwoordelijkheid valt. Een speciale openbaarmakingsregeling is van toepassing op deepfakes die deel uitmaken van artistieke, creatieve, fictieve, satirische of analoge werken of programma's.

(7) In de punten 3 tot en met 6 van deze richtsnoeren wordt nader ingegaan op elk van deze vier verplichtingen. Daarnaast is in artikel 50, lid 5, van de AI-verordening bepaald dat de in de leden 1 tot en met 4 bedoelde informatie uiterlijk op het moment van de eerste interactie of blootstelling op duidelijke en te onderscheiden wijze aan de betrokken natuurlijke personen moet worden verstrekt en moet voldoen aan de toepasselijke toegankelijkheidsvoorschriften. In punt 7 worden deze eisen verduidelijkt.

(8) De verschillende transparantieverplichtingen die zijn vastgelegd in artikel 50 van de AI-verordening kunnen cumulatief van toepassing zijn op (de output van) één AI-systeem, waarbij mogelijk de verantwoordelijkheid van verschillende actoren (aanbieders of gebruiksverantwoordelijken) wordt betrokken. Als een AI-systeem bijvoorbeeld afbeeldingen genereert in het kader van een directe interactie met een natuurlijke persoon, dan moet de aanbieder van het systeem mogelijk voldoen aan zowel lid 1 als lid 2 van artikel 50 van de AI-verordening. Wanneer dat systeem ook beelden kan genereren die deepfakes of gepubliceerde teksten over aangelegenheden van algemeen belang vormen, moet de gebruiksverantwoordelijke mogelijk ook voldoen aan artikel 50, lid 4, van de AI-verordening.

2.2. **Motivering en doelstellingen**

(9) Zoals uiteengezet in de overwegingen 132 tot en met 136 van de AI-verordening, hebben de transparantieverplichtingen tot doel de risico's van imitatie, misleiding, desinformatie, manipulatie op grote schaal en fraude te verminderen en potentiële negatieve gevolgen voor democratische processen en het maatschappelijk vertrouwen als gevolg van door AI gegenereerde of gemanipuleerde content of interactie te beperken. Door gebruikers te informeren over de AI-oorsprong van interacties en content, kunnen ze onderbouwde beslissingen nemen, terwijl dat tevens bijdraagt aan het behoud van het vertrouwen en de integriteit van het informatie-ecosysteem. Transparantie is niet alleen een technische nalevingsvereiste, maar ook een voorwaarde om natuurlijke personen te laten herkennen dat er sprake is van door AI tot stand gebrachte interacties en content of van het gebruik van systemen voor het herkennen van emoties of biometrische categorisering. Door te zorgen voor een transparant en betrouwbaar gebruik van interactieve en generatieve AI-systemen kunnen tevens de persoonlijke autonomie en Europese waarden, menselijke waardigheid en andere grondrechten (bv. vrijheid van gedachte, privacy, gelijkheid en non-discriminatie, consumentenbescherming, intellectuele eigendom enz.) worden gewaarborgd, terwijl tegelijkertijd de vrijheid van meningsuiting en de geloofwaardigheid van mediacontent behouden blijven.

2.3. **Verantwoordelijke actoren uit hoofde van artikel 50 van de AI-verordening**

(10) Overeenkomstig artikel 3, punt 3, van de AI-verordening zijn **aanbieders** natuurlijke of rechtspersonen, overheidsinstanties, agentschappen of andere organen die AI-systemen ontwikkelen of laten ontwikkelen en die systemen in de Unie in de handel brengen of in gebruik stellen onder de eigen naam of merk, al dan niet tegen betaling. Aanbieders moeten voldoen aan de transparantieverplichtingen van artikel 50, leden 1 en 2, van de AI-verordening die op hen van toepassing zijn, ongeacht of ze in de Unie of in een derde land zijn gevestigd of zich daar bevinden ([3]). Ook aanbieders die buiten de Unie zijn gevestigd of zich daar bevinden, vallen onder de AI-verordening indien de output van hun AI-systeem in de Unie wordt gebruikt ([4]). Het incidenteel, onvoorzien of ongeoorloofd gebruik in een later stadium mag echter op zichzelf geen aanleiding vormen voor de toepassing van de verplichtingen op dergelijke aanbieders uit derde landen van interactieve of generatieve AI-systemen die niet in de Unie in de handel worden gebracht of in gebruik worden gesteld, onverminderd de bevoegdheden van de bevoegde markttoezichtautoriteiten om de output daarvan als niet in overeenstemming met de AI-verordening te beschouwen.

(11) Aanbieders moeten ervoor zorgen dat in het ontwerp en de werking van hun interactieve of generatieve AI-systemen transparantie is ingebouwd en dat uiterlijk wanneer die systemen in de handel worden gebracht of in gebruik worden gesteld, aan alle relevante transparantieverplichtingen van artikel 50, leden 1, 2 en 5, van de AI-verordening wordt voldaan. Alle natuurlijke personen die met het AI-systeem interageren of aan de output ervan worden blootgesteld, moeten gedurende de hele levenscyclus van het AI-systeem, zoals dat door die aanbieder in de handel wordt gebracht of in gebruik wordt gesteld, doeltreffend en consistent worden geïnformeerd (ongeacht wie de gebruiksverantwoordelijken of andere gebruikers van het systeem zijn).

> Een voorbeeld is een bedrijf dat onder zijn eigen naam of handelsmerk een generatieve of interactieve AI-toepassing (zoals een chatbot, beeldgenerator of AI-agent) in de Unie in de handel brengt voor consumenten, professionele gebruikers of andere juridische entiteiten, die het AI-systeem direct kunnen gebruiken voor interactie of het genereren van content (waaronder door het systeem in hun bestaande processen te integreren en het dienovereenkomstig te gebruiken zonder dat aanpassingen of aanvullende componenten nodig zijn). Het bedrijf is een aanbieder die verantwoordelijk is voor de naleving van de transparantieverplichtingen van artikel 50, lid 1 en/of 2, van de AI-verordening, ongeacht of het AI-systeem gratis of tegen betaling wordt aangeboden en ongeacht of de aanbieder in de Unie of in een derde land is gevestigd of zich daar bevindt.
>
> Een ander voorbeeld van een aanbieder is een onderneming of andere organisatie (bv. een overheidsinstantie) die intern een interactief AI-systeem (bv. chatbot) heeft ontwikkeld en het in de Unie in gebruik stelt voor eigen gebruik en onder zijn naam of handelsmerk.
>
> Indien een onderneming een reeds bestaand generatief AI-systeem, dat door een andere aanbieder in de handel is gebracht, aanpast (bv. met nieuwe trainingsdata) en dat systeem vervolgens onder haar eigen naam of handelsmerk in gebruik neemt, wordt die onderneming de aanbieder van het nieuwe systeem, zonder dat dit afbreuk doet aan de aansprakelijkheid van de aanbieder van het oorspronkelijke AI-systeem voor dat systeem.

(12) Gebruiksverantwoordelijken zijn natuurlijke of rechtspersonen, overheidsinstanties, agentschappen of andere organen die een AI-systeem onder eigen verantwoordelijkheid gebruiken, tenzij het AI-systeem wordt gebruikt in het kader van een persoonlijke niet-beroepsactiviteit ([5]). Met de "verantwoordelijkheid" voor een AI-systeem wordt bedoeld: het aanvaarden van de verantwoordelijkheid voor het besluit om het systeem te gebruiken en voor de wijze waarop het systeem daadwerkelijk wordt gebruikt (waaronder de output ervan). Dit vereist niet noodzakelijkerwijs technisch toezicht op de werking van het AI-systeem, mits de gebruiksverantwoordelijke beslist voor welke doeleinden en hoe het AI-systeem wordt ingezet (ook in gedecentraliseerde workflows en bedrijfsstructuren van een groep). Gebruiksverantwoordelijken die betrokken zijn bij complexe waardeketens voor de productie en distributie van content moeten evenredige maatregelen nemen om ervoor te zorgen dat de labeling van de content die ze overeenkomstig artikel 50, lid 4, van de AI-verordening hebben uitgevoerd, op een duidelijke en te onderscheiden wijze worden weergegeven voor de doelgroep en het te verwachten publiek op het moment van eerste blootstelling overeenkomstig artikel 50, lid 5, van de AI-verordening (bv. via contractuele voorwaarden met distributiepartners, instellingen voor gebruikerservaring (UX) en weer te geven interfaces).

([3]) Artikel 2, lid 1, punt a), en artikel 3, punten 3, 9 en 11, van de AI-verordening. Voor meer informatie over de begrippen "in de handel brengen", "in gebruik stellen" en "gebruik", zie de richtsnoeren van de Commissie betreffende verboden artificiële-intelligentiepraktijken (C(2025) 5052) en punt 2 van de mededeling van de Commissie – De Blauwe Gids van 2022: richtlijnen voor de uitvoering van de productvoorschriften van de EU (PB C 247 van 29.6.2022, blz. 1).

([4]) Artikel 2, lid 1, punt c), en overweging 22 van de AI-verordening.

([5]) Artikel 3, lid 4, van de AI-verordening.

(13) Gebruiksverantwoordelijken vallen binnen het toepassingsgebied van de transparantieverplichtingen van artikel 50, leden 3 en 4, van de AI-verordening indien hun plaats van vestiging of locatie zich in de Unie bevindt of indien ze zijn gevestigd of zich bevinden in een derde land waar de output van het AI-systeem in de Unie wordt gebruikt ([6]). In het laatste geval zijn de transparantieverplichtingen van toepassing op entiteiten die zich buiten de Unie bevinden of gevestigd zijn, wanneer de gebruiksverantwoordelijke zelf voorziet in de verspreiding en het gebruik van de AI-output in de Unie (d.w.z. door distributie binnen de Unie te sturen of toe te staan, onder meer door deepfakes op het wereldwijd toegankelijke internet te plaatsen). Gebruiksverantwoordelijke uit derde landen zijn echter niet gebonden aan de transparantieverplichtingen wanneer de content van het AI-systeem het publiek in de Unie bereikt via kanalen die onvoorzienbaar zijn en waarop zij geen invloed hebben.

(14) Wanneer de gebruiksverantwoordelijke van een AI-systeem een rechtspersoon is onder wiens gezag het systeem wordt gebruikt (bv. een reclamebureau), mogen de individuele werknemers die handelen op instructie en onder toezicht van die rechtspersoon (bv. digitale animatoren, webontwerpers, makers van content, journalisten) niet worden beschouwd als afzonderlijke gebruiksverantwoordelijken voor dat systeem. Een rechtspersoon blijft een gebruiksverantwoordelijke, ook wanneer hij derden (bv. contractanten, freelancers) inschakelt bij de exploitatie van het systeem namens hem en onder zijn verantwoordelijkheid en toezicht.

> Zo is een in de Unie gevestigd mediakanaal waarvan de redacteuren een AI-systeem gebruiken om hun schriftelijke berichtgeving over actuele gebeurtenissen te ondersteunen met door AI gegenereerde tekstpublicaties over aangelegenheden van algemeen belang die op een wereldwijd toegankelijke website worden geplaatst, een gebruiksverantwoordelijke die binnen het toepassingsgebied van de AI-verordening valt. Evenzo is een in een derde land gevestigd bedrijf dat een AI-systeem gebruikt om een deepfake van een beroemdheid te genereren die voorkomt in een reclame die in de Unie wordt getoond, ook een gebruiksverantwoordelijke die binnen het toepassingsgebied van de AI-verordening valt.
>
> Daarentegen is een onderneming die zich ertoe beperkt een reclamebureau opdracht te geven om een advertentie te produceren, zonder beslissingen te nemen en controle uit te oefenen over de vraag of en hoe het reclamebureau AI in het productieproces inzet, geen gebruiksverantwoordelijke.

(15) Exploitanten (b.v. aanbieders en gebruiksverantwoordelijken) kunnen tegelijkertijd meer dan één rol vervullen met betrekking tot een AI-systeem. Een bedrijf kan bijvoorbeeld een intern generatief AI-systeem ontwikkelen en gebruiken om deepfakes te genereren. In een dergelijk geval wordt het aangemerkt als zowel een aanbieder als een gebruiksverantwoordelijke die moet zorgen voor naleving van de verplichtingen inzake machineleesbare markering op grond van artikel 50, lid 2, en de vereisten voor een duidelijke en waarneembare labeling van deepfakes op grond van artikel 50, lid 4, van de AI-verordening.

(16) Andere actoren (zoals aanbieders van hostingdiensten, waaronder onlineplatforms, of omroepen) wier rol beperkt is tot het verspreiden of doorgeven van door AI gegenereerde of gemanipuleerde content die door derden is gecreëerd, of die door AI gegenereerde of gemanipuleerde content ontvangen of daaraan worden blootgesteld zonder zeggenschap te hebben over het gebruik van het AI-systeem voor het genereren of manipuleren van de AI-content, zijn daarentegen geen gebruiksverantwoordelijken in de zin van de AI-verordening. Hoewel die actoren niet verantwoordelijk of aansprakelijk zijn op grond van de AI-verordening, kunnen ze niettemin een belangrijke rol spelen in de waardeketen om het vertrouwen in en de integriteit van het informatie-ecosysteem te waarborgen. Ze worden daarom sterk aangemoedigd om de overeenkomstig artikel 50 van de AI-verordening ingevoerde markering en labeling te handhaven en evenredige en technisch haalbare maatregelen te nemen om natuurlijke personen in staat te stellen door AI gegenereerde of gemanipuleerde content te detecteren, onverminderd de verantwoordelijkheid van aanbieders en gebruiksverantwoordelijken van generatieve AI-systemen of enige verplichtingen of aansprakelijkheid uit hoofde van Verordening (EU) 2022/2065 (de digitaledienstenverordening) ([7]) of andere toepasselijke Unie- of nationale wetgeving.

([6]) Artikel 2, lid 1, punten b) en c), van de AI-verordening.

([7]) Verordening (EU) 2022/2065 van het Europees Parlement en de Raad van 19 oktober 2022 betreffende een eengemaakte markt voor digitale diensten (PB L 277 van 27.10.2022, blz. 1).

(17) Ook andere actoren die in hun professionele hoedanigheid handelen en content verspreiden (zelfs als ze niet rechtstreeks onder artikel 50 van de AI-verordening vallen), worden aangemoedigd om evenredige en passende maatregelen te nemen, zodat de natuurlijke personen die uiteindelijk aan de content worden blootgesteld, daadwerkelijk worden geïnformeerd en kunnen detecteren dat die kunstmatig is gegenereerd of gemanipuleerd. Daarnaast kan van die partijen worden verlangd dat ze de respectieve deepfakes of tekstpublicaties over aangelegenheden van algemeen belang labelen op grond van andere toepasselijke wettelijke of deontologische regels (bv. een omroeporganisatie die een door een particulier gecreëerde deepfake labelt en als onderdeel van een tv-programma laat zien). Het enkele feit dat een actor content labelt of gelabelde content gebruikt, betekent echter niet dat die kan worden aangemerkt als een "gebruiksverantwoordelijke" in de zin van de AI-verordening. De AI-verordening is niet van toepassing en heeft geen invloed op de wettelijke verantwoordelijkheid van die actoren op grond van de toepasselijke rechtskaders indien zij niet tegelijkertijd als aanbieders of gebruiksverantwoordelijken van AI-systemen kunnen worden aangemerkt.

2.4. **Uitsluitingen van het toepassingsgebied van de AI-verordening**

(18) Artikel 2 van de AI-verordening voorziet in een aantal algemene uitsluitingen van het toepassingsgebied van de AI-verordening die reeds zijn verduidelijkt in de richtsnoeren van de Commissie betreffende verboden AI-praktijken ([8]). In dit punt wordt dan ook alleen ingegaan op de aspecten die relevant zijn voor de transparantieverplichtingen van artikel 50 van de AI-verordening.

2.4.1. *Gebruiksverantwoordelijken die handelen in het kader van een louter persoonlijke niet-professionele activiteit*

(19) Volgens artikel 2, lid 10, van de AI-verordening is de verordening "niet van toepassing op verplichtingen van gebruiksverantwoordelijken die natuurlijke personen zijn die AI-systemen gebruiken in het kader van een louter persoonlijke niet-professionele activiteit" ([9]). Activiteiten waarmee natuurlijke personen regelmatig economische voordelen behalen of beroepsmatige, bedrijfsmatige, handels- of freelanceactiviteiten waarbij zij anderszins betrokken zijn, moeten als een "professionele activiteit" worden beschouwd. De toevoeging "louter persoonlijk" vormt een verdere specificatie van niet-professioneel, waarmee wordt aangegeven dat de persoon zowel in een persoonlijke als in een niet-professionele hoedanigheid moet handelen om onder het toepassingsgebied van de uitzondering te vallen. Elk gebruik door natuurlijke personen die handelen namens of onder het gezag van een in een professionele hoedanigheid optredende gebruiksverantwoordelijke, valt onder de transparantieverplichtingen van de professionele gebruiksverantwoordelijke zoals vastgelegd in artikel 50, leden 3 en 4, van de AI-verordening.

Voorbeelden van persoonlijk niet-professioneel gebruik van AI-systemen die van het toepassingsgebied zijn uitgesloten

— Personen die een AI-systeem gebruiken om deepfakes van de leden van hun huishouden te genereren.

— Personen die persoonlijk een AI-systeem gebruiken om tekst over verschillende onderwerpen te genereren (bv. leerlingen die AI gebruiken om hun huiswerk te maken).

— Personen die persoonlijk een AI-systeem gebruiken om deepfakes te genereren die op sociale media worden verspreid.

(20) De uitsluitingsgrond van artikel 2, lid 10, van de AI-verordening is alleen van toepassing op de verplichtingen van gebruiksverantwoordelijken wanneer zij het systeem gebruiken voor louter persoonlijke niet-professionele activiteiten. Het systeem als zodanig valt nog steeds binnen het toepassingsgebied van de AI-verordening wat betreft de verplichtingen van aanbieders die het systeem in de handel brengen of in gebruik stellen en van gebruiksverantwoordelijken die het systeem voor professioneel gebruik inzetten. Deze uitzondering laat tevens de toepassing van ander toepasselijk Unie- en nationaal recht onverlet.

([8]) Zie punt 2.5 van de richtsnoeren van de Commissie betreffende verboden artificiële-intelligentiepraktijken zoals vastgesteld bij Verordening (EU) 2024/1689 (AI-verordening), Brussel, 29.7.2025, C(2025) 5052 final.

([9]) De definitie van gebruiksverantwoordelijke sluit ook gebruikers uit die dit soort activiteiten verrichten (zie punt 2.3).

Zo moet de aanbieder van het AI-systeem dat wordt gebruikt om deepfakes voor op een persoonlijke kerstkaart te genereren, die deepfakes op machineleesbare wijze markeren overeenkomstig de transparantieverplichting van artikel 50, lid 2, van de AI-verordening. Tegelijkertijd valt het gebruik van het systeem door natuurlijke personen voor louter persoonlijke niet-professionele activiteiten niet onder de verplichting voor gebruiksverantwoordelijken van artikel 50, lid 4, van de AI-verordening.

2.4.2. *Onderzoek & ontwikkeling*

(21) Artikel 2, lid 6, van de AI-verordening bevat een uitsluiting van het toepassingsgebied op AI-systemen of AI-modellen, met inbegrip van hun output, die specifiek zijn ontwikkeld en in gebruik gesteld met wetenschappelijk onderzoek en wetenschappelijke ontwikkeling als enig doel. Die uitsluiting laat de toepasselijke normen inzake ethisch onderzoek en de EU-wetgeving inzake gegevensbescherming en andere onderwerpen onverlet. In het kader van artikel 50 van de AI-verordening heeft de uitsluiting niet alleen betrekking op interactieve AI-systemen, maar ook op de output van generatieve AI-systemen (zoals deepfake-content) die in het kader van wetenschappelijk onderzoek worden gebruikt. Indien die systemen echter in gebruik worden gesteld of hun output wordt gebruikt voor andere doeleinden dan wetenschappelijk onderzoek en wetenschappelijke ontwikkeling als enig doel, moet nog steeds aan de desbetreffende transparantieverplichtingen van artikel 50 van de AI-verordening worden voldaan.

Als onderzoekers bijvoorbeeld een interactief AI-systeem ontwikkelen en in de handel brengen of in gebruik stellen en willen onderzoeken of natuurlijke personen in staat zijn een onderscheid te maken tussen AI-systemen en mensen bij mondelinge interactie met dat AI-systeem, hoeven zij de respectieve transparantieverplichtingen van artikel 50, leden 1 en 2, van de AI-verordening niet te volgen. Als hetzelfde interactieve AI-systeem echter ook voor andere doeleinden dan onderzoek en ontwikkeling in gebruik wordt gesteld, moet aan de desbetreffende transparantieverplichtingen van artikel 50 van de AI-verordening worden voldaan.

(22) In artikel 2, lid 8, van de AI-verordening worden bovendien onderzoeks-, test- of ontwikkelingsactiviteiten met betrekking tot AI-systemen of AI-modellen uitgesloten voordat zij in de handel worden gebracht of in gebruik worden gesteld. Testen onder reële omstandigheden (binnen of buiten AI-testomgevingen voor regelgeving) valt niet onder die uitsluiting. Bovendien moet een AI-systeem, zodra het als gevolg van dergelijke onderzoeks- en testactiviteiten in de handel wordt gebracht of in gebruik wordt gesteld, niettemin voldoen aan de relevante transparantieverplichtingen van artikel 50 van de AI-verordening ([10]).

2.4.3. *AI die wordt vrijgegeven onder vrije en opensource licenties*

(23) Overeenkomstig artikel 2, lid 12, van de AI-verordening vallen AI-systemen die worden vrijgegeven onder vrije en opensource licenties buiten het toepassingsgebied van de AI-verordening, mits ze niet onder de verbodsbepalingen van artikel 5 van de AI-verordening of de transparantieverplichtingen van artikel 50 van de AI-verordening vallen, of niet als systemen met een hoog risico zijn geclassificeerd ([11]). Dit betekent dat aanbieders en gebruiksverantwoordelijken van opensource-AI-systemen die binnen het toepassingsgebied van artikel 50 van de AI-verordening vallen, nog steeds moeten zorgen voor naleving van hun respectieve transparantieverplichtingen.

(24) Aanbieders van vrije en opensource AI-componenten die betrekking hebben op software, data, modellen, instrumenten, diensten of processen voor aanpassing aan en integratie in een AI-systeem en die op zichzelf (geheel of gedeeltelijk) geen AI-systeem vormen, zijn daarentegen niet onderworpen aan de transparantieverplichtingen van artikel 50 van de AI-verordening. Aanbieders van dergelijke systemen worden niettemin aangemoedigd om die modellen, componenten en diensten op een in punt 27 beschreven manier te ontwerpen, voor zover dat evenredig en haalbaar is.

([10]) Overweging 25 van de AI-verordening.
([11]) In overweging 102 van de AI-verordening wordt toegelicht dat het vrijgeven van software en gegevens in het kader van een vrije en opensource licentie het mogelijk maakt om deze "vrijelijk [te delen] en gebruikers ze in al dan niet gewijzigde vorm vrijelijk kunnen raadplegen, gebruiken, wijzigen en herdistribueren".

2.5. **Wisselwerking met de verboden praktijken, de vereisten voor AI-systemen met een hoog risico en AI-geletterdheid**

(25) In overweging 137 van de AI-verordening wordt verduidelijkt dat de naleving van de transparantieverplichtingen van artikel 50 van de AI-verordening voor een bepaald AI-systeem niet aldus kan worden uitgelegd dat het gebruik van dat AI-systeem of de output ervan rechtmatig is op grond van de AI-verordening. Met name AI-systemen die onder de verplichtingen van artikel 50 van de AI-verordening vallen, kunnen onder artikel 5 van de AI-verordening vallen en dus verboden zijn (bv. het gebruik van systemen voor het herkennen van emoties op de werkplek of in onderwijsinstellingen) ([12]). Bovendien kunnen AI-systemen die binnen het toepassingsgebied van artikel 50 van de AI-verordening vallen, ook worden geclassificeerd als systemen met een hoog risico ([13]) en onderworpen zijn aan de relevante voorschriften en verplichtingen voor dergelijke systemen. De naleving van artikel 50 van de AI-verordening vervangt, beperkt of beïnvloedt derhalve geen verplichtingen die voortvloeien uit andere bepalingen van de AI-verordening ([14]). Tegelijkertijd kunnen AI-systemen ook binnen het toepassingsgebied van artikel 50 van de AI-verordening vallen zonder te worden geclassificeerd als systemen met een hoog risico op grond van artikel 6 van de AI-verordening. De vereisten inzake geletterdheid op grond van artikel 4 van de AI-verordening zijn ook van toepassing op zowel aanbieders als gebruiksverantwoordelijken van AI-systemen die binnen het toepassingsgebied van artikel 50 van de AI-verordening vallen.

2.6. **Wisselwerking met regels die van toepassing zijn op AI-modellen en/of -systemen voor algemene doeleinden**

(26) Artikel 50 van de AI-verordening is van toepassing op bepaalde AI-systemen, waaronder AI-systemen voor algemene doeleinden ("AI-systemen voor algemene doeleinden" of "GPAI-systemen") ([15]), zoals chatbots die directe interactie met natuurlijke personen ondersteunen of AI-systemen die synthetische content kunnen genereren of manipuleren (waaronder deepfakes) ([16]).

(27) Artikel 50 van de AI-verordening is niet uitdrukkelijk van toepassing op GPAI-modellen. Indien een interactief of generatief AI-systeem dat binnen het toepassingsgebied van artikel 50, lid 1 of 2, van de AI-verordening valt, echter is gebaseerd op een GPAI-model dat door dezelfde aanbieder wordt aangeboden, kunnen de in deze richtsnoeren beschreven transparantiemaatregelen ook op modelniveau worden genomen (zie ook punt 4). Bovendien worden aanbieders van andere GPAI-modellen en van andere generatieve AI-modellen, hoewel ze buiten het toepassingsgebied van artikel 50 van de AI-verordening vallen, ook aangemoedigd om op modelniveau passende transparantiemaatregelen te nemen om interactieve AI-systemen te herkennen en door AI gegenereerde en gemanipuleerde content te markeren. Dat is vooral van belang om het voor aanbieders verder in de waardeketen van AI-systemen gemakkelijker te maken aan hun verplichtingen op grond van artikel 50, leden 1 en 2, van de AI-verordening ([17]) te voldoen en bij te dragen aan het behoud van het vertrouwen in en de integriteit van het informatie-ecosysteem. Voor GPAI-modellen die een systeemrisico inhouden, kunnen veiligheidsmaatregelen op modelniveau voor het herkennen van AI-interacties en het markeren van door AI gegenereerde of gemanipuleerde content indien nodig deel uitmaken van een bredere reeks risicobeperkende maatregelen die aanbieders kunnen nemen in het kader van hun verplichting om systeemrisico's te beperken overeenkomstig artikel 55, lid 1, punt b), van de AI-verordening.

3. **ARTIKEL 50, LID 1, VAN DE AI-VERORDENING: TRANSPARANTIE VOOR INTERACTIEVE AI-SYSTEMEN**

(28) Artikel 50, lid 1, van de AI-verordening is bedoeld voor aanbieders van AI-systemen die direct interageren met natuurlijke personen; zij moeten hun systemen zodanig ontwerpen en ontwikkelen dat de betrokken natuurlijke personen worden geïnformeerd dat zij interageren met een AI-systeem. Zoals uiteengezet in overweging 132 van de AI-verordening, heeft deze informatieverplichting tot doel die natuurlijke personen in staat te stellen onderbouwde beslissingen te nemen over de output van het systeem, overmatig vertrouwen in dergelijke systemen te voorkomen en die natuurlijke personen in staat te stellen hun vertrouwen in de content en de interacties daar dienovereenkomstig op af te stemmen.

([12]) Voor meer informatie over artikel 5 van de AI-verordening, zie de richtsnoeren van de Commissie betreffende verboden artificiële-intelligentiepraktijken zoals vastgesteld bij Verordening (EU) 2024/1689 (AI-verordening), C(2025) 5052 final.
([13]) Zie artikel 6 van de AI-verordening.
([14]) Met name in artikel 26, lid 11, van de AI-verordening wordt aan gebruiksverantwoordelijken van AI-systemen met een hoog risico die beslissingen nemen of helpen nemen, een transparantieverplichting opgelegd om natuurlijke personen te informeren dat zij onderworpen zijn aan het gebruik van het AI-systeem met een hoog risico. Deze transparantieverplichtingen vormen een aanvulling op die van artikel 50 van de AI-verordening, aangezien ze betrekking hebben op andere informatie die aan de natuurlijke personen moet worden verstrekt, maar ze kunnen in één kennisgeving worden gebundeld.
([15]) In artikel 3, punt 66, van de AI-verordening wordt een GPAI-systeem gedefinieerd als een AI-systeem dat is gebaseerd op een AI-model voor algemene doeleinden en dat verschillende doeleinden kan dienen, zowel voor direct gebruik als voor integratie in andere AI-systemen.
([16]) Artikel 50, lid 2, van de AI-verordening bevat een expliciete verwijzing naar GPAI-systemen.
([17]) Overweging 133 van de AI-verordening.

3.1. **Belangrijkste componenten, concepten en daarmee samenhangende transparantieverplichting(en)**

(29) De transparantieverplichting van artikel 50, lid 1, van de AI-verordening is van toepassing op aanbieders van AI-systemen als het systeem voor directe interactie met natuurlijke personen is bedoeld en die interactie niet onder een van de volgende uitzonderingen valt: 1) het is duidelijk vanuit het oogpunt van een normaal geïnformeerde en redelijk omzichtige en oplettende natuurlijke persoon of 2) het systeem is bij wet toegestaan voor het opsporen, voorkomen, onderzoeken of vervolgen van strafbare feiten, tenzij het systeem voor het publiek beschikbaar is om een strafbaar feit te melden.

3.1.1. *AI-systemen die voor directe interactie met natuurlijke personen bedoeld zijn*

(30) Een systeem valt pas onder het toepassingsgebied van artikel 50, lid 1, van de AI-verordening indien aan vier voorwaarden is voldaan: i) het systeem moet een AI-systeem zijn ii) dat voor interactie, iii) met name directe, iv) met natuurlijke personen is bedoeld.

i) **Een AI-systeem**: het systeem moet voldoen aan de criteria van de definitie van een AI-systeem ([18]), waarbij eenvoudige, niet op AI gebaseerde geautomatiseerde responsmechanismen (bv. traditionele afwezigheidsberichten of op regels gebaseerde snelle antwoorden op berichten) zijn uitgesloten.

ii) **Voor interactie is bedoeld:** Het AI-systeem moet voor interactie met natuurlijke personen bedoeld zijn. Interactie houdt in dat er sprake is van een uitwisseling in twee richtingen van informatie of handelingen tussen natuurlijke personen en het AI-systeem, met het karakter van een echt gesprek of een echte reactie. Die kan in verschillende vormen plaatsvinden (bijvoorbeeld auditief, visueel en fysiek). Natuurlijke personen moeten input (bv. geschreven tekst of andere vormen van content, spraak of fysieke handelingen) aan het AI-systeem kunnen leveren (bv. via technische communicatiemiddelen, waaronder online-interfaces, telefonisch of per e-mail, of rechtstreeks), terwijl het AI-systeem in staat moet zijn te reageren met contextuele output (waaronder alle soorten content of handelingen) en omgekeerd. De interactie kan betrekking hebben op elk type output of formaat dat door de mens kan worden waargenomen en begrepen. Ze hoeft niet door een mens te worden geïnitieerd; ze kan ook betrekking hebben op AI-content of handelingen die een menselijke respons vereisen. De interactie kan een eenmalige uitwisseling zijn (bv. één prompt gevolgd door één antwoord) of plaatsvinden over een bepaalde periode (meerdere beurten). Systemen die alleen passief gegevens verzamelen en geen gegevens uit kunnen wisselen met natuurlijke personen (bv. geautomatiseerde toegangscontroles met gezichtsherkenning) of slechts eenmalige feedback verzamelen (bv. spamfilters), worden niet geacht te interageren met natuurlijke personen.

iii) **Directe interactie:** De interactie tussen het AI-systeem en de natuurlijke personen moet direct zijn. Dit omvat doorgaans realtime of bijna-realtime interactie (onder meer door de mogelijkheden van het systeem om berichten te schrijven aan en te verzenden naar natuurlijke personen of anderszins te interageren met de fysieke of virtuele omgeving die door natuurlijke personen kan worden waargenomen). Directe interactie omvat geen indirecte of bemiddelde menselijke interactie waarbij een natuurlijke persoon wordt blootgesteld aan AI-output zonder directe interactie met het systeem (bv. als vertegenwoordigers van klantendiensten AI-ondersteuningsinstrumenten gebruiken om hen te helpen communiceren met natuurlijke personen of als de AI-output niet aan de persoon beschikbaar wordt gesteld door het AI-systeem zelf, maar door een andere persoon die de content verspreidt). De loutere mogelijkheid voor mensen om in te grijpen of de output van het AI-systeem te toetsen, mag echter niet worden aangegrepen om de toepassing van de openbaarmakingsvereiste op grond van artikel 50, lid 1, van de AI-verordening. Gevallen waarin producten en diensten door AI gegenereerde antwoorden combineren met door mensen samengestelde content vallen onder het toepassingsgebied van artikel 50, lid 1, van de AI-verordening en vereisen dat wordt openbaar gemaakt welke specifieke output van de interactie door AI is gegenereerd, tenzij die AI-output naar behoren is gecontroleerd en verzonden door mensen die optreden als de belangrijkste gesprekspartners van de natuurlijke personen.

iv) **Met natuurlijke personen:** Artikel 50, lid 1, van de AI-verordening vereist dat het AI-systeem interageert met natuurlijke personen. Het kan daarbij gaan om professionele gebruiksverantwoordelijken, andere gebruikers (die het systeem ook voor louter persoonlijke activiteiten gebruiken) of andere personen die het systeem namens hen gebruiken. AI-systemen die werken in een gesloten fysieke omgeving (bv. als onderdeel van een industriële machine-opstelling), backend-machine-to-machine-communicatie tussen AI-systemen (waaronder AI-agenten) waarvan de output niet bedoeld is om natuurlijke personen te bereiken, of virtuele omgevingen zonder enig direct contact met natuurlijke personen, vallen buiten het toepassingsgebied van de verplichting.

([18]) Zie artikel 3, punt 1, van de AI-verordening. Voor meer informatie over het begrip "AI-systeem", zie de richtsnoeren van de Commissie betreffende de definitie van een AI-systeem, 29.7.2025, C(2025) 5053 final.

(31) AI-agenten vallen onder artikel 50, lid 1, van de AI-verordening indien ze in staat zijn om bij de uitvoering van de taken te interageren met de personen die hen instructies geven of met andere natuurlijke personen (bv. boekingen maken, correspondentie beheren, onderhandelen over contracten of die sluiten, aankopen doen enz.). AI-agenten moeten zodanig worden ontworpen en ontwikkeld dat ze bekendmaken dat ze kunstmatig zijn en wie de persoon is namens wie ze handelen, rekening houdend met de noodzaak van transparantie ten aanzien van de oorsprong en de delegatie van bevoegdheden, alsmede de verantwoordelijkheid voor de gevolgen van hun handelingen. Dit omvat gevallen waarin AI-agenten in staat zijn om in complexe architecturen met meerdere agenten te interageren met andere AI-agenten die direct interageren met de natuurlijke personen bij de uitvoering van de taak. Niettemin is het voor de aanbieder van de AI-agent wellicht niet altijd mogelijk om vooraf individuele gevallen van directe interactie tussen een AI-agent en andere natuurlijke personen te herkennen, zodra de AI-agent in verschillende contexten wordt ingezet. Indien de aanbieder niet op betrouwbare wijze kan bepalen of de AI-agent direct met een natuurlijke persoon zal interageren voordat hij de AI-agent in de handel brengt of in gebruik stelt, moet de agent op architectuurniveau worden ontworpen en geïnstrueerd om zichzelf als zodanig openbaar te maken in elke situatie waarin het redelijkerwijs waarschijnlijk is dat de agent met een natuurlijke persoon zal interageren (ook als een natuurlijke persoon mogelijk een juridische entiteit vertegenwoordigt of daarvoor optreedt). Bovendien moeten AI-agenten zich ook openbaar maken aan de personen die hen instructies geven bij belangrijke stappen (bv. op het moment van autorisatie, rapportage, validatie enz., ook wanneer de gebruikte AI-agent output ontvangt, verwerkt of gebruikt die is gegenereerd door andere AI-systemen in plaats van rechtstreeks door een natuurlijke persoon) en bij elke nieuwe interactie.

Voorbeelden van AI-systemen die voor directe interactie met natuurlijke personen zijn bedoeld die binnen het toepassingsgebied van artikel 50, lid 1, van de AI-verordening vallen

— Op AI gebaseerde spraakassistenten, chatbots/conversationele agenten in verschillende contexten (bv. overheidsdienst, klantendienst, klachtenbeheer, e-commerce, financiën, gezondheidszorg, onderwijs enz.).

— AI-meldpunt voor incidenten, fraude of andere meldingen, (humanoïde) robots/cobots, AI-metgezellen; robothuisdieren; AI-avatars (bv. in virtual-reality-omgevingen).

— AI-bots op sociale netwerken en media, coderingsagenten en andere AI-agenten die direct kunnen interageren met natuurlijke personen (waaronder in complexe architecturen met meerdere agenten).

3.2.1. **Voorbeelden van AI-systemen die niet direct interageren met natuurlijke personen en die buiten het toepassingsgebied van artikel 50, lid 1, van de AI-verordening vallen**

— Op AI gebaseerde traditionele industriële robots die in een gesloten industriële omgeving werken en niet voor interactie mensen bedoeld zijn (bv. robots voor elektronische assemblage).

— Op AI gebaseerde aanbevelingssystemen, spamfilters, geautomatiseerde vertaal- of transcriptie-hulpmiddelen.

— Op AI gebaseerd zoeken en opvragen van informatie zonder content te genereren of te wijzigen.

— Op AI gebaseerde methoden voor de automatische aanvulling van tekst en code, authenticatie-systemen/systemen voor biometrische herkenning, back-end besluitvormingssystemen waarbij de gebruiker alleen de AI-output ziet zonder de mogelijkheid van directe interactie met het systeem.

— Op AI gebaseerde voorspellende onderhouds- en optimalisatiesystemen in fabrieken en industriële toepassingen, door AI ondersteunde productvisualisatie (bv. virtuele try-on, kamerindeling, productweergave), op AI gebaseerde “single-turn”-routeplanners, transcriptie van oproepen van klanten die worden verwerkt voor gedragsanalyse enz.

3.1.2. *Informatieverplichting uit hoofde van artikel 50, lid 1, van de AI-verordening*

(32) Op grond van artikel 50, lid 1, van de AI-verordening zijn aanbieders van AI-systemen die direct interageren met natuurlijke personen verplicht ervoor te zorgen dat hun AI-systemen (inclusief AI-agenten en GPAI-systemen) zodanig worden ontworpen en ontwikkeld dat de betrokken natuurlijke personen worden geïnformeerd dat ze interageren met een AI-systeem.

(33) Hoewel het informatiemechanisme moet worden geïntegreerd in het ontwerp en de ontwikkeling van het systeem en in de wijze waarop het systeem met natuurlijke personen interageert, moet de kennisgeving plaatsvinden tijdens de werking van het AI-systeem en uiterlijk op het moment van de eerste interactie met de natuurlijke persoon (zoals vereist in artikel 50, lid 5, van de AI-verordening; zie punt 7). Aanbieders zijn ervoor verantwoordelijk dat, na het in de handel brengen of in gebruik nemen van het AI-systeem, natuurlijke personen die met het AI-systeem interageren, gedurende de gehele levenscyclus van het AI-systeem – zoals dat door die aanbieder in de handel is gebracht of in gebruik is genomen – doeltreffend en consistent worden geïnformeerd.2.2

(34) **Vorm van de openbaarmaking:** In artikel 50, lid 1, van de AI-verordening wordt geen specifieke methode voor openbaarmaking voorgeschreven. Aanbieders kunnen daarom elke geschikte methode kiezen, mits aan twee vereisten wordt voldaan: i) er is voldaan aan de voorschriften van artikel 50, lid 5, van de AI-verordening met betrekking tot de verstrekking van duidelijke en te onderscheiden informatie uiterlijk op het moment van de eerste interactie (zie punt 7); en ii) er wordt rekening gehouden met de kenmerken van tot kwetsbare groepen behorende natuurlijke personen vanwege hun leeftijd (bv. kinderen of ouderen) of handicap, voor zover het AI-systeem bedoeld is om met natuurlijke personen uit die groepen te interageren of het redelijkerwijs waarschijnlijk is dat dat gebeurt (bv. ook wanneer redelijkerwijs kan worden verwacht dat het AI-systeem interageert met het grote publiek) ([19]). De informatie moet ook worden verstrekt in een formaat dat toegankelijk is voor natuurlijke personen met een handicap, zoals gespecificeerd in artikel 50, lid 5, en overweging 132 van de AI-verordening (zie punt 7). Als het systeem met kinderen kan interageren, moeten die kennisgevingen kindvriendelijk, leeftijdsgeschikt, gemakkelijk te begrijpen en gemakkelijk toegankelijk zijn voor alle kinderen, ook kinderen met een handicap of extra toegankelijkheidsbehoeften. Indien beschikbaar en passend wordt aanbieders aanbevolen gebruik te maken van gestandaardiseerde openbaarmakingswerkwijzen en labels om de doeltreffendheid en consistentie van de kennisgevingen en een beter bewustzijn van de gebruikers te waarborgen.

(35) **Inhoud van de kennisgeving:** Op grond van artikel 50, lid 1, van de AI-verordening zijn aanbieders van interactieve AI-systemen verplicht die zodanig te ontwerpen en te ontwikkelen dat alle natuurlijke personen die met dergelijke AI-systemen interageren, uitdrukkelijk worden geïnformeerd over de kunstmatige en niet-menselijke aard van de gesprekspartner. Indien het voor de aanbieder redelijkerwijs te voorzien is dat het AI-systeem kan interageren met natuurlijke personen die tot de bovengenoemde kwetsbare groepen behoren, moet de informatie dienovereenkomstig worden aangepast en in een toegankelijk formaat worden verstrekt, zoals vereist krachtens het toepasselijke Unierecht, zodat ook die natuurlijke personen de informatie volledig kunnen begrijpen ([20]).

(36) In de kennisgevingen moet informatie worden verstrekt over de kunstmatige oorsprong van de interactie op een manier die past binnen het kader waarin de interactie plaatsvindt. Dat moet gebeuren op een manier die waarborgt dat de betrokken personen naar behoren en doeltreffend worden geïnformeerd, waarbij elk risico van misleiding of manipulatie vóór en tijdens de interactie wordt vermeden en tegelijkertijd wordt gezorgd voor haalbare en evenredige openbaarmakingen. Voorbeelden van het verstrekken van die informatie zijn onder meer schriftelijke kennisgevingen (bv. een chatbot die een gesprek begint met de melding dat het is gebaseerd op AI-technologie), visuele middelen (bv. een e-mail gegenereerd door een AI-agent die naar een natuurlijke persoon wordt gestuurd en bovenaan voorzien is van een AI-label), auditieve middelen (bv. een spraakassistent die aan het begin van een sessie aangeeft dat hij door AI wordt aangestuurd), het openbaar maken van AI-identificatiegegevens en referenties (bv. AI-agenten die hun AI-identiteit voor zover mogelijk op verifieerbare wijze openbaar maken) ([21]).

([19]) Zie overweging 132 van de AI-verordening. Dit vereist niet dat het systeem gericht is op die kwetsbare groepen. Indien een systeem bedoeld is om te interageren met leden van het publiek, dan vallen ook leden van kwetsbare groepen hieronder en moeten aanbieders daar rekening mee houden bij de uitvoering van de kennisgevingsmaatregelen.

([20]) Zie punt 7.3 voor de toepasselijke toegankelijkheidsvoorschriften.

([21]) Zo kunnen elektronische attesteringen van attributen zoals vastgesteld bij Verordening (EU) nr. 910/2014 en beschikbaar gesteld in de EU-portemonnees voor digitale identiteit die bij dezelfde verordening zijn vastgesteld, alsook de voorgestelde Europese portemonnees voor ondernemingen, veilige en efficiënte middelen bieden om AI-agenten te identificeren. De portemonnees kunnen elektronische attesteringen opslaan en beheren waarmee de identiteit, attributen en machtigingen van de AI-agent worden geverifieerd, waardoor een naadloze en betrouwbare openbaarmaking aan natuurlijke personen mogelijk wordt.

(37) **Vorm van de kennisgevingen**: Het ontwerp van het AI-systeem met multimodale en toegankelijke benaderingen voor openbaarmaking wordt aanbevolen als beste praktijk. In dergelijke kennisgevingen kunnen verschillende methoden worden gecombineerd om het begrip bij de gebruiker te vergroten, waar dat past en naargelang de context, de doelgroep en het te verwachten publiek, alsmede de beschikbare middelen van de aanbieders, met name kleine en middelgrote ondernemingen (kmo's). Voorbeelden van geschikte methoden zijn:

— **Tekstuele openbaarmaking** (op basis van gebruikersinterface (UI)): opvallende labels of banners in gewone taal (bv. "U interageert met een AI-systeem") en begroetingen bij de aanhef in chatbots die in voorkomend geval kunnen worden gecombineerd met badges of labels die gedurende de interactie blijvend zichtbaar zijn en die in verhouding staan tot de specifieke context en risico's. Voorts wordt aanbevolen om bijbehorende openbaarmakingen dicht bij de interactie-interface te plaatsen (bv. in de buurt van het input-/outputveld) en eenvoudige bewoordingen te gebruiken, vooral voor gebruikers met een lagere digitale geletterdheid en voor kinderen.

— **Auditieve kennisgeving:** in een context met spraakgestuurde of telefonische interacties moeten expliciete gesproken mededelingen aan het begin van de interactie (bv. "Dit is een assistent die werkt op basis van AI") worden aangevuld met, indien van toepassing, periodieke herinneringen tijdens langere interacties, met name in geval van onderbrekingen of een verandering in de rol van het AI-systeem tijdens de interactie. Afzonderlijke audiosignalen (bv. tonen of "earcons" (meldingsgeluiden)) kunnen helpen bij herkenning, vooral voor visueel gehandicapte gebruikers, maar worden op zichzelf niet voldoende geacht.

— **Visuele/grafische signalen:** ontwerp van het AI-systeem met blijvend zichtbare pictogrammen, watermerken, gekleurde frames of herkenbare "AI"-symbolen ter aanvulling van tekstuele openbaarmakingen. Het gebruik van gestandaardiseerde visuele indicatoren, ook voor andere diensten van een aanbieder, wordt aanbevolen, aangezien dergelijke indicatoren de cognitieve belasting aanzienlijk kunnen verminderen en de herkenbaarheid kunnen vergroten.

— **Multimodale combinaties:** het combineren van aanwijzingen in tekst, audio en beeld (bv. een chatbot die een label weergeeft en tegelijkertijd begroetingen bij de aanhef tekstuele informatie verstrekt) om de toegankelijkheid te waarborgen en de duidelijkheid voor verschillende gebruikersgroepen te vergroten.

(38) Sommige methoden zijn niet noodzakelijk en, wanneer zij afzonderlijk worden toegepast, onvoldoende om doeltreffend te voldoen aan de transparantieverplichting zoals bedoeld in artikel 50, leden 1 en 5, van de AI-verordening, met name:

— openbaarmakingen die alleen in voorwaarden, URL's of documentatie zijn opgenomen (dergelijke openbaarmakingen kunnen openbaarmaking binnen de context aanvullen maar niet vervangen);

— machineleesbare markeringen (bv. metadata of watermerken) die tijdens de interactie niet waarneembaar zijn voor gebruikers. Dit doet geen afbreuk aan de geschiktheid van dergelijke markeringen om te voldoen aan artikel 50, lid 2, van de AI-verordening (zie punt 4);

— onduidelijke of dubbelzinnige signalen (bv. algemene verwijzingen naar "assistent") of op mensen lijkende voorstellingen die gebruikers kunnen misleiden;

— algemene openbaarmakingen die niet specifiek genoeg zijn voor de output en interacties van het AI-systeem (bv.: op een platform dat verschillende diensten of verschillende soorten individuele input aanbiedt, is een algemene openbaarmaking zoals "De diensten op deze website gebruiken AI" ontoereikend);

— technische of capaciteitsgebonden beschrijvingen: verklaringen die uitsluitend verwijzen naar onderliggende technologieën (bv. "Dit systeem gebruikt LLM's") zonder de functie of implicaties van het systeem voor de gebruiker en de kunstmatige oorsprong ervan toe te lichten.

(39) Openbaarmakingen moeten evenredig zijn en afgestemd op de verschillende contexten en doelgroepen van het AI-systeem, zodat ze doeltreffend en gebruiksvriendelijk blijven en, voor zover mogelijk, verstoringen van de gebruikerservaring of gewenningseffecten (zogenaamde "bannerblindheid") worden voorkomen.

(40) Eén opvallende kennisgeving vóór de eerste interactie van het AI-systeem met een bepaalde natuurlijke persoon zal in de meeste gevallen waarschijnlijk volstaan om te voldoen aan de verplichting van artikel 50, lid 1, van de AI-verordening. In bepaalde risicovollere contexten zijn periodieke herinneringen en contextgerichte openbaarmakingen waarschijnlijk echter noodzakelijk om ervoor te zorgen dat gebruikers zich er voortdurend van bewust blijven. Dat is met name het geval bij AI-systemen die interageren met kwetsbare personen (bv. kinderen, ouderen, personen met een handicap en laaggeletterden), bij langdurige en veranderende interacties in gevoelige of immersieve contexten (bv. wanneer gebruikers emotionele stress of kwetsbaarheid kunnen uiten of ervaren of het risico lopen op verslavingsachtig gedrag), wanneer er een verhoogd risico bestaat dat gebruikers worden misleid (bv. bij financieel advies, verzekering, rechtsbijstand, gezondheidsadvies en klachtenbehandeling) of emotionele banden of afhankelijkheden ontwikkelen (bv. AI-metgezellen), alsook wanneer het nodig is om inzicht in en controle over de handelingen van AI-agenten te waarborgen. Aanbieders moeten het AI-systeem ook zo ontwerpen en ontwikkelen dat er in alle situaties waarin het AI-systeem vragen wordt gesteld over de aard ervan of over de oorsprong van de interactie, of waarin op basis van de communicatie met de natuurlijke persoon redelijkerwijs kan worden aangenomen dat die persoon waarschijnlijk zal worden misleid of in verwarring worden gebracht over de AI-oorsprong van de interactie, duidelijkheid wordt verschaft.

3.2. **Uitzonderingen op de informatieverplichting uit hoofde van artikel 50, lid 1, van de AI-verordening**

(41) Artikel 50, lid 1, van de AI-verordening voorziet in twee uitzonderingen op de informatieverplichting met betrekking tot interactieve AI-systemen, waarop hieronder wordt ingegaan.

3.2.1. *Uitzondering voor duidelijke interactie met een AI-systeem*

(42) Om de eerste uitzondering te kunnen inroepen, moeten aanbieders het volgende beoordelen en aantonen: i) de duidelijk kunstmatige aard van de interactie voor ii) een normaal geïnformeerde en redelijk omzichtige en oplettende natuurlijke persoon, rekening houdend met de omstandigheden en de gebruikscontext.

(43) Die norm is gebaseerd op gevestigde EU-wetgeving inzake consumentenbescherming met betrekking tot het begrip “gemiddelde consument” en moet consequent worden uitgelegd, rekening houdend met de specifieke en uiteenlopende contexten en toepassingen van AI-systemen die direct interageren met natuurlijke personen en de doelstellingen van artikel 50, lid 1, van de AI-verordening.

(44) Voor het definiëren van een dergelijke “gemiddelde” natuurlijke persoon zijn twee stappen nodig. Ten eerste moeten aanbieders rekening houden met de doelgroep waarvoor het AI-systeem is bedoeld en waarmee het naar alle waarschijnlijkheid zal interageren. Ten tweede moet de aanbieder nagaan hoe normaal geïnformeerd en redelijk omzichtig en oplettend een gemiddeld lid van die doelgroep is, rekening houdend met de omstandigheden en de gebruikscontext. Bij die beoordeling moeten verschillende factoren in aanmerking worden genomen:

— aanbieders moeten rekening houden met de doelgroep van het Ai-systeem en met een breder potentieel publiek dat redelijkerwijs te verwachten is. Dat zal met name het geval zijn als het grote publiek gemakkelijk toegang heeft tot het AI-systeem of als het systeem in virtuele of fysieke openbare ruimten functioneert en kan interageren met natuurlijke personen die in die ruimten aanwezig zijn. Het publiek dat mogelijk in aanraking komt met een AI-systeem, is niet altijd hetzelfde als de doelgroep.

— Aanbieders moeten rekening houden met de mogelijk diverse samenstelling van de doelgroep en het redelijkerwijs te verwachten publiek, in overeenstemming met overweging 132 van de AI-verordening. Wanneer personen met een handicap, ouderen of minderjarigen waarschijnlijk deel uitmaken van die doelgroep, zullen de verwachte niveaus van informatie, omzichtigheid en oplettendheid van een gemiddeld lid van een dergelijke doelgroep lager uitvallen dan van een gemiddeld lid van een doelgroep die deze categorieën van personen niet omvat. Als het interactieve AI-systeem daarentegen uitsluitend beschikbaar is voor een professionele of gespecialiseerde doelgroep, zou dat een positief effect kunnen hebben op het verwachte niveau van informatie, omzichtigheid en oplettendheid, en kan een gemiddeld lid van die doelgroep worden beschouwd als beter geïnformeerd, omzichtiger en oplettender.

— Als onderdeel van de bredere omstandigheden en gebruikscontext moeten aanbieders rekening houden met het niveau van digitale en AI-geletterdheid van de doelgroep en het redelijkerwijs te verwachten publiek en relevante sociale, culturele en taalkundige factoren.

(45) **Duidelijkheid**: Bij het vaststellen van de hypothetische normaal geïnformeerde, redelijk omzichtige en oplettende natuurlijke persoon moeten aanbieders beoordelen of het voor een dergelijke persoon duidelijk is dat die direct interageert met een AI-systeem. De uitzondering moet restrictief worden geïnterpreteerd, aangezien ze natuurlijke personen de bescherming ontneemt, evenals het recht om via duidelijke en te onderscheiden openbaarmakingen te informeren dat ze interageren met AI-systemen. Dat consumenten en andere natuurlijke personen in het algemeen weten dat er AI-systemen (zoals chatbots en agents) bestaan, betekent nog niet dat ze die tijdens interacties herkennen. Verschillende factoren kunnen relevant zijn en in aanmerking worden genomen bij de beoordeling of de interactie duidelijk is:

— De verwachte aard van de interactie: in gevallen waarin het AI-systeem fysiek interageert met natuurlijke personen, kunnen relevante elementen de zichtbaarheid van mechanische componenten omvatten (wat de duidelijkheid zou vergroten) of de mate waarin het AI-systeem zijn niet-kunstmatige equivalent op authentieke wijze repliceert (wat de duidelijkheid zou verminderen). Bij interacties van visuele of auditieve aard kunnen elementen zoals schrijf- of spraakpatronen, de klank van de stem (robotstem versus een authentiek menselijk klinkende stem), het ontwerp van de gebruikersinterface (bv. een profielfoto bij een chatbot waarop een mens te zien is) en de mogelijkheid tot geavanceerde, gepersonaliseerde interactie van invloed zijn op deze beoordeling.

— De samenstelling van de doelgroep en het te verwachten publiek: indien het interactieve AI-systeem uitsluitend beschikbaar is voor een professioneel of gespecialiseerd publiek, zonder dat kwetsbare groepen ermee in aanraking kunnen komen, zou dat ertoe kunnen leiden dat de duidelijkheid van de kunstmatige aard van de interactie voor een gemiddeld lid van het betreffende publiek toeneemt. Indien het AI-systeem daarentegen toegankelijk is voor het grote publiek of consumenten die ook kwetsbare personen omvatten (zoals kinderen, ouderen, personen met een cognitieve, lichamelijke of geestelijke handicap of personen met een lagere AI-geletterdheid) voor wie de interactie minder duidelijk is, kan geen beroep worden gedaan op de uitzondering om die kwetsbare gebruikers efficiënt te beschermen.

— Er moet ook rekening worden gehouden met het feit dat het voor natuurlijke personen steeds moeilijker wordt onderscheid te maken tussen een interactie met een AI-systeem en een met een echt mens. Vanwege het gebruik van avatars en agenten, de algemene anonimiteit van online-interacties en de mogelijkheden van AI-systemen om op een mensachtige manier te schrijven en te interageren, moet de uitzondering van duidelijkheid worden beperkt tot gevallen waarin er bijna geen twijfel overblijft over de aard van de interactie voor een gemiddelde persoon van de doelgroep en het redelijkerwijs te verwachten publiek.

Voorbeelden van duidelijkheid waarop de transparantieverplichting van artikel 50, lid 1, van de AI-verordening niet van toepassing is

— Door AI aangestuurde chatbots die helpen bij het coderen en het beoordelen van code en uitsluitend beschikbaar zijn voor professionele ontwikkelaars die, gezien hun expertise en de gebruikscontext, redelijkerwijs niet verwachten dat ze met een mens interageren en gemakkelijk kunnen herkennen dat de suggesties door een AI-systeem worden gegenereerd.

— Een interne assistent voor goed opgeleide medewerkers die AI-geletterd zijn en zich ervan bewust zijn dat ze AI-systemen gebruiken voor interne organisatorische doeleinden (bv. hr, juridische zaken, inkoop, naleving of IT-ondersteuning).

— Interactieve AI-systemen die uitsluitend bedoeld zijn voor gebruik door goed opgeleide zorgverleners ter ondersteuning van medische diagnoses en het voorstellen van mogelijke behandelingen.

— Ambient AI die is geïntegreerd in huishoudelijke apparaten, waarbij de functionaliteiten beperkt blijven tot ondersteuning en bediening van die apparaten.

— Interacties met op AI gebaseerde niet-speelbare personages (NPC's) in een videospel met één speler waarbij uit de aard van het videospel duidelijk blijkt geen andere natuurlijke persoon kan deelnemen aan en interageren met de gebruiker en de AI-aard van de interactie ook duidelijk is voor alle gebruikers (inclusief kwetsbare personen).

Voorbeelden die niet voldoen aan de uitzondering van duidelijkheid en waarvoor de transparantieverplichting van artikel 50, lid 1, van de AI-verordening van toepassing is

— Een door AI aangestuurd robothuisdier, dat sterk lijkt op zijn natuurlijke tegenhanger en is ontworpen om de typische interactie tussen mens en huisdier na te bootsen, waardoor onduidelijkheid ontstaat over de vraag of de interactie al dan niet met een AI-systeem plaatsvindt.

— AI-systemen die zijn ingebed in een immersieve omgeving (bv. virtuele of toegevoegde realiteit) waarbij gebruik wordt gemaakt van realistische avatars of stemmen die op die van mensen lijken, en waarbij gebruikers en met name kinderen, ouderen en personen met een handicap mogelijk niet gemakkelijk onderscheid kunnen maken tussen menselijke en AI-interactie.

— AI-chatbots die zijn ingebed in onlineplatforms of ondersteunende instrumenten (helpdesks) waarmee gebruikers direct interageren en AI-output ontvangen (bv. antwoorden op vragen of andere door AI gegenereerde content), die ze als door mensen gegenereerde content kunnen beschouwen.

3.2.2. *Uitzondering voor AI-systemen die bij wet zijn toegestaan met het oog op de rechtshandhaving*

(46) Aanbieders van interactieve AI-systemen zijn vrijgesteld van de transparantieverplichting van artikel 50, lid 1, van de AI-verordening indien ze bij wet zijn toegestaan strafbare feiten op te sporen, te voorkomen, te onderzoeken of te vervolgen, met inachtneming van passende waarborgen voor de rechten en vrijheden van derden.

(47) Met "bij wet toegestaan" wordt bedoeld: toegestaan op grond van het Unierecht, alsmede op grond van het nationale recht van de lidstaten dat in overeenstemming met het Unierecht is vastgesteld. In het recht waarbij het gebruik van het AI-systeem wordt toegestaan, moeten de doeleinden en omstandigheden van het toegestane gebruik duidelijk worden gespecificeerd en moet worden voorzien in passende waarborgen ter bescherming van de rechten en vrijheden van derden. Gezien de technologisch neutrale benadering van de AI-verordening is het niet nodig specifieke AI-systemen bij wet aan te wijzen of toe te staan wanneer het gebruik van het AI-systeem in het algemeen binnen het toepassingsgebied van de wettelijk toegestane activiteit valt en in overeenstemming is met het toepasselijke rechtskader.

(48) Om onder de uitzondering te vallen, moet het AI-systeem tot doel hebben strafbare feiten op te sporen, te voorkomen, te onderzoeken of te vervolgen (bv. AI-undercoveragent, AI-systemen die onder menselijke controle van onderzoekers worden gebruikt voor onlineonderzoeken of communicatie met personen die betrokken kunnen zijn bij criminele activiteiten). De uitzondering is niet beperkt tot het gebruik van dergelijke AI-systemen door rechtshandhavingsinstanties zoals bedoeld in artikel 3, punt 48, van de AI-verordening, maar kan ook betrekking hebben op interactieve (of generatieve) AI-systemen die door andere instanties van de Unie of nationale overheidsinstanties worden gebruikt, mits het gebruik ervan bij wet is toegestaan voor het opsporen, voorkomen, onderzoeken of vervolgen van strafbare feiten, en met inachtneming van passende waarborgen ter bescherming van de rechten en vrijheden van derden. De vrijstelling houdt niet in dat het AI-systeem in staat moet zijn om de onderzoeken zelfstandig uit te voeren. Indien een bepaald systeem zowel voor rechtshandhavingsdoeleinden als voor andere reguliere doeleinden kan worden gebruikt, moet de toepasselijkheid van de vrijstelling worden beoordeeld in het licht van het specifieke gebruik. De transparantieverplichtingen blijven van toepassing als hetzelfde AI-systeem wordt gebruikt voor andere doeleinden dan het wettelijk toegestane opsporen, voorkomen, onderzoeken of vervolgen van strafbare feiten. Er moet nog steeds transparantie worden geboden aan natuurlijke personen als die met het systeem interageren in het kader van de reguliere doeleinden ervan.

(49) De uitzondering is niet van toepassing indien i) het AI-systeem beschikbaar is voor het publiek en ii) het systeem personen een functionaliteit biedt om strafbare feiten te melden. Dergelijke systemen mogen worden gebruikt, maar blijven onderworpen aan de transparantieverplichting van artikel 50, lid 1, van de AI-verordening, waarbij natuurlijke personen geïnformeerd moeten worden gesteld dat ze interageren met dergelijke systemen.

Voor alle duidelijkheid: deze uitzondering heeft enkel en alleen betrekking op de verplichting van artikel 50, lid 1, om natuurlijke personen te informeren dat ze interageren met een AI-systeem ([22]). Ze doet geen afbreuk aan het toepasselijke Unierecht en nationale recht, inclusief procedureregels betreffende de documentatie, openbaarmaking, toelaatbaarheid en beoordeling van bewijsmateriaal.

([22]) Dat geldt ook voor de soortgelijke uitzonderingen ten behoeve van rechtshandhavingsactiviteiten als bedoeld in artikel 50, leden 2 tot en met 4, van de AI-verordening.

Voorbeelden van AI-systemen die ter beschikking van het publiek worden gesteld om strafbare feiten te melden die niet zijn vrijgesteld en binnen het toepassingsgebied van artikel 50, lid 1, van de AI-verordening vallen

— Op AI gebaseerde chatbots van de politie die worden ingezet op officiële websites of mobiele applicaties van de politie waarmee personen strafbare feiten of klachten kunnen melden. Die systemen leiden gebruikers door gestructureerde vragen, verzamelen relevante informatie en sorteren de meldingen, alvorens ze door te sturen naar menselijke medewerkers voor beoordeling.

— ·Door AI aangestuurde telefonische meldpunten die door de politie of een andere rechtshandhavingsinstantie worden beheerd voor algemene openbare onderzoeken.

— Door AI ondersteunde meldpunten of digitale portalen om fraude te melden, die worden beheerd door financiële instellingen of overheidsinstanties en waar gebruikers vermoedelijke financiële misdrijven kunnen melden. Het AI-systeem interageert met de gebruiker om details te verzamelen, de melding te categoriseren en zaken te prioriteren, terwijl menselijke onderzoekers de informatie valideren voordat er actie wordt ondernomen.

— Virtuele assistenten die worden gebruikt voor het verzamelen van getuigenverklaringen, waarbij slachtoffers of getuigen informatie kunnen indienen via een AI-gestuurde interface. Dergelijke AI-systemen bieden wellicht ondersteuning voor meertalige input, toegankelijkheidsfuncties en het indienen van gestructureerd bewijsmateriaal, maar de uiteindelijke beoordeling en de verdere stappen in het onderzoek blijven voorbehouden aan menselijke autoriteiten.

3.3. Wisselwerking met andere rechtshandelingen van de Unie

(50) De transparantieverplichting van artikel 50, lid 1, van de AI-verordening doet geen afbreuk aan bestaande Uniewetgeving inzake consumentenbescherming, gegevensbescherming, digitalisering of non-discriminatie op grond waarvan mogelijk andere informatie moet worden verstrekt die niet specifiek is voor de AI-oorsprong van de interactie. Meer bepaald zijn krachtens Richtlijn 2005/29/EG (de richtlijn oneerlijke handelspraktijken) ([23]) misleidende handelingen of omissies met betrekking tot de belangrijkste kenmerken van een product of dienst verboden. Bovendien verplicht Richtlijn 2011/83/EG (de richtlijn consumentenrechten) ([24]) handelaren ertoe om duidelijke en begrijpelijke precontractuele informatie te verstrekken over de belangrijkste kenmerken van de goederen of diensten. In het geval van een AI-gestuurde dienst (zoals een chatbot op abonnementsbasis, een app met een AI-metgezel of een virtuele coachingdienst) kan de AI-functionaliteit worden aangemerkt als een essentieel kenmerk dat openbaar gemaakt moet worden voordat de consument zich aan een overeenkomst bindt. De informatieverplichtingen uit hoofde van het Unierecht inzake consumentenbescherming zijn van toepassing ongeacht of de interactie als "duidelijk" wordt beschouwd in de zin van artikel 50, lid 1, van de AI-verordening. Deze verplichtingen vormen tevens een aanvulling op de algemene informatieverplichtingen inzake consumentenbescherming, terwijl de kennisgevingsverplichting uit hoofde van de AI-verordening meer contextgebonden is en ervoor moet zorgen dat alle natuurlijke personen zich ten minste vóór aanvang van de interactie en gedurende de gehele interactie bewust zijn van het feit dat zij rechtstreeks met een AI-systeem interageren.

(51) De digitaledienstenverordening bevat ook transparantievereisten voor aanbevelingssystemen van onlineplatforms, waaronder de verplichting om de belangrijkste rangschikkingsparameters toe te lichten. Die verplichtingen vormen een aanvulling op de transparantieverplichtingen van artikel 50, lid 1, van de AI-verordening, aangezien zij een ander toepassingsgebied en doel hebben en betrekking hebben op systemen die in principe niet direct interageren met gebruikers in de zin van artikel 50, lid 1, van de AI-verordening.

(52) Wat de informatieverplichtingen ten aanzien van betrokkenen uit hoofde van het Unierecht inzake gegevensbescherming betreft, vervult de verplichting van artikel 50, lid 1, van de AI-verordening een andere doelstelling dan die verplichtingen en doet ze als zodanig geen afbreuk aan die informatieverplichtingen.

([23]) Richtlijn 2005/29/EG van het Europees Parlement en de Raad van 11 mei 2005 betreffende oneerlijke handelspraktijken van ondernemingen jegens consumenten op de interne markt (PB L 149 van 11.6.2005, blz. 2).

([24]) Richtlijn 2011/83/EU van het Europees Parlement en de Raad van 25 oktober 2011 betreffende consumentenrechten (PB L 304 van 22.11.2011, blz. 64).

(53) Evenzo is de transparantieverplichting van artikel 50, lid 1, van de AI-verordening van toepassing onverminderd het Unie- en nationaal recht over de verstrekking van informatie aan personen in het kader van strafprocedures (bv. Richtlijn 2012/13/EU en Richtlijn 2012/29/EU, zoals gewijzigd bij Richtlijn (EU) 2026/1472) ([25]).

4. **ARTIKEL 50, LID 2, VAN DE AI-VERORDENING: MARKERING EN DETECTIE VAN DOOR AI GEGENEREERDE OF GEMANIPULEERDE CONTENT**

(54) Op grond van artikel 50, lid 2, van de AI-verordening moeten aanbieders van AI-systemen die synthetische content genereren, voorzien in technische oplossingen die voldoen aan bepaalde kwaliteitseisen voor machineleesbare markering en detectie van de output van hun AI-systemen. Het doel is natuurlijke personen de mogelijkheid te geven door AI gegenereerde of gemanipuleerde content te onderscheiden van andere content (bijvoorbeeld door mensen gecreëerde content) en de oorsprong ervan te verifiëren, en zo ook bij te dragen tot meer integriteit en vertrouwen in het informatie-ecosysteem.

(55) De begunstigden van deze transparantieverplichting zijn daarom alle natuurlijke personen die waarschijnlijk zullen worden blootgesteld aan de door AI gegenereerde of gemanipuleerde content en actoren die een belangrijke rol spelen voor het vertrouwen in en de integriteit van het informatie-ecosysteem, zoals bevoegde markttoezicht- en andere autoriteiten, onafhankelijke onderzoekers, maatschappelijke organisaties, media, betrouwbare flaggers, verdedigers van de grondrechten, aanbieders van onlineplatforms en onlinezoekmachines, rechthebbenden en andere makers en distributeurs van content enz.

4.1. **Belangrijkste componenten en concepten van artikel 50, lid 2, van de AI-verordening**

(56) Artikel 50, lid 2, van de AI-verordening is van toepassing indien aan verschillende voorwaarden tezamen is voldaan:

i) het systeem moet aan de criteria voor een AI-systeem voldoen;

ii) het AI-systeem moet synthetische content kunnen genereren of manipuleren;

iii) de content moet in een of meer van de volgende modaliteiten worden aangeboden: audio, beeld, video of tekst;

iv) het AI-systeem mag niet onder een van de volgende uitzonderingen vallen: 1) het AI-systeem vervult een ondersteunende functie voor standaardbewerking; 2) het AI-systeem brengt geen wezenlijke wijzigingen aan in de door de gebruiksverantwoordelijke verstrekte inputdata of de semantiek daarvan, of 3) het AI-systeem is bij wet toegestaan om strafbare feiten op te sporen, te voorkomen, te onderzoeken of te vervolgen.

4.1.1. *AI-systemen die synthetische content genereren of manipuleren*

(57) Artikel 50, lid 2, van de AI-verordening is van toepassing op AI-systemen ([26]) (inclusief GPAI-systemen) die synthetische audio-, beeld-, video- of tekstcontent genereren. Artikel 50, lid 2, van de AI-verordening heeft betrekking op een subset van systemen die voldoen aan de definitie van AI-systeem in artikel 3, punt 1, van de AI-verordening en die dergelijke output genereren. Artikel 50, lid 2, van de AI-verordening heeft ook betrekking op AI-systemen die worden gebruikt om content te manipuleren.

(58) Het genereren van content heeft betrekking op het genereren van synthetisch materiaal door het AI-systeem (bv. op basis van een menselijke prompt), zoals het genereren van een synthetisch beeld of een lied. In de praktijk betekent dat dat artikel 50, lid 2, van de AI-verordening van toepassing kan zijn op generatieve AI-systemen die zijn ontworpen met een beperkt beoogd doel om specifieke output te produceren. Tegelijkertijd is artikel 50, lid 2, van de AI-verordening ook van toepassing op AI-systemen die meerdere doeleinden kunnen dienen (zoals GPAI-systemen) en die verschillende soorten content kunnen produceren, alsook agentische AI-systemen, mits ze synthetische audio-, beeld-, video- of tekstcontent genereren.

([25]) Richtlijn 2012/13/EU van het Europees Parlement en de Raad van 22 mei 2012 betreffende het recht op informatie in strafprocedures (PB L 142 van 1.6.2012, blz. 1); Richtlijn 2012/29/EU van het Europees Parlement en de Raad van 25 oktober 2012 tot vaststelling van minimumnormen voor de rechten, de ondersteuning en de bescherming van slachtoffers van strafbare feiten (PB L 315 van 14.11.2012, blz. 57).

([26]) Zoals gedefinieerd in artikel 3, punt 1, van de AI-verordening.

(59) Het manipuleren van content heeft daarentegen betrekking op reeds bestaande (al dan niet synthetische) content die door een AI-systeem wordt veranderd (bv. een bestaand beeld of een stemopname die door een AI-systeem wordt gemanipuleerd in overeenstemming met menselijke instructies). Om onder artikel 50, lid 2, van de AI-verordening te vallen, volstaat het dat content door AI zodanig wordt gegenereerd of gemanipuleerd dat dat verder gaat dan standaardbewerking op grond van de uitzonderingen van artikel 50, lid 2, van de AI-verordening (zie punt 4.3). Bovendien vereist artikel 50, lid 2, van de AI-verordening niet dat de content uitsluitend door AI wordt gegenereerd of gemanipuleerd. Content die wordt vermengd met door mensen gecreëerd materiaal wordt ook als synthetische content aangemerkt, indien die is gemanipuleerd of gegenereerd volgens een van de in artikel 50, lid 2, van de AI-verordening genoemde methoden.

4.1.2. *Modaliteiten van synthetische content die binnen het toepassingsgebied vallen*

(60) Artikel 50, lid 2, van de AI-verordening bevat een uitputtende lijst van de modaliteiten van de synthetische content die onder de transparantieverplichting valt: audio, beeld, video of tekst. Daaronder valt ook multimodale content, dat wil zeggen content die bestaat uit een mix van deze modaliteiten. Voor de toepassing van artikel 50, lid 2, van de AI-verordening moeten de verschillende modaliteiten worden geïnterpreteerd in overeenstemming met praktische en technologische ontwikkelingen die in de loop van de tijd kunnen evolueren. Voorlopig moeten ze als volgt worden opgevat:

— Het begrip "tekst" heeft betrekking op afzonderlijke symbolische content die bestaat uit tekens, of cijfers die in een bepaalde volgorde zijn gerangschikt, die door mensen semantisch kunnen worden gelezen en geïnterpreteerd.

— Het begrip "beeld" heeft betrekking op statische ruimtelijke voorstellingen waarin visuele informatie is vastgelegd en die door mensen kunnen worden gezien.

— Het begrip "audio" heeft betrekking op een in de tijd variërend signaalcoderingsgeluid dat door mensen kan worden waargenomen door middel van gehoor. Dat kan betrekking hebben op spraak, instrumentale muziek of andere audiosignalen.

— Het begrip "video" heeft betrekking op een op tijd gebaseerde reeks beelden, die kan worden gesynchroniseerd met audio en door mensen kan worden gezien. In dat geval moet zowel de video- als de audiocontent worden gemarkeerd en moet de detecteerbaarheid worden gewaarborgd, zoals hieronder uiteengezet.

(61) AI-systemen die 3D-beelden, -video's en -audio genereren of manipuleren, alsook virtuele, toegevoegde of gemengde realiteit, vallen binnen het toepassingsgebied van artikel 50, lid 2, van de AI-verordening. "Virtuele realiteit" heeft betrekking op AI-technologieën waarmee gebruikers een door de computer gegenereerde omgeving waarin fysieke aanwezigheid wordt gesimuleerd, kunnen ervaren en ermee interageren, en die realtime interactie van gebruikers met die omgeving mogelijk maken. Aangezien virtuele en toegevoegde realiteit bestaat uit een op tijd gebaseerde reeks beelden, eventueel gesynchroniseerd met audio, die de visuele verandering in de loop van de tijd weergeeft, kunnen ze worden beschouwd als "video"-content. De "output" die moet worden gemarkeerd, is de definitieve vorm van het afzonderlijke, door AI gegenereerde materiaal, met speciale uitzonderingen voor de tijdelijke, interactieve realtime-sessies zoals bedoeld in punt 88. Digitale artefacten die worden gebruikt om 3D-beelden, -video's en -audiocontent te genereren, zoals een puntenwolk en een 3D-mesh, hoeven niet te worden gemarkeerd.

(62) Ook digitale tweelingen die een virtuele replica vormen van personen, fysieke voorwerpen of systemen en waarin meerdere soorten gegevens zijn geïntegreerd, vallen onder het toepassingsgebied van artikel 50, lid 2, van de AI-verordening als ze door AI zijn gegenereerd of gemanipuleerd en de vorm aannemen van audio, beeld, video of tekst, tenzij ze onder de uitzonderingen vallen en in aanmerking komen als "industriële toepassingen" of "business-to-business-toepassingen" overeenkomstig punt 87 of als tijdelijke realtime-content overeenkomstig punt 88.

(63) Artikel 50, lid 2, van de AI-verordening kan ook van toepassing zijn op AI-agenten indien de AI-agent een actie onderneemt waarvan de output door AI gegenereerde of gemanipuleerde content is die door natuurlijke personen waarneembaar is in de vorm van audio, beeld, video of tekst. Deze output moeten daarom worden gemarkeerd en detecteerbaar zijn zoals hieronder beschreven. Daarentegen worden handelingen van een AI-agent die geen door AI gegenereerde of gemanipuleerde content bevatten (bv. tussentijdse verwerkingsstappen zoals redeneringen en chain-of-thought) of door AI gegenereerde handelingen die niet bedoeld zijn om rechtstreeks door natuurlijke personen te worden waargenomen (bv. een webverzoek of browseractie), niet aangemerkt als synthetische content in de zin van artikel 50, lid 2, van de AI-verordening en hoeven ze niet te voldoen aan de markerings- en detectieverplichtingen.

4.1.3. *Content die buiten het toepassingsgebied van artikel 50, lid 2, van de AI-verordening valt*

(64) Content die door een AI-systeem wordt gegenereerd op basis van eenvoudige gegevensverwerking en die niet specifiek door AI wordt gegenereerd of gemanipuleerd, bijvoorbeeld een weergegeven kader, valt buiten het toepassingsgebied van artikel 50, lid 2, van de AI-verordening.

(65) Bovendien valt de output van een AI-systeem waarbij alleen bestaande content wordt gereproduceerd of de presentatie of ordening van reeds bestaande content mogelijk wordt gemaakt, niet onder artikel 50, lid 2, van de AI-verordening. Dat kan bijvoorbeeld het geval zijn bij afspeellijsten voor muziek of andere aanbevelingssystemen die uitsluitend bestaande content selecteren of rangschikken op basis van de voorkeuren of activiteiten van gebruikers, zonder dat er nieuwe content wordt gecreëerd of de content zelf wordt gemanipuleerd. Hetzelfde geldt voor AI-systemen die worden gebruikt in interne analytische processen die gegevens extraheren en structureren, maar die niet samenvatten.

(66) Hetzelfde geldt voor de input en output van een AI-systeem die louter waarnemingen en registraties van data uit fysieke of virtuele omgevingen vormen, die niet door AI worden gegenereerd of gemanipuleerd (bv. door robots of andere op AI gebaseerde sensoren) of datatransmissie door AI-systemen zonder enige wijziging. Dat heeft betrekking op verschillende soorten input en output van AI-systemen, met name in industriële omgevingen, zoals productiegegevens voor robots die gegevens registreren; verbruik geregistreerd door op AI gebaseerde slimme meters, netfrequentie en spanningsmetingen; de registratie van gps-locatiegegevens van voertuigen enz.

(67) Aangezien de transparantieverplichting van artikel 50, lid 2, van de AI-verordening tot doel heeft dat mensen in staat zijn door AI gegenereerde of gemanipuleerde content te onderscheiden om de risico's van misleiding en manipulatie aan te pakken en de integriteit van en het vertrouwen in het informatie-ecosysteem te waarborgen, valt content die geen verband houdt met die doelstelling en die niet waarneembaar is of niet bedoeld is om door natuurlijke personen te worden geïnterpreteerd, geverifieerd of als basis voor handelen te dienen, niet onder de bepaling.

(68) Met name de volgende content moet worden geacht buiten het toepassingsgebied van artikel 50, lid 2, van de AI-verordening te vallen:

— output die worden gegenereerd in de vorm van een korte reeks cijfers, symbolen of letters (bv. afzonderlijke woorden, beeldopnamen, alt-text, UI-labels en afbeeldingen ter grootte van een icoon, beeldopnamen en andere gegevenslabels);

— broncode bedoeld als content geschreven in een programmeer-, script-, markup-, query- of configuratietaal die bedoeld is om door een computersysteem te worden geïnterpreteerd, gecompileerd of uitgevoerd, ongeacht of die uit teksttekens bestaat. Dat kan ook opmerkingen in natuurlijke taal en contextuele informatie omvatten die integraal deel uitmaken van de broncode. Evenzo uitgesloten zijn Software Development Kits (SDK's), Structured Query Language (SQL), infrastructure-as-code en "YAML Ain't Markup Language" (YAML), JavaScript Object Notation (JSON), configuratieschema's, scripts, machineleesbare specificaties, applicatieprogramma-interfaces (API's & softwarebibliotheken);

— output van een AI-systeem die uitsluitend bedoeld is om van machine naar machine te worden gecommuniceerd en automatisch te worden verwerkt zonder dat mensen eraan blootgesteld worden, bijvoorbeeld communicatie tussen agenten en antispamsignalen, zolang die niet door natuurlijke personen wordt waargenomen;

— output die uitsluitend wordt gebruikt in omgevingen met gesloten programmalussen in workflows voor industriële toepassingen en productontwikkeling (bv. voor de productie van films, animaties, games of reclame), tenzij het gaat om de uiteindelijke output van het AI-systeem in de vorm van door AI gegenereerde of gemanipuleerde tekst-, audio-, beeld- of videocontent. In die gevallen hoeft alleen de uiteindelijke output van de stappen in de workflow te worden gemarkeerd en detecteerbaar te zijn.

4.2. **De markerings- en detectieverplichtingen van artikel 50, lid 2, van de AI-verordening**

(69) Artikel 50, lid 2, van de AI-verordening bevat een transparantieverplichting voor aanbieders van generatieve AI-systemen die binnen het toepassingsgebied van die bepaling vallen, bestaande uit twee afzonderlijke, maar inherent met elkaar verbonden elementen. Ten eerste moeten aanbieders ervoor zorgen dat de output van het AI-systeem in een machineleesbaar formaat worden gemarkeerd. Ten tweede moeten aanbieders ervoor zorgen dat de output detecteerbaar is als kunstmatig gegenereerde of gemanipuleerde content. Voor elke toegepaste markeringsoplossing moeten aanbieders ervoor zorgen dat er overeenkomstige detectiemiddelen beschikbaar zijn (zoals hieronder uiteengezet) om natuurlijke personen die aan de content worden blootgesteld en andere relevante actoren in staat te stellen de kunstmatig gegenereerde content van het AI-systeem te herkennen en te onderscheiden van andere content. Een technische oplossing moet worden verstaan als een combinatie van markeringstechnieken en detectiemiddelen waarin de aanbieder heeft voorzien om te voldoen aan de transparantieverplichting uit hoofde van artikel 50, lid 2, van de AI-verordening.

(70) Aan elk van de twee bovengenoemde elementen moet worden voldaan om de doelstellingen van de transparantieverplichtingen van artikel 50, lid 2, van de AI-verordening doeltreffend te verwezenlijken. Het vervullen van slechts één element (bv. machineleesbare markering van output zonder dat de detectiemiddelen ervan beschikbaar zijn) volstaat niet om aan die bepaling te voldoen ([27]).

4.2.1. *De markeringsverplichting*

(71) Het toepassingsgebied van de markeringsverplichting is beperkt tot het aanbrengen van markeringen in een machineleesbaar formaat. Een machineleesbaar formaat betekent dat markeringen zodanig zijn gestructureerd dat softwaretoepassingen ze gemakkelijk kunnen waarnemen, herkennen en extraheren zonder menselijke tussenkomst. Waarneembare markeringen en labels worden niet uitgesloten als aanvullende maatregel, indien van toepassing, om het voor gebruiksverantwoordelijken gemakkelijker te maken te voldoen aan hun verplichting om deepfakes te labelen overeenkomstig artikel 50, lid 4, van de AI-verordening.

(72) Om aan de markeringsverplichting te voldoen, kunnen aanbieders gebruikmaken van één enkele markeringstechniek of een combinatie van technieken, mits hun algehele technische oplossing machineleesbaar is en voldoet aan de eisen inzake doeltreffendheid, interoperabiliteit, robuustheid en betrouwbaarheid in de wettelijk vereiste mate, zoals bepaald in artikel 50, lid 2, tweede zin, van de AI-verordening (zie punt 4.2.3.).

(73) Overweging 133 van de AI-verordening bevat voorbeelden van dergelijke technieken, zoals watermerken, metagegevens, cryptografische methoden voor het aantonen van de herkomst en authenticiteit van content, loggingmethoden, vingerafdrukken of andere technieken, en een combinatie van dergelijke technieken. Hoewel in overweging 133 van de AI-verordening wordt verwezen naar methoden voor het aantonen van de herkomst en authenticiteit, zijn aanbieders niet verplicht een volledige herkomstketen te registreren of bij te houden met informatie over de oorsprong en wijzigingen van de content of andere beweringen over de geschiedenis van de content. Dergelijke methoden voor het aantonen van de herkomst kunnen echter ook worden gebruikt voor de naleving van artikel 50, lid 2, van de AI-verordening en kunnen natuurlijke personen in staat stellen door AI gegenereerde of gemanipuleerde content te onderscheiden van authentieke content.

(74) Aanbieders kunnen voorzien in de markeringsoplossing voor de output in verschillende stadia van de waardeketen (bv. na het genereren van content op het niveau van het AI-systeem, op het niveau van het onderliggende AI-model of geïntegreerd in het inferentieproces van het AI-systeem). Ze kunnen gebruikmaken van de markeringsoplossing van een upstream-modelaanbieder of een derde die de oplossing aanbiedt (bv. als een open norm of een gespecialiseerde dienst), voor zover de markeringsoplossing in overeenstemming is met artikel 50, lid 2, van de AI-verordening. Dat gebruik doet geen afbreuk aan de verantwoordelijkheid van de aanbieder van het AI-systeem om de naleving van artikel 50, lid 2, van de AI-verordening aan te tonen.

([27]) In overweging 133 van de AI-verordening wordt verduidelijkt dat het passend is aanbieders van die systemen voor te schrijven dat zij technische oplossingen inbouwen die markering in een machineleesbaar formaat mogelijk maken en aantonen dat de output is gegenereerd of gemanipuleerd door een AI-systeem en niet door een mens. Overweging 135 van de AI-verordening heeft betrekking op de detectieverplichting.

4.2.2. *De detectieverplichting*

(75) Aanbieders van AI-systemen die binnen het toepassingsgebied van artikel 50, lid 2, van de AI-verordening vallen, moeten ervoor zorgen dat de output van hun systemen detecteerbaar is als door AI gegenereerde of gemanipuleerde content. Dit betekent dat de aanbieder verplicht is ervoor te zorgen dat de detectiemiddelen beschikbaar zijn voor de personen die mogelijk aan de content worden blootgesteld. Bovendien moeten dergelijke detectieoplossingen overeenkomstig artikel 50, lid 5, van de AI-verordening voor mensen leesbare resultaten kunnen opleveren die aangeven of de content door AI is gegenereerd of gemanipuleerd.

(76) Een detectie-instrument is een mechanisme dat detecteert of content door AI is gegenereerd of gemanipuleerd, waarbij doorgaans technische markeringen of handtekeningen worden geïdentificeerd die de oorsprong ervan verifiëren. Om volledige interoperabiliteit te waarborgen, moeten aanbieders gebruikmaken van openbaar beschikbare detectieoplossingen volgens de industrienorm, waarmee derden kunnen voorzien in detectieoplossingen die bij voorkeur lokaal op het digitale apparaat kunnen worden uitgevoerd en die een naadloze integratie op zichzelf staande, geïsoleerde netwerken mogelijk maken. Indien dergelijke normen niet beschikbaar zijn, met name in de initiële fase van de uitvoering van artikel 50, lid 2, van de AI-verordening voor watermerktechnologie, kan de aanbieder gebruikmaken van zijn eigen detectieoplossing of van een derde of een gezamenlijke detectieoplossing, mits die oplossingen de interoperabiliteit waarborgen met de detectieoplossingen die door andere aanbieders worden gebruikt (zie punt 4.2.3). Die mogelijkheid moet qua duur beperkt blijven totdat geharmoniseerde normen en gestandaardiseerde, aanbiederonafhankelijke en interoperabele detectieoplossingen beschikbaar worden die veilig zijn, de privacy waarborgen en lokaal uitvoerbaar zijn.

(77) In artikel 50, lid 5, van de AI-verordening wordt verder bepaald dat de in artikel 50, lid 2, van de AI-verordening bedoelde informatie uiterlijk op het moment van de eerste interactie of blootstelling aan de content op duidelijke en te onderscheiden wijze aan de betrokken natuurlijke personen wordt verstrekt. Deze informatie is met name het resultaat van de detectie, waarbij wordt aangegeven of de content door AI is gegenereerd of gemanipuleerd, en niet de door AI gegenereerde of gemanipuleerde content zelf. Met “op het moment van eerste blootstelling” wordt bedoeld: het moment waarop natuurlijke personen bereid zijn de oorsprong van een specifiek element van de content te verifiëren en de detectieoplossing te raadplegen om de detectieresultaten te verkrijgen.

(78) Indien de aanbieder van het AI-systeem afhankelijk is van een andere actor voor de levering van of de toegang tot een detectieoplossing, moet hij ervoor zorgen dat de detectieoplossing voldoet aan artikel 50, leden 2 en 5, van de AI-verordening en dat die actoren de detectieresultaten weergeven overeenkomstig artikel 50, lid 5, van de AI-verordening (zie punt 7).

4.2.3. *Naleving van de eisen voor de technische oplossingen: doeltreffend, interoperabel, robuust en betrouwbaar*

(79) In artikel 50, lid 2, tweede zin, van de AI-verordening is bepaald dat de technische oplossingen voor markering en detectie doeltreffend, interoperabel, robuust en betrouwbaar moeten zijn. Die eisen moeten als volgt worden opgevat:

— **Doeltreffendheid** heeft betrekking op het vermogen van de door aanbieders ingevoerde technische oplossing voor de detectie van hun markeringen en het vermogen van natuurlijke personen om kunstmatig gegenereerde of gemanipuleerde content die door hun AI-systeem wordt geproduceerd, te onderscheiden en zo bij te dragen tot het vertrouwen in en de integriteit van het informatie-ecosysteem.

— **Betrouwbaarheid** heeft betrekking op het vermogen van de technische oplossing om, onder normale omstandigheden, door AI gegenereerde of gemanipuleerde content nauwkeurig te herkennen en te onderscheiden van andere content, en dit voor de gehele verscheidenheid aan content die door het AI-systeem van de aanbieder wordt gegenereerd of gemanipuleerd.

— **Robuustheid** heeft betrekking op het vermogen van de technische oplossing om de door het AI-systeem van de aanbieder gegenereerde of gemanipuleerde content nauwkeurig te herkennen en te onderscheiden van andere content onder uiteenlopende omstandigheden, waarbij zowel gangbare veranderingen als vijandige aanvallen worden meegenomen.

— **Interoperabiliteit** heeft betrekking op het vermogen van de verschillende technische oplossingen voor markering en detectie om naadloos te functioneren binnen meerdere systemen, bij verschillende actoren, in diverse contexten en bij technische uitvoeringen om de detectie van door AI gegenereerde of gemanipuleerde content mogelijk te maken, ongeacht de markeringstechniek(en) die de aanbieders van verschillende AI-systemen toepassen.

(80) Aanbieders moeten ervoor zorgen dat de combinatie van de technische oplossingen voor markering en detectie algeheel voldoet aan alle bovengenoemde eisen in de wettelijk vereiste mate, rekening houdend met de beperkingen en de complementariteit van de verschillende toegepaste oplossingen. De technische oplossingen moeten aan alle eisen voldoen voor zover dat technisch haalbaar is, rekening houdend met de specifieke kenmerken en beperkingen van de verschillende soorten content, de uitvoeringskosten en de algemeen erkende stand van de techniek, zoals tot uiting kan komen in desbetreffende technische normen.

(81) "Technisch haalbare" oplossingen in de zin van artikel 50, lid 2, van de AI-verordening zijn oplossingen die kunnen worden uitgevoerd voor de modaliteiten die onder het toepassingsgebied van de bepaling vallen, met gebruikmaking van de momenteel beschikbare technologie, methoden en technische werkwijzen, binnen de specifieke technische architectuur en operationele omgeving in kwestie. De aanbieder is niet verplicht gebruik te maken van een technische oplossing die nog niet is ontwikkeld of nog niet op de markt beschikbaar is of waarvan de uitvoering technisch niet haalbaar is. Technische haalbaarheid is een objectief begrip dat niet afhankelijk is van de specifieke middelen en capaciteiten van individuele aanbieders.

(82) Op grond van artikel 50, lid 2, van de AI-verordening moeten aanbieders voorzien in technisch haalbare oplossingen die in overeenstemming zijn met de algemeen erkende stand van de techniek, om te voldoen aan de vier kwaliteitseisen van doeltreffendheid, robuustheid, betrouwbaarheid en interoperabiliteit.

(83) Met "stand van de techniek" wordt bedoeld: een op een bepaald moment ontwikkelde fase van technisch vermogen met betrekking tot producten, processen en diensten, gebaseerd op geconsolideerde bevindingen van de wetenschap, technologie en ervaring en aanvaard als goede praktijk op technologisch gebied ([28]). De stand van de techniek houdt niet noodzakelijkerwijs in dat het om het recentste wetenschappelijk onderzoek gaat dat zich nog in een experimentele fase bevindt of nog onvoldoende technologisch uitgewerkt is. Aanbieders moeten hun oplossingen voor markering en detectie voortdurend, op tijd en evenredig aanpassen aan de ontwikkelingen op technologisch gebied en de stand van de techniek.

(84) Op grond van artikel 50, lid 7, van de AI-verordening kunnen aanbieders een praktijkcode hanteren die overeenkomstig die bepaling als toereikend is beoordeeld, om aan te tonen dat ze voldoen aan de markerings- en detectieverplichtingen als bedoeld in artikel 50, leden 2 en 5, van de AI-verordening, die in overeenstemming is met de stand van de techniek. Zie punt 8.1 voor meer informatie over de effecten van de naleving van een dergelijke code en het toezicht op andere aanbieders die zich niet aan een praktijkcode houden.

(85) Ook de kosten van de uitvoering van bepaalde technische oplossingen voor markering en detectie kunnen in aanmerking worden genomen. In bepaalde gevallen, zoals beschreven in de volgende punten, kunnen de kosten onevenredig zijn aan de marginale voordelen, die van beperkte waarde zijn om mensen in staat te stellen door AI gegenereerde of gemanipuleerde content te onderscheiden en om de integriteit van en het vertrouwen in het informatie-ecosysteem te waarborgen.

(86) Er kunnen derhalve eng omschreven gevallen zijn waarin een technische oplossing op basis van minder robuuste metadatamarkeringen volstaat om te voldoen aan artikel 50, lid 2, van de AI-verordening, rekening houdend met de specifieke kenmerken van de output van het AI-systeem, indien die inherent een lager risico op misleiding, manipulatie of negatieve effecten op het informatie-ecosysteem met zich meebrengt. Dat geldt vooral als een generatief AI-systeem is ingebouwd in fysieke producten die output genereren in een technisch gecontroleerde en gesloten omgeving die voornamelijk instructief van aard is (bv. een AI-systeem dat is ingebouwd in navigatiesystemen in voertuigen), voor zover er effectieve technische maatregelen in het product zijn ingebouwd om te voorkomen dat de output de productomgeving verlaat, bijvoorbeeld door vastlegging of export en verspreiding buiten het product.

([28]) Zie bijlage II bij Uitvoeringsbesluit C(2025) 3871 final van de Commissie betreffende een normalisatieverzoek aan het Europees Comité voor Normalisatie en het Europees Comité voor elektrotechnische normalisatie met betrekking tot AI-systemen met een hoog risico, ter ondersteuning van Verordening (EU) 2024/1689 van het Europees Parlement en de Raad en tot intrekking van Uitvoeringsbesluit C(2023) 3215 final.

(87) Er kunnen ook beperkte gevallen zijn van “industriële AI-toepassingen” of “business-to-businesstoepassingen” (met uitzondering van openbare en consumentgerichte AI-systemen) waarbij het evenredigheidsbeginsel en de beperkte voordelen voor de transparantiedoelstellingen inhouden dat er geen markering en detectie van door AI gegenereerde of gemanipuleerde output vereist is. Dit geldt uitsluitend indien aan alle volgende eisen is voldaan:

i) De door het AI-systeem gegenereerde output is uitsluitend van technische aard, bijvoorbeeld technische ontwerpen, industriële productieprocessen, technische instructies, output die voortkomt uit voorspellende systeemonderhoudsprocessen in een industriële omgeving, interne documentatieprocessen of productiestappen en werkstromen voordat de output wordt voltooid en ter beschikking wordt gesteld aan andere externe personen of het publiek.

ii) De door het AI-systeem gegenereerde output is uitsluitend bedoeld om te worden waargenomen en verwerkt door een beperkt, vooraf bepaald aantal natuurlijke personen dat in een professionele hoedanigheid handelt binnen de organisatie van de aanbieder en de gebruiksverantwoordelijke.

iii) De door het AI-systeem gegenereerde output mag niet bedoeld zijn om buiten het bedrijf te worden gedeeld of door externe personen te worden gebruikt, waarbij passende waarborgen zijn getroffen om redelijkerwijs te voorzien misbruik te voorkomen (bv. cloudisolatie, op rollen gebaseerde controles).

(88) Het genereren van realtime-content die van voorbijgaande aard is en onmiddellijk wordt geconsumeerd, zonder dat hij wordt opgenomen, opgeslagen of verder wordt verspreid (bv. in videogames, virtualrealitytoepassingen), kan eveneens worden vrijgesteld als markering technisch niet haalbaar is en de personen die aan de content worden blootgesteld erop gewezen worden dat de content door AI is gegenereerd of gemanipuleerd (bv. openbaarmaking tijdens de sessie, kennisgevingen op sessieniveau).

4.3. **Uitzonderingen op de verplichtingen op grond van artikel 50, lid 2, van de AI-verordening**

(89) Artikel 50, lid 2, van de AI-verordening voorziet in drie uitdrukkelijke uitzonderingen op de in die bepaling vastgelegde transparantieverplichtingen.

(90) De eerste uitzondering betreft AI-systemen voor zover die een ondersteunende functie voor standaardbewerking vervullen. Met “standaardbewerking” wordt bedoeld: het proces waarbij bestaande content wordt voorbereid voor publicatie of verspreiding (bv. kleine bewerkingen ter verbetering van de leesbaarheid, grammatica, kwaliteit en het formaat) en waarbij geen nieuwe content wordt gegenereerd. Standaardbewerking heeft tot doel ervoor te zorgen dat content onder meer vrij is van duidelijke technische of grammaticale taalfouten, is afgestemd op de toepasselijke voorschriften inzake opmaak, presentatie, formattering of toegankelijkheid of in overeenstemming is met sectorale werkwijzen. Bewerking gaat verder dan standaardbewerking als de content wezenlijk wordt gewijzigd (inhoudelijke aanpassingen, structurele wijzigingen enz.) met een invloed op de betekenis, stijl of bedoeling ervan.

(91) De tweede uitzondering is van toepassing als een AI-systeem de door de gebruiksverantwoordelijke verstrekte inputdata of de semantiek daarvan niet substantieel wijzigt. Een wijziging moet als substantieel worden beschouwd indien de inputdata of de semantiek daarvan aanzienlijk zijn gemanipuleerd door het AI-systeem tijdens het genereren van output. Dit moet gebaseerd zijn op een beoordeling van relevante factoren, zoals de vorm, het type mediacontent, de stijl en veranderingen in de content die van invloed zijn op de betekenis, stijl of bedoeling ervan. Of dat zo is, moet per geval worden beoordeeld.

(92) Indien een AI-systeem zowel kan worden gebruikt om content te genereren of manipuleren voor niet-substantiële geringe wijzigingen van inputdata, is de transparantieverplichting uit hoofde van artikel 50, lid 2, van de AI-verordening niet van toepassing op de content die op een geringe, niet-substantiële wijze wordt gewijzigd.

Voorbeelden van standaardbewerking en geringe wijzigingen waarvoor de uitzondering van artikel 50, lid 2, van de AI-verordening geldt

— Grammaticale correcties en spellingcontrole, taalkundige en kleine stilistische verbeteringen die geen invloed hebben op de inhoud, betekenis, stijl of boodschap van de tekst, door AI gegenereerde vertalingen van tekst.

— Formattering, formaatconversies, technische compressie, ruisonderdrukking of -verwijdering ter verbetering van de duidelijkheid, zonder de betekenis of de inhoud van de content te wijzigen.

— Kleine bijsnijdingen, kleine kleuraanpassingen of -correcties, lichter of donker maken, scherper maken met het oog op meer helderheid of andere standaardbewerkingen en technische correcties.

— Verwijdering van stofvlekken veroorzaakt door een vuile lens of sensor, verwijdering van rode ogen veroorzaakt door flitsfotografie; het verwijderen en verbergen van achtergronden die zichtbaar zijn in het oorspronkelijke bestand, pixeleren of vervagen van gezichten.

— Schaalaanpassing van een videoclip, compressie van het dynamische bereik en egalisatie; beperkte videostabilisatie.

— Standaardaanpassingen van kleur en contrast, kleine aanpassingen aan de afspeelsnelheid, kleine correcties om de horizon van een beeld recht te trekken, het toepassen van pixelfilters om bepaalde delen van een beeld te vergroten, het toepassen van kleurenkaarten op beelden in grijswaarden, het toevoegen van randen in beelden, het omzetten van een zwart-witbeeld of -video in kleur en omgekeerd.

— Pixelvulling voor aanpassing van de beeldverhouding of verbeterde helderheid, automatische overgangsfragmenten en andere niet-inhoudelijke bewerkingen voor esthetische en technische doeleinden.

— Technische verwerking of reconstructie van medische beelden in medische apparatuur (bv. annotaties, kaders, contouren, cirkels, warmtekaarten).

— Transcripties van gesprekken; door AI gegenereerde content die alleen authentieke menselijke input transformeert door middel van ondersteunende technologieën zodat personen met een handicap kunnen communiceren (bv. ondersteunende en alternatieve communicatie (AAC) of customized neural voices (CNV)), aangezien die de betekenis van de content niet semantisch veranderen.

Voorbeelden van semantische wijzigingen die markering vereisen op grond van artikel 50, lid 2, van de AI-verordening

— Door AI gegenereerde samenvattingen van tekst; het parafraseren of herschrijven van tekst waarbij de stijl, structuur en betekenis zodanig worden gewijzigd dat het meer inhoudt dan louter grammaticale en kleine stilistische correcties.

— Het verwijderen, vervangen of invoegen van voorwerpen of personen in bestaande beelden en video's, waardoor de betekenis en inhoud van de content verandert; vervanging of ingrijpende aanpassing van een gezicht.

— Synthese van realistische spraak in de stem van een specifieke persoon of het genereren van realistische video's waarin gebeurtenissen worden weergegeven die niet hebben plaatsgevonden; wijziging van de lichaamsvorm of huidskleur van een persoon; extreme oplichting, verduistering, kleur- en contrastveranderingen waardoor de betekenis, bedoeling en boodschap van de content veranderen.

— Het maken van samengestelde beelden of videoclips waarbij de weergave van personen, voorwerpen, gebeurtenissen of feiten wordt gewijzigd, of elke andere wezenlijke wijziging van de content.

(93) De derde uitzondering is van toepassing op generatieve AI-systemen die bij wet zijn toegestaan om synthetische content te genereren of te manipuleren om strafbare feiten op te sporen, te voorkomen, te onderzoeken of te vervolgen. De verduidelijkingen in de punten 46 tot en met 48 zijn voor deze uitzondering ook van toepassing.

4.4. **Wisselwerking met andere rechtshandelingen van de Unie**

(94) De markerings en detectieverplichtingen van artikel 50, lid 2, van de AI-verordening zijn gericht op de manier waarop de content is gecreëerd en de kunstmatige oorsprong ervan, niet op wie de content heeft gecreëerd. De oplossingen voor markering en detectie die door aanbieders van AI-systemen worden gebruikt, moeten daarom in overeenstemming zijn met het toepasselijke Unierecht inzake gegevensbescherming, waaronder de beginselen en verplichtingen inzake gegevensbescherming, zoals de verplichting van gegevensbescherming door ontwerp en door standaardinstellingen, de beginselen van minimale gegevensverwerking en de beveiligings- en vertrouwelijkheidsbeginselen. Met het oog op markering en detectie mag de informatie over de maker van de content bijvoorbeeld niet worden verwerkt en moeten alle andere verwerkte persoonsgegevens worden gewist zodra het doel van de detectie is bereikt, moeten de veiligheid en vertrouwelijkheid worden gewaarborgd, enz.

(95) In overweging 136 van de AI-verordening wordt voorts gewezen op het bijzondere belang van de transparantieverplichtingen op grond van artikel 50, leden 1, 2 en 4, van de AI-verordening om de doeltreffende uitvoering van de digitaledienstenverordening te vergemakkelijken. Dat geldt met name voor de verplichtingen van aanbieders van zeer grote onlineplatforms (VLOP's) of zeer grote onlinezoekmachines (VLOSE's) om systeemrisico's die kunnen voortvloeien uit de verspreiding van kunstmatig gegenereerde of gemanipuleerde content te identificeren en te beperken, en dan met name het risico op feitelijke of te verwachten negatieve effecten op democratische processen, het maatschappelijk debat en verkiezingsprocessen, onder meer van desinformatie (artikelen 34 en 35 van de digitaledienstenverordening). Met machineleesbare markeringen kunnen die aanbieders door AI gegenereerde of gemanipuleerde content detecteren en hiervoor passende labeling en openbaarmaking verstrekken, wat het risico op door AI gegenereerde desinformatie binnen hun diensten vermindert.

(96) De verplichtingen om door AI-systemen gegenereerde of gemanipuleerde content te markeren en te labelen op grond van artikel 50, leden 2 en 4, van de AI-verordening doen geen afbreuk aan de verplichting van artikel 16, lid 6, van de digitaledienstenverordening voor aanbieders van hostingdiensten om meldingen van illegale content te verwerken overeenkomstig artikel 16, lid 1, van de digitaledienstenverordening ([29]). In het bijzonder mogen markeringen of labels die worden aangebracht op door AI gegenereerde of gemanipuleerde content geen invloed hebben op de beoordeling van en de beslissing over de onwettigheid van de betreffende content op grond van andere regelgevingskaders. Die beoordeling moet uitsluitend worden verricht aan de hand van de regels inzake de rechtmatigheid van de content overeenkomstig die kaders ([30]). Als bijvoorbeeld een gelabelde deepfake-afbeelding wordt gesignaleerd als potentieel materiaal van seksueel misbruik van kinderen of als inbreukmakend op het merken- of auteursrecht, heeft het aanbrengen van een label geen invloed op de mogelijke onwettigheid van de content binnen die regelgevingskaders. Die beoordeling mag uitsluitend worden uitgevoerd op basis van de respectieve toepasselijke wetgeving (bv. strafrecht, merkenrecht of auteursrechtelijke regels).

(97) Omgekeerd kan de machineleesbare markering die overeenkomstig artikel 50, lid 2, van de AI-verordening in de content is ingebed, aanbieders van VLOP's en VLOSE's helpen bij het nakomen van hun verplichtingen uit hoofde van de artikelen 34 en 35 van de digitaledienstenverordening om systeemrisico's die kunnen voortvloeien uit de verspreiding van kunstmatig gegenereerde of gemanipuleerde content te identificeren, te analyseren en te beoordelen, met name het risico op feitelijke of te verwachten negatieve effecten op democratische processen, het maatschappelijk debat en verkiezingsprocessen, onder meer van desinformatie ([31]), en om redelijke, evenredige en doeltreffende maatregelen te nemen om die risico's te beperken, afgestemd op de vastgestelde specifieke systeemrisico's. Aanbieders van VLOP's en VLOSE's kunnen ook als aanbieders van AI-systemen worden aangemerkt indien ze voldoen aan de relevante voorwaarden van de AI-verordening wanneer ze AI-systemen in hun diensten integreren (zie ook punt 6.1.5).

(98) Aanbieders van onlineplatforms, onlinezoekmachines en andere actoren die content verspreiden binnen de waardeketen, worden in het bijzonder aangemoedigd om de door aanbieders overeenkomstig artikel 50, lid 2, van de AI-verordening ingevoerde markeringen te behouden en de herkenning daarvan mogelijk te maken, zodat de natuurlijke personen die aan de content worden blootgesteld, kunnen worden geïnformeerd over de oorsprong ervan.

5. **ARTIKEL 50, LID 3, VAN DE AI-VERORDENING: SYSTEMEN VOOR HET HERKENNEN VAN EMOTIES EN SYSTEMEN VOOR BIOMETRISCHE CATEGORISERING**

(99) In artikel 50, lid 3, van de AI-verordening wordt aan gebruiksverantwoordelijken van systemen voor het herkennen van emoties en systemen voor biometrische categorisering de verplichting opgelegd om natuurlijke personen die aan die systemen worden blootgesteld, te informeren over de werking van het systeem.

(100) In overweging 132 van de AI-verordening wordt toegelicht dat het doel van die verplichting erin bestaat te waarborgen dat natuurlijke personen (inclusief personen met een handicap) zich ervan bewust zijn dat ze worden blootgesteld aan systemen voor het herkennen van emoties en biometrische categorisering die inbreuk op hun privacy kunnen vormen. De verplichting geldt derhalve ongeacht of de personen in realtime aan dergelijke systemen worden blootgesteld of dat die achteraf worden gebruikt.

([29]) Zie ook overweging 11 van de AI-verordening.
([30]) Zie ook overweging 137 van de AI-verordening.
([31]) Overweging 136 van de AI-verordening.

5.1. **Belangrijkste componenten, begrippen en daarmee samenhangende transparantieverplichting(en) van artikel 50, lid 3, van de AI-verordening**

5.1.1. *Het begrip "systeem voor het herkennen van emoties"*

(101) In artikel 3, punt 39, van de AI-verordening wordt een "systeem voor het herkennen van emoties" gedefinieerd als "een AI-systeem dat is bedoeld voor het vaststellen of afleiden van de emoties of intenties van natuurlijke personen op basis van hun biometrische gegevens". De gedetailleerde toelichting van dit begrip in de richtsnoeren van de Commissie voor de classificatie van AI-systemen met een hoog risico, inclusief de daarin opgenomen praktische voorbeelden, is evenzeer relevant in het kader van artikel 50, lid 3, van de AI-verordening.

(102) Aangezien alle systemen voor het herkennen van emoties ook als systemen met een hoog risico worden geclassificeerd, moet de transparantieverplichting van artikel 50, lid 3, van de AI-verordening, samen met de overige waarborgen en eisen die gelden voor AI-systemen met een hoog risico, van toepassing zijn, tenzij dat op grond van artikel 5, lid 1, punt f), van de AI-verordening op de werkplek en in het onderwijs verboden is.

5.1.2. *Het begrip "systeem voor biometrische categorisering"*

(103) In artikel 3, punt 40, van de AI-verordening wordt een "systeem voor biometrische categorisering" gedefinieerd als "een AI-systeem dat is bedoeld voor het indelen van natuurlijke personen in specifieke categorieën op basis van hun biometrische gegevens, tenzij dit een aanvulling vormt op een andere commerciële dienst en om objectieve technische redenen strikt noodzakelijk is". De gedetailleerde toelichting van dit begrip in de richtsnoeren van de Commissie voor de classificatie van AI-systemen met een hoog risico, inclusief de specifieke elementen ervan en de erin opgenomen praktische voorbeelden, is evenzeer relevant in het kader van artikel 50, lid 3, van de AI-verordening.

(104) Belangrijk is dat artikel 50, lid 3, van de AI-verordening van toepassing is op alle systemen voor biometrische categorisering (bv. indeling naar leeftijd of gender op basis van biometrische gegevens), ongeacht of die al dan niet onder het toepassingsgebied van AI-systemen met een hoog risico in het kader van de AI-verordening vallen, tenzij dat op grond van artikel 5, lid 1, punt g), van de AI-verordening verboden is.

5.1.3. *De informatieverplichting uit hoofde van artikel 50, lid 3, van de AI-verordening*

(105) **De omvang van de te verstrekken informatie**: gebruiksverantwoordelijken moeten de betrokken natuurlijke personen informeren dat ze worden blootgesteld aan een systeem voor het herkennen van emoties of aan een systeem voor biometrische categorisering dat in gebruik is. In de AI-verordening wordt niet vereist dat informatie wordt verstrekt over de redenen voor het gebruik van het systeem, noch over de vraag of het systeem ook gegevens verwerkt voor andere doeleinden dan het herkennen van emoties of biometrische categorisering, hetgeen wel verplicht is op grond van het Unierecht inzake gegevensbescherming (zie punt 5.3.).

(106) **De personen aan wie de informatie is gericht:** in artikel 50, lid 3, van de AI-verordening is bepaald dat gebruiksverantwoordelijken alle natuurlijke personen die aan het gebruikte AI-systeem worden blootgesteld, inclusief kinderen, daarover moeten informeren.

(107) **De manier waarop informatie wordt verstrekt:** in de AI-verordening wordt geen specifieke manier voorgeschreven waarop de informatie moet worden verstrekt, behalve dat in artikel 50, lid 5, van de AI-verordening wordt bepaald dat de informatie op een duidelijke en te onderscheiden wijze en in overeenstemming met de toepasselijke toegankelijkheidsvoorschriften moet worden verstrekt (om tegemoet te komen aan de behoeften van natuurlijke personen met een handicap). Hoe natuurlijke personen die aan het systeem worden blootgesteld, moeten worden geïnformeerd, kan afhangen van de plaats van gebruik (bv. virtueel spelplatform, fysieke winkel, treinstation), de mogelijke geadresseerden (bv. kinderen, ouderen, personen met een handicap, klanten), de duur van de en de soort blootstelling (bv. langdurig, eenmalig, continu, met tussenpauzes) en het mogelijke bestaan van een verband tussen de gebruiksverantwoordelijke en de blootgestelde personen (bv. of er een bestaand communicatiekanaal met de persoon is). Afhankelijk van de specifieke situatie kan de informatie schriftelijk, door middel van standaardpictogrammen (ook als die elektronisch worden weergegeven), mondeling of door een combinatie van deze middelen worden verstrekt.

Voorbeelden van het informeren van natuurlijke personen

— Een centraal geplaatste pop-up met een welkomstbericht dat verschijnt voordat een computerspel wordt gestart, waarin wordt gemeld dat het gezicht van de speler wordt opgenomen om zijn of haar emoties vast te leggen.

— Een zichtbare mededeling bij elke mogelijke ingang van een tentoonstellingsruimte, waarin wordt vermeld dat bij het betreden van de ruimte gezichtsopnamen van bezoekers worden gemaakt om hen aan een specifieke leeftijdsgroep toe te wijzen.

(108) **Het tijdschema voor het verstrekken van de informatie**: de informatie moet uiterlijk op het moment van de eerste interactie met of blootstelling aan het AI-systeem worden verstrekt overeenkomstig artikel 50, lid 5, van de AI-verordening. In voorkomende gevallen kan de vereiste informatie ook vooraf worden verstrekt.

5.2. Uitzondering met het oog op de rechtshandhaving

(109) Deze informatieverplichting geldt niet voor systemen voor het herkennen van emoties en biometrische categorisering die bij wet zijn toegestaan om strafbare feiten op te sporen, te voorkomen of te onderzoeken, met inachtneming van passende waarborgen voor de rechten en vrijheden van derden en overeenkomstig het Unierecht. De verduidelijkingen in de punten 46 tot en met 48 met betrekking tot deze uitzondering zijn in dit geval ook relevant, zij het met één verschil. Terwijl de overige uitzonderingen in artikel 50 van de AI-verordening een uitdrukkelijke wettelijke machtiging vereisen voor het niet-transparante gebruik van de AI-systemen, wordt dergelijk gebruik op grond van artikel 50, lid 3, van de AI-verordening mogelijk gemaakt als dat is toegestaan krachtens de rechtsbepalingen over de bevoegdheden van rechtshandhavingsinstanties om systemen voor het herkennen van emoties en biometrische categorisering te gebruiken voor het opsporen, voorkomen of onderzoeken van strafbare feiten zonder de betrokken personen daarover te informeren, met inachtneming van passende waarborgen.

5.3. Wisselwerking met andere rechtshandelingen van de Unie

(110) De informatieverplichting uit hoofde van artikel 50, lid 3, van de AI-verordening maakt het gebruik van een systeem voor het herkennen van emoties en biometrische categorisering op zich niet rechtmatig. Evenmin rechtvaardigt ze ingrijpend of discriminerend gebruik dat onwettig en verboden zou kunnen zijn op grond van artikel 5 van de AI-verordening of ander Unierecht. Gebruiksverantwoordelijken moeten voldoen aan de informatieverplichting van artikel 50, lid 3, van de AI-verordening, naast alle toepasselijke informatie-eisen uit hoofde van het Unierecht inzake gegevensbescherming. In bepaalde gevallen kunnen gebruiksverantwoordelijken overwegen kennisgevingen overeenkomstig artikel 50, lid 3, van de AI-verordening toe te voegen aan de informatie die op grond van het Unierecht inzake gegevensbescherming aan betrokkenen wordt verstrekt.

6. ARTIKEL 50, LID 4: LABELEN VAN DEEPFAKES EN BEPAALDE TEKSTPUBLICATIES

(111) Artikel 50, lid 4, van de AI-verordening bevat twee afzonderlijke verplichtingen voor gebruiksverantwoordelijken van generatieve AI-systemen die duidelijke en te onderscheiden openbaarmaking vereisen van i) deepfakes en ii) door AI gegenereerde of gemanipuleerd tekst die wordt gepubliceerd om het publiek te informeren over aangelegenheden van algemeen belang, behalve in welbepaalde gevallen. Die labelingsverplichtingen gelden naast en onverminderd de verplichtingen voor machineleesbare markering en detectie op grond van artikel 50, lid 2, van de AI-verordening die van toepassing zijn op aanbieders van AI-systemen die synthetische beelden, video-, audio- of tekstcontent genereren of manipuleren.

6.1. Belangrijkste componenten, concepten en daarmee samenhangende transparantieverplichting(en) voor deepfakes op grond van artikel 50, lid 4, van de AI-verordening

(112) Artikel 50, lid 4, eerste alinea, van de AI-verordening is van toepassing indien aan verschillende voorwaarden is voldaan:

i) het systeem moet een AI-systeem zijn;

ii) dat door gebruiksverantwoordelijken voor professionele doeleinden worden gebruikt (zie de punten 2.3 en 2.4.1);

iii) om beeld-, audio- of videocontent te genereren of manipuleren dat een deepfake vormt;

iv) en het AI-systeem mag niet vallen onder de uitzondering voor gebruik dat bij wet is toegestaan om strafbare feiten op te sporen, te voorkomen, te onderzoeken of te vervolgen.

6.1.1. *Het begrip "deepfake"*

(113) In artikel 3, punt 60, van de AI-verordening wordt onder een "deepfake" verstaan: "door AI gegenereerd of gemanipuleerd beeld-, audio- of videomateriaal dat een gelijkenis vertoont met bestaande personen, voorwerpen, plaatsen, entiteiten of gebeurtenissen, en door een persoon ten onrechte voor authentiek of waarheidsgetrouw zou worden aangezien". Het verschil tussen manipuleren (of "bewerken") en genereren is verduidelijkt in punt 4.1. Ter verduidelijking van het begrip deepfake wordt hieronder nader ingegaan op vier cumulatieve criteria uit die definitie: i) gelijkenis met ii) bestaande iii) personen, voorwerpen, plaatsen, entiteiten of gebeurtenissen iv) die door een persoon ten onrechte voor authentiek of waarheidsgetrouw zou worden aangezien.

i) **Gelijkenis:** in artikel 3, punt 60, van de AI-verordening is bepaald dat door AI gegenereerd of gemanipuleerd beeld-, audio- of videomateriaal een "gelijkenis" moet vertonen met bestaande onderwerpen om als een deepfake te worden beschouwd. In overweging 134 van de AI-verordening wordt daaraan toegevoegd toe dat een dergelijke gelijkenis "aanzienlijk" moet zijn. Door AI gegenereerd of gemanipuleerd beeld-, audio- of videomateriaal vertoont een aanzienlijke gelijkenis met een persoon indien er sprake is van een hoge mate van gelijkenis tussen de deepfake-content en het onderwerp (waaronder elk herkenbaar element ervan) dat door de deepfake wordt gesimuleerd. De content hoeft niet identiek te zijn aan het onderwerp. In hoeverre de gelijkenis aanzienlijk is, moet per geval worden beoordeeld door de gebruiksverantwoordelijke op basis van een objectieve vergelijking tussen het gesimuleerde onderwerp en de deepfake-content, waarbij onder andere wordt gekeken naar de mate waarin karakteristieke of onderscheidende kenmerken in de deepfake tot uiting komen.

ii) **Bestaand:** In artikel 3, punt 60, van de AI-verordening is bepaald dat door AI gegenereerd of gemanipuleerd beeld-, audio- of videomateriaal een gelijkenis moet vertonen met "bestaande" onderwerpen om als een deepfake te worden beschouwd. Om het doel van artikel 50, lid 4, van de AI-verordening om de risico's van imitatie, misleiding, desinformatie, manipulatie en fraude te verminderen daadwerkelijk te verwezenlijken, houdt het kenmerk "bestaand" in dat het door AI gegenereerde of gemanipuleerde beeld-, audio- of videomateriaal een gelijkenis moet vertonen met realistische onderwerpen (bv. personen, voorwerpen, plaatsen). Daarom volstaat het dat gesimuleerde personen, voorwerpen, plaatsen, entiteiten of gebeurtenissen lijken op iemand die of iets dat in werkelijkheid bestaat, waarschijnlijk zou kunnen bestaan of zou kunnen hebben bestaan, om als een deepfake te worden beschouwd, aangezien dergelijke content een belangrijke bron van misleiding, manipulatie en desinformatie kan zijn. Daarentegen worden gesimuleerde personen, voorwerpen, plaatsen, entiteiten of gebeurtenissen die bijvoorbeeld de natuurwetten of de wetten van de fysica doorbreken of levensvormen weergeven die in de biologie niet algemeen aanvaard zijn (zoals bv. mensen die zonder mechanische hulpmiddelen vliegen, draken of olifanten die auto's besturen) en die geen potentieel hebben om te misleiden, als onrealistisch beschouwd en vallen die derhalve buiten het toepassingsgebied van de transparantieverplichting.

iii) **Personen, voorwerpen, plaatsen, entiteiten of evenementen:** in artikel 3, punt 60, van de AI-verordening is bepaald dat door AI gegenereerd of gemanipuleerd beeld-, audio- of videomateriaal een gelijkenis moet vertonen met bestaande "personen, voorwerpen, plaatsen, entiteiten of gebeurtenissen" om als een deepfake te worden beschouwd. Met "personen" wordt bedoeld: realistische, menselijke wezens (inclusief digitale replica's van echte personen, realistische, door AI gegenereerde menselijke avatars of figuren, en persoonlijke kenmerken of uitingen, zoals uiterlijk, stem, gedrag, optredens enz.). Met "voorwerpen" wordt bedoeld: realistische, levenloze materiële zaken, waaronder gebouwen, kunstwerken, machines, consumptiegoederen enz. Met "plaatsen" wordt bedoeld: realistische locaties. Met "entiteiten" wordt bedoeld: realistische, niet-menselijke maar levende wezens, waaronder dieren of andere biologische levensvormen. Met "gebeurtenissen" wordt bedoeld: realistische scènes of situaties waarin personen, voorwerpen, plaatsen en entiteiten kunnen voorkomen (bv. het uitbeelden van historische gebeurtenissen of de weergave van professionele of consumentendiensten).

iv) **Door een persoon ten onrechte voor authentiek of waarheidsgetrouw zou worden aangezien:** ten vierde is in artikel 3, punt 60, van de AI-verordening bepaald dat door AI gegenereerd of gemanipuleerd beeld-, audio- of videomateriaal dat een gelijkenis vertoont met bestaande personen, voorwerpen, plaatsen, entiteiten of gebeurtenissen, "door een persoon ten onrechte voor authentiek of waarheidsgetrouw zou worden aangezien". Dat criterium heeft betrekking op het wezenlijke kenmerk van deepfake-content en het vermogen daarvan om een persoon mogelijk te misleiden of in verwarring te brengen ten aanzien van de authenticiteit of waarheidsgetrouwheid van de content ([32]). Authenticiteit van content heeft betrekking op de vraag of de content daadwerkelijk is wat die pretendeert te zijn wat betreft de bron of het totstandkomingsproces (waaronder bv. de betrokkenheid van echte mensen of dieren en hun daadwerkelijke gedrag of handelingen, het daadwerkelijke uiterlijk of gebruik van voorwerpen of de levering van diensten, en het nauwkeurige verloop of de gang van zaken van een gebeurtenis). Waarheidsgetrouwheid heeft betrekking op de echtheid van de content (bv. de feitelijke juistheid van de weergaven in de deepfake).

([32]) Daarom moet dit criterium los worden gezien van en worden onderscheiden van het begrip misleiding als bedoeld in de richtlijn oneerlijke handelspraktijken (https://eur-lex.europa.eu/eli/dir/2005/29/oj).

(114) Of content “door een persoon ten onrechte voor authentiek of waarheidsgetrouw wordt aangezien”, moet in zijn geheel worden beoordeeld, waarbij rekening moet worden gehouden met de mate van gelijkenis, de mogelijke inhoudelijke boodschap van de content, de beoogde en voorzienbare gebruikscontexten, de omgeving waarin de content wordt gepresenteerd en de samenstelling en verwachtingen van de doelgroep en het redelijkerwijs te verwachten publiek. Die beoordeling is echter objectief en vereist niet dat de gebruiksverantwoordelijke voornemens is de natuurlijke personen die aan de content worden blootgesteld, te misleiden of te bedriegen om van deepfake te kunnen spreken. Wat de context van het gebruik en de verwachtingen van de doelgroep betreft, betekent dit dat, indien de content wordt gebruikt in specifieke gebruikscontexten waarin de doelgroep en het redelijkerwijs te verwachten publiek niet verwacht dat de content authentiek of waarheidsgetrouw is, de door AI gegenereerde of gemanipuleerde content niet *ten onrechte* de indruk mag wekken authentiek of waarheidsgetrouw te zijn, zelfs als die in werkelijkheid niet-authentiek of niet-waarheidsgetrouw is. In dergelijke gevallen wordt niet voldaan aan het vierde criterium van de definitie van een deepfake. Zo is het bijvoorbeeld onwaarschijnlijk dat het door AI genereren of manipuleren van achtergrondbeelden, speciale effecten of technische voor- en nabewerking als onderdeel van gebruikelijke filmproductieprocessen ertoe leidt dat de content bij de doelgroep *ten onrechte* de indruk wekt authentiek of waarheidsgetrouw te zijn. Omgekeerd zal het gebruik van AI voor het genereren of manipuleren van essentiële elementen die van invloed zijn op de perceptie van de doelgroep waarschijnlijk wel een dergelijke invloed hebben (zoals volledig door AI gegenereerde acteurs, digitale replica’s van echte of overleden acteurs, verjonging, gesimuleerde optredens en ook delen daarvan, of een niet-authentieke of niet-waarheidsgetrouwe weergave van personen, voorwerpen, plaatsen of gebeurtenissen in documentaires). Wat beeld- en videocontent betreft, wordt erkend dat een hoge mate van fotorealisme de kans vergroot dat dergelijke content als een deepfake moet worden beschouwd (aangezien die waarschijnlijk op bestaande onderwerpen lijkt), maar fotorealisme op zich is niet doorslaggevend voor de beoordeling. Wat nog moet worden beoordeeld, is of een persoon door dergelijke content misleid of in verwarring gebracht kan worden ten aanzien van de authenticiteit of waarheidsgetrouwheid ervan, waarbij onder meer rekening moet worden gehouden met de beoogde en te verwachten context waarin de content wordt gebruikt.

(115) In tegenstelling tot artikel 50, lid 1, van de AI-verordening mag de beoordeling of een valse weergave voor een persoon authentiek of waarheidsgetrouw lijkt, niet worden gebaseerd op een hypothetische “gemiddelde” persoon van wie wordt verwacht dat die aan de content wordt blootgesteld (zie punt 3.2.1), maar moet naar behoren rekening worden gehouden met de mogelijke uiteenlopende samenstelling van het redelijkerwijs te verwachten publiek dat aan de deepfake-content kan worden blootgesteld. Het is van groot belang om terdege rekening te houden met de samenstelling van het publiek, vooral als redelijkerwijs te verwachten is dat de deepfake-content op enig moment onder de aandacht kan komen van kinderen, ouderen of andere groepen personen met een lagere digitale en AI-geletterdheid of een lager algemeen kennisniveau, aangezien die wellicht gemakkelijker misleid of in de war gebracht kunnen worden wat betreft de authenticiteit of waarheidsgetrouwheid van de content. In dergelijke gevallen kan het feit dat de content voor dat deel van het te verwachten publiek de indruk wekt authentiek of waarheidsgetrouw te zijn, voldoende zijn om de content als een deepfake aan te merken. Van gebruiksverantwoordelijken wordt echter niet verwacht dat ze rekening houden met de mogelijke verdere verspreiding van deepfake-content door derden buiten het redelijkerwijs te verwachten publiek door de gebruiksverantwoordelijke, rekening houdend met de beoogde distributiekanalen (bv.: als deepfake-content uitsluitend wordt getoond op een deel van een website dat alleen voor abonnees toegankelijk is of als onderdeel van een bedrijfsnieuwsbrief, betekent dat niet dat de gebruiksverantwoordelijke standaard moet uitgaan van brede toegankelijkheid voor het publiek).

(116) In dezelfde geest kan door AI ondersteunde manipulatie van onbeduidende inhoudelijke of technische aspecten van reeds bestaande content van gering belang zijn voor de beoordeling door een persoon van de authenticiteit of waarheidsgetrouwheid van de content, waardoor de resulterende content geen deepfake wordt. Dat kan bijvoorbeeld bestaan uit het bewerken van achtergronddetails (zoals het verwijderen van voorbijgangers), het aanpassen van de belichting, het aanpassen van audioparameters, kleurcorrectie, ruisonderdrukking of -verwijdering, het verbeteren van de toegankelijkheid of bestandscompressie, cosmetische aanpassingen en verbeteringen. Of de manipulatie onder deze categorie valt, hangt af van de context en de impact op de perceptie van de betrokkenen ten aanzien van de authenticiteit of waarheidsgetrouwheid van de content in het specifieke geval. Zo zullen bijvoorbeeld door AI aangestuurde kleurcorrectie, uitbreidingen van de achtergrond van bestaande content, aanpassingen of vervangingen van achtergronden met duidelijk esthetische doeleinden, composities en opstellingen van bestaande producten of het aanpassen van de schaal van beelden in productadvertenties of op verpakkingen waarschijnlijk slechts een geringe invloed hebben op de perceptie van een persoon ten aanzien van de authenticiteit en waarheidsgetrouwheid van de advertentie en het product. Omgekeerd kan substantiële, door AI aangestuurde bewerking van details op de achtergrond van journalistieke beelden die verder gaat dan de gebruikelijke technische en redactionele werkwijzen, een negatieve invloed hebben op de perceptie van een persoon ten aanzien van de authenticiteit en waarheidsgetrouwheid van de content.

Voorbeelden van deepfakes uit hoofde van artikel 3, punt 60, van de AI-verordening

— Door AI gemanipuleerd beeld van een scène met twee echte profvoetballers voor een gebouw dat op een voetbalstadion lijkt.

— Door AI gegenereerde audio waarin de stemmen van de vaste presentatoren van een krantenpodcast zijn gekloond en waarin een gastgesprekspartner enkele nieuwsfeiten bespreekt.

— Door AI gegenereerde video van een persoon die lijkt op een politicus en een toespraak houdt voor een publiek.

— Door AI gegenereerde video met een door AI gegenereerd beeld van een beroemde influencer voor reclame of promotie.

— Door AI gegenereerde video met een realistische synthetische avatar van de CEO van een bedrijf die werknemers feliciteert met hun werk en de bedrijfsresultaten van het afgelopen jaar.

— Door AI gegenereerd beeld van een product in reclame of verpakkingen dat van invloed kan zijn op de perceptie van het publiek en misleidend kan zijn wat betreft het werkelijke uiterlijk, de werkelijke kenmerken of het werkelijke gebruik van het product (bv. door het product anders, aantrekkelijker of kwalitatiever te laten lijken dan het in werkelijkheid is).

Voorbeelden die geen deepfakes zijn uit hoofde van artikel 3, punt 60, van de AI-verordening

— Door AI gegenereerd beeld van een scène met een sfinx die over de Eiffeltoren vliegt.

— Door AI gegenereerde video waarin muizen in menselijke taal discussiëren over de beste kaassoort in het kader van een reclamecampagne voor een kaasfabrikant.

— AI-stemreproductie voor fictieve karakters (bv. voor audioboeken, spellen of animatie) wanneer er geen sprake is van misleiding met betrekking tot de identiteit van de vertellers.

— Door AI gemanipuleerde radio-uitzending waarbij technische audioparameters worden aangepast (bv. normalisatie van het volume, ruisonderdrukking, audiocompressie) zonder de daadwerkelijk door de sprekers uitgesproken woorden of hun manier van spreken te wijzigen.

— Door AI gegenereerde cartoon van een bestaand beeld dat een historische gebeurtenis weergeeft.

— Video van een presentator van een wetenschappelijk tv-programma die door een bergachtig gebied wandelt terwijl hij spreekt over het terugtrekken van gletsjers, dat wordt geïllustreerd aan de hand van een door AI gemanipuleerde achtergrondanimatie die laat zien hoe de gletsjers zich hebben teruggetrokken van hun historische niveau naar hun huidige niveau.

— Door AI gegenereerde fictieve omgevingen (bv. fictieve bossen en kastelen) in videospellen.

— Films waarin echte bestaande actoren (niet gemanipuleerd door AI) spelen tegen een door AI gegenereerde achtergrond (bv. een fictief stuk prairie, historische gebouwen van oude steden enz.).

— Een echt product (bv. een auto) dat in een advertentie wordt getoond tegen een door AI gegenereerde achtergrond en omgeving, mits het niet waarschijnlijk is dat de advertentie het publiek misleidt ten aanzien van de werkelijke weergave van het product en de kenmerken en het gebruik ervan.

6.1.2. *De verplichting tot openbaarmaking van artikel 50, lid 4, eerste alinea, van de AI-verordening*

(117) In artikel 50, lid 4, eerste alinea, van de AI-verordening wordt de gebruiksverantwoordelijken van AI-systemen die beeld-, audio- of videocontent genereren of manipuleren die een deepfake vormt, verplicht om "bekend [te maken] dat de [deepfake-content] kunstmatig is gegenereerd of gemanipuleerd". In overweging 134 van de AI-verordening

wordt verduidelijkt dat gebruiksverantwoordelijken die AI-systemen gebruiken om deepfake-content te genereren of te manipuleren, klaar en duidelijk kenbaar moeten maken dat de content kunstmatig is gecreëerd of gemanipuleerd door de AI-output als zodanig aan te merken en de kunstmatige oorsprong ervan bekend te maken. De methoden voor labeling of openbaarmaking die worden toegepast overeenkomstig artikel 50, lid 4, eerste alinea, van de AI-verordening moeten begrijpelijk en waarneembaar zijn voor natuurlijke personen (bv. door middel van zicht- of hoorbare labels), zonder dat ze daarvoor een beroep hoeven te doen op specifieke technische hulpmiddelen of speciale handelingen hoeven te verrichten. Daarom kunnen gebruiksverantwoordelijken geen gebruik maken van de machineleesbare markering die door de aanbieder in de content is ingebed op grond van artikel 50, lid 2, van de AI-verordening, aangezien die markeringen niet onmiddellijk duidelijk en te onderscheiden zijn voor de natuurlijke personen die aan de deepfake-content worden blootgesteld.

(118) In het licht van artikel 50, lid 7, van de AI-verordening kunnen gebruiksverantwoordelijken zich houden aan een praktijkcode die overeenkomstig artikel 50, lid 7, van de AI-verordening als toereikend is beoordeeld om de naleving van de verplichting tot bekendmaking aan te tonen. Zie punt 8.1 voor meer informatie over de effecten van de naleving van een dergelijke code en het toezicht op andere gebruiksverantwoordelijken die zich niet aan een code houden.

6.1.3. *Transparantie van artistieke, creatieve, satirische, fictieve of analoge deepfake-content*

(119) Artikel 50, lid 4, eerste alinea, van de AI-verordening voorziet in een versoepelde transparantieverplichting voor deepfakes die deel uitmaken van kennelijk artistieke, creatieve, satirische, fictieve of analoge werken of programma's. Voor dergelijke content is de transparantieverplichting beperkt tot de openbaarmaking van de deepfake op een passende wijze die de weergave of het genot van het werk niet belemmert.

(120) **Relevante categorieën content**: om in aanmerking te komen voor de versoepelde transparantieverplichting, moet de deepfake (beeld, audio of video) behoren tot ten minste een van de volgende categorieën content: artistieke, creatieve, satirische, fictieve of analoge werken of programma's. Of een deepfake onder een of meer van deze categorieën valt, moet per geval door de gebruiksverantwoordelijke worden beoordeeld. Voor de toepassing van artikel 50, lid 4, van de AI-verordening wordt onder die categorieën het volgende verstaan:

— met "artistieke werken" wordt bedoeld: werken die zijn gecreëerd met de bedoeling als kunst beschouwd te worden, zoals muziek, cinematografische werken en beeldende kunst;

— met "creatieve werken" wordt bedoeld: werken waarbij creatieve keuzes een rol spelen. Werken die hoofdzakelijk zijn ingegeven door functionele of technische overwegingen kunnen niet als creatief worden beschouwd;

— met "satirische werken" wordt bedoeld: werken die tot doel hebben de samenleving, de politiek, het bedrijfsleven of publieke figuren te bekritiseren door middel van humoristische technieken (waaronder ironie, sarcasme, spot, pastiche enz.);

— met "fictieve werken" wordt bedoeld: culturele werken waarin personen, voorwerpen, plaatsen, entiteiten of gebeurtenissen voorkomen in een fictieve, maar levensechte setting;

— met "analoge werken" wordt bedoeld: werken die belangrijke kenmerken of vergelijkbare expressieve of functionele eigenschappen delen met de bovengenoemde categorieën, maar die niet duidelijk onder één daarvan vallen.

(121) De bovengenoemde categorieën content kunnen ook van toepassing zijn op programma's die moeten worden beschouwd als afzonderlijke onderdelen van een door een aanbieder van mediadiensten opgesteld(e) programma of catalogus en die qua vorm en content vergelijkbaar zijn met televisie-uitzendingen ([33]). De verwijzing naar "werken" in de context van artikel 50, lid 4, van de AI-verordening heeft geen invloed op de vraag of de content in aanmerking komt voor bescherming uit hoofde van het Unierecht inzake auteursrecht en naburige rechten.

([33]) Zie in dezelfde zin artikel 1, punt b), van de richtlijn audiovisuele mediadiensten.

(122) **Duidelijkheid:** dat een deepfake onder een (of meer) van de bovenstaande categorieën content valt, moet duidelijk zijn voor de natuurlijke personen die eraan worden blootgesteld. Dat houdt in dat de categorieën strikt moeten worden geïnterpreteerd, wat wordt gerechtvaardigd doordat een lichtere openbaarmakingsregeling van toepassing is vanwege de mogelijke gevolgen voor de vrijheid van meningsuiting en de vrijheid van kunsten en wetenschappen, terwijl de risico's van desinformatie, manipulatie en misleiding nog steeds toereikend moeten worden aangepakt. Daarom valt content waarvan de aard voor het publiek mogelijk onduidelijk of dubbelzinnig is, buiten het toepassingsgebied van deze lichtere transparantieregeling. Relevante factoren om de kennelijke aard van artistieke, creatieve, satirische of fictieve content te beoordelen, zijn onder meer: i) of de content formaten of stijlen weergeeft die kenmerkend zijn voor de categorieën content (bv. ironie of overdrijving voor satirische werken, bepaalde kunststijlen, onderscheidende thema's, technieken of esthetiek enz.); ii) de context waarin de content wordt gepresenteerd (bv. of het platform, het medium of de plaats waar de content gepresenteerd, beschikbaar gesteld of verspreid wordt, verband houdt met artistiek, creatief, satirisch of fictief gebruik); en iii) verwachtingen van het publiek (bv. een film, spelomgeving of virtual-realityscène). Bovendien sluit de voorwaarde dat content "kennelijk" in een van de bovengenoemde categorieën content moet vallen, content uit van het toepassingsgebied van die categorieën content als de aard ervan uitsluitend informatief of commercieel is en als zodanig herkenbaar is (bv. nieuwsverslaggeving). In dit verband kunnen sommige soorten content (bv. advertenties of documentaires) met deepfakes in bepaalde, specifieke situaties worden beschouwd als duidelijk creatieve of fictieve werken, maar in andere niet, aangezien de beoordeling specifiek is voor het geval. Wanneer in een deepfake meerdere aarden worden gecombineerd (bv. informatief en creatief), moet de informatieve aard altijd de overhand hebben en moet aan de standaardeisen voor labeling worden voldaan.

(123) **Passende openbaarmaking die de weergave of het genot van het werk niet belemmert**: deepfakes die deel uitmaken van kennelijk artistieke, creatieve, satirische of fictieve werken of programma's zijn niet uitgesloten van de transparantieverplichting van artikel 50, lid 4, eerste alinea, van de AI-verordening. Gebruiksverantwoordelijken moeten nog steeds de AI-oorsprong van de content of de manipulatie ervan openbaar maken, maar kunnen dat doen op een passende wijze die de weergave of het genot van het werk niet belemmert. In overweging 134 van de AI-verordening wordt verduidelijkt dat een dergelijke passende openbaarmaking de normale exploitatie en het normale gebruik niet mag belemmeren, met behoud van het nut en de kwaliteit van het werk. Bij het bepalen welke maatregelen inzake openbaarmaking als passend moeten worden beschouwd, gaat het om een beoordeling per geval, waarbij gebruiksverantwoordelijken rekening kunnen houden met alle relevante factoren (waaronder de aard van het werk, het publiek, de context enz.). Gebruiksverantwoordelijken moeten in elk geval voldoen aan artikel 50, lid 5, van de AI-verordening (zie punt 7). In het licht van artikel 50, lid 7, van de AI-verordening kunnen gebruiksverantwoordelijken zich houden aan een praktijkcode die op grond van die bepaling als toereikend wordt beoordeeld, om de naleving van deze verplichting tot openbaarmaking aan te tonen. Zie punt 8.1 voor meer informatie over de effecten van de naleving van een dergelijke code en toezicht op andere gebruiksverantwoordelijken die zich niet aan een code houden.

(124) **Passende waarborgen voor de rechten en vrijheden van derden:** in overweging 134 van de AI-verordening wordt verduidelijkt dat bij de naleving van de versoepelde transparantieverplichting en het gebruik van artistieke, creatieve, satirische of fictieve deepfakes sprake moet zijn van "inachtneming van passende waarborgen voor de rechten en vrijheden van derden". Daarom moeten gebruiksverantwoordelijken ervoor zorgen dat de rechten en vrijheden van derden (waaronder bv. hun recht op de bescherming van persoonsgegevens of intellectuele-eigendomsrechten) naar behoren worden gewaarborgd en geëerbiedigd bij het creëren of publiceren van deepfakes, zelfs als de deepfake als artistiek, creatief, satirisch of fictief moet worden beschouwd. Een beroep op de versoepelde transparantieverplichting kan geen rechtvaardiging zijn voor het niet eerbiedigen van de grondrechten van personen of rechthebbenden uit hoofde van het Unierecht inzake intellectuele eigendom of inzake gegevensbescherming ([34]).

Voorbeelden van deepfakes die creatieve, satirische, fictieve of analoge werken zijn

— Artistiek/fictief werk: films (of filmtrailers) met door AI gegenereerde of gemanipuleerde deepfakes van verjongde, bestaande acteurs of digitale replica's van overleden acteurs of andere personen, beschikbaar in een bioscoop of op een streamingplatform.

— Artistiek/creatief werk: door AI gegenereerde muziek in elk genre dat lijkt op de individuele stijl van bestaande artiesten.

([34]) Dat geldt ook voor gepubliceerde teksten met als doel het publiek informeren over aangelegenheden van algemeen belang.

— Satirisch/fictief werk: door AI gemanipuleerd beeld van een bestaande politicus, waarin die is geplaatst in een scène die duidelijk bedoeld is om op humoristische wijze kritiek te leveren op bepaalde beleidsbeslissingen die deze persoon heeft genomen of gesteund.

— Analoog creatief/fictief werk: door AI gegenereerde gamebeelden met deepfakesimulaties van echte, bestaande personen.

Voorbeelden van deepfakes die geen artistieke, creatieve, satirische, fictieve of analoge werken zijn

— Door AI gemanipuleerde video in de stijl van een telewinkelzender, waarbij gebruik wordt gemaakt van deepfakesimulaties van mensen die reclame maken voor een product in een door AI gegenereerde scène waarin het functionele gebruik van het product door de gesimuleerde consumenten wordt weergegeven, met als doel de kijker over te halen het product te kopen.

— Door AI gegenereerde beelden van beroemdheden waarin wordt gesuggereerd dat die betrokken zijn bij gebeurtenissen die nooit hebben plaatsgevonden, zonder dat dat een fictief, satirisch of analoog doel dient.

— Door AI gemanipuleerde video met een realistische synthetische influencer die een gesponsord echt product test en zich uitsluitend richt op het weergeven van de productfunctionaliteiten.

— Door AI gegenereerde video's met realistische beelden van de Holocaust die worden gedeeld op openbaar toegankelijke sociale-mediaplatforms.

6.1.4. *Uitzondering in verband met rechtshandhaving*

(125) Indien het gebruik van een deepfake (bv. een door AI gegenereerde video waarin gebeurtenissen worden gesimuleerd die nooit hebben plaatsgevonden) bij wet is toegestaan om strafbare feiten op te sporen, te voorkomen, te onderzoeken of te vervolgen, zijn gebruiksverantwoordelijken vrijgesteld van de transparantieverplichting van artikel 50, lid 4, eerste alinea, van de AI-verordening. De verduidelijkingen in de punten 46 tot en met 48 voor een soortgelijke uitzondering in artikel 50, lid 1, van de AI-verordening zijn ook relevant voor deze bepaling.

6.1.5. *Wisselwerking met andere rechtshandelingen van de Unie*

(126) De transparantieverplichting van artikel 50, lid 4, eerste alinea, van de AI-verordening hangt samen met de verplichtingen van aanbieders van VLOP's en VLOSE's om systeemrisico's te identificeren en te beperken overeenkomstig de artikelen 34 en 35 van de digitaledienstenverordening, met name artikel 35, lid 1, punt k) ([35]). In die laatste bepaling worden, onder de mogelijke risicobeperkende maatregelen die aanbieders van VLOP's en VLOSE's kunnen nemen, maatregelen genoemd die "waarborgen dat informatie, of het nu gaat om gegenereerd of gemanipuleerd beeld-, audio- of videomateriaal dat een merkbare gelijkenis vertoont met bestaande personen, voorwerpen, plaatsen of andere entiteiten of gebeurtenissen, en door een persoon ten onrechte voor authentiek of waarheidsgetrouw wordt aangezien, door opvallende markeringen te onderscheiden is bij het presenteren ervan op hun online-interfaces, en daarnaast, de verstrekking van een gebruiksvriendelijke functionaliteit die afnemers van de dienst de mogelijkheid biedt dergelijke informatie aan te geven". Daarnaast kunnen aanbieders van VLOP's en VLOSE's ook een gebruiksvriendelijke functionaliteit verstrekken die afnemers van de dienst de mogelijkheid biedt dergelijke informatie aan te geven. Hoewel die bepaling van de digitaledienstenverordening lijkt op artikel 50, lid 4, van de AI-verordening, vormt ze daar een aanvulling op en zijn er twee belangrijke verschillen:

— **Verschillende materiële werkingssferen**: vergeleken met artikel 50, lid 4, van de AI-verordening heeft artikel 35, lid 1, punt k), van de digitaledienstenverordening betrekking op een breder scala aan content die ten onrechte voor authentiek of waarheidsgetrouw kan worden aangezien, ongeacht de technologieën die worden gebruikt om de content te creëren (d.w.z. de bepaling heeft betrekking op content die is gegenereerd of gemanipuleerd met behulp van AI-systemen of andere instrumenten).

([35]) Zie ook punt 4.4 voor meer informatie over de wisselwerking tussen artikel 50 van de AI-verordening en de digitaledienstenverordening.

— **Verschillende personele werkingssferen**: in artikel 50, lid 4, van de AI-verordening worden gebruiksverantwoordelijken van AI-systemen die deepfakes genereren of manipuleren verplicht om dergelijke content te labelen, terwijl artikel 35, lid 1, punt k), van de digitaledienstenverordening van toepassing is op aanbieders van VLOP's en VLOSE's die de content verspreiden of door middel van zoekopdrachten vindbaar maken. Gebruiksverantwoordelijken van AI-systemen die op grond van de AI-verordening in een VLOP zijn geïntegreerd, kunnen bijvoorbeeld als "afnemers van de dienst" in de zin van artikel 3, punt b), van de digitaledienstenverordening ([36]) worden beschouwd als ze het VLOP gebruiken om de met het respectieve AI-systeem gegenereerde content te verspreiden, terwijl het VLOP als de aanbieder van het AI-systeem kan worden beschouwd indien aan de voorwaarden van artikel 3, punt 3, van de AI-verordening is voldaan. Als aanbieders van VLOP's of VLOSE's labelinginstrumenten ter beschikking stellen van dergelijke gebruiksverantwoordelijken waarmee die hun deepfake-content kunnen labelen in overeenstemming met artikel 50, lid 4, van de AI-verordening (d.w.z. een duidelijke en te onderscheiden openbaarmaking van de AI-oorsprong), kunnen die gebruiksverantwoordelijken gebruikmaken van dergelijke instrumenten om hun transparantieverplichting van die bepaling na te komen in het kader van het/de gebruikte VLOP of VLOSE. Het aanbieden van een dergelijke functionaliteit doet geen afbreuk aan de verantwoordelijkheid van de gebruiksverantwoordelijken op grond van de AI-verordening om hun labelingsverplichtingen van artikel 50, lid 4, van de AI-verordening na te komen. Bovendien worden aanbieders van VLOP's en VLOSE's ook als gebruiksverantwoordelijken van een AI-systeem in de zin van de AI-verordening beschouwd indien ze onder hun gezag een AI-systeem gebruiken voor hun eigen professionele doeleinden (bv. het maken van beeldmateriaal voor gebruik als onderdeel van marketingmateriaal).

(127) Indien de deepfake informatie bevat over een direct of indirect identificeerbare (levende) natuurlijke persoon, moeten gebruiksverantwoordelijken als verwerkingsverantwoordelijken uit hoofde van het Unierecht inzake gegevensbescherming ook voldoen aan de respectieve verplichtingen uit hoofde van het Unierecht inzake gegevensbescherming ([37]). Dat omvat bijvoorbeeld het gebruik van een passende rechtsgrondslag voor de verwerking van de persoonsgegevens, informatie over de verwerking van gegevens met betrekking tot de afgebeelde betrokkenen enz.

(128) Gebruiksverantwoordelijken van AI-systemen die deepfakes genereren, moeten er bovendien voor zorgen dat ze voldoen aan het desbetreffende Unierecht inzake intellectuele-eigendomsrechten (zoals het merkenrecht van de Unie en het Unierecht inzake auteursrechten en naburige rechten) indien hun deepfake materiaal bevat dat door die wetgeving wordt beschermd. Het aanbrengen van een label of pictogram op een deepfake heeft op zich geen invloed op de vraag of die in aanmerking komt voor auteursrechtelijke bescherming. Dat moet worden beoordeeld overeenkomstig het toepasselijke auteursrecht van de Unie.

(129) Als bestaande personen in deepfakes worden afgebeeld, moeten gebruiksverantwoordelijken zich tevens houden aan de wettelijke regelingen ter bescherming van persoonlijkheidsrechten met betrekking tot persoonlijke afbeeldingen of stemmen. De transparantieverplichting van artikel 50, lid 4, eerste alinea, van de AI-verordening betekent niet dat door AI gegenereerde of gemanipuleerde deepfakes die schadelijk en onwettig zijn op grond van het toepasselijke Unie- of nationale recht (bv. misleidende reclame of strafrechtelijke bepalingen, zoals materiaal met seksueel misbruik van kinderen en intieme beelden waarvoor geen toestemming is gegeven) kunnen worden gegenereerd en verspreid (zie ook punt 4.4).

6.2. Belangrijkste componenten, concepten en daarmee samenhangende transparantieverplichting(en) voor door AI gegenereerde of gemanipuleerde tekst uit hoofde van artikel 50, lid 4, van de AI-verordening

(130) Artikel 50, lid 4, tweede alinea, van de AI-verordening is van toepassing indien aan verschillende voorwaarden is voldaan:

i) het systeem moet een AI-systeem zijn;

ii) het moet worden gebruikt om tekst te genereren of te manipuleren (zie punt 4.2.1) die wordt gepubliceerd om het publiek te informeren over aangelegenheden van algemeen belang;

iii) het gebruik van het AI-systeem valt niet onder een van de uitzonderingen wanneer 1) de door AI gegenereerde content een proces van menselijke toetsing of redactionele controle heeft ondergaan en wanneer een natuurlijke of rechtspersoon redactionele verantwoordelijkheid draagt voor de publicatie van de content of 2) het gebruik bij wet is toegestaan om strafbare feiten op te sporen, te voorkomen, te onderzoeken of te vervolgen.

([36]) Overweging 2 van de digitaledienstenverordening.

([37]) Zie voor een verdere verduidelijking de gezamenlijke richtsnoeren van de Commissie en het EDPB over de wisselwerking tussen de AI-verordening en de EU-wetgeving inzake gegevensbescherming (in voorbereiding).

6.2.1. *Tekst die wordt gepubliceerd om het publiek te informeren over aangelegenheden van algemeen belang*

(131) Ter verduidelijking van het toepassingsgebied van artikel 50, lid 4, tweede alinea, van de AI-verordening wordt hieronder ingegaan op drie elementen: i) gepubliceerde tekst ii) om het publiek te informeren iii) over aangelegenheden van algemeen belang.

i) **Gepubliceerde tekst:** de door AI gegenereerde of gemanipuleerde tekst moet worden gepubliceerd. Dit houdt in dat de tekst tegelijkertijd en/of achtereenvolgens toegankelijk moet zijn voor een niet nader bepaald, vrij groot aantal onderling niet-verwante, potentiële lezers, al dan niet tegen betaling (bv. via een abonnement) ([38]). Daarentegen wordt tekst als niet gepubliceerd beschouwd indien de toegang beperkt is tot specifieke personen die behoren tot een besloten, particuliere groep (bv. een kleine, besloten groep op een app voor directe berichten, of indien de groep te klein of onbeduidend is). Voorbeelden van teksten die als niet gepubliceerd worden beschouwd, zijn bv. persoonlijke correspondentie (voor professionele doeleinden) of interne teksten of mededelingen van organisaties (bv. publicaties op interne bedrijfsnetwerken).

ii) **Het publiek informeren:** de door AI gegenereerde of gemanipuleerde gepubliceerde tekst moet tot doel hebben het publiek te informeren. Dit houdt in dat de tekst bedoeld moet zijn om kennis, meningen of feiten over te brengen. Korte teksten die dat niet doen, kunnen daarentegen niet worden geacht het publiek te informeren.

iii) **Over aangelegenheden van algemeen belang:** de tekst moet het publiek informeren "over aangelegenheden van algemeen belang". In het algemeen wordt met dergelijke aangelegenheden bedoeld: aangelegenheden die van belang zijn voor de samenleving als geheel, hetzij op lokaal, nationaal, EU- of internationaal niveau, en die een openbaar debat of publieke toetsing verdienen. In dat verband moeten teksten worden beschouwd als teksten met betrekking tot aangelegenheden van algemeen belang indien ze betrekking hebben op onderwerpen als politiek, democratische processen, openbaar bestuur, openbare dienstverlening, rechtspraak, wetshandhaving, de bescherming van grondrechten, openbare veiligheid, volksgezondheid, milieubescherming, consumentenveiligheid en elke economische, financiële, politieke, wetenschappelijke of culturele ontwikkeling die een relevant onderwerp van publiek debat kan zijn. Zaken die als van algemeen belang kunnen worden beschouwd, kunnen in de loop van de tijd en naar gelang de context veranderen.

Voorbeelden van tekst die wordt gepubliceerd om het publiek te informeren over aangelegenheden van algemeen belang uit hoofde van artikel 50, lid 4, van de AI-verordening en die moet worden gelabeld als kunstmatig gegenereerd of gemanipuleerd, tenzij de tekst een proces van menselijke toetsing en/of redactionele controle moet ondergaan (zie punt 6.2.3)

— Door AI gegenereerde samenvatting van een door mensen geschreven artikel op de website van een krant over een recent besluit van een gemeenteraad.

— Door AI gemanipuleerde delen van een artikel op een lifestyle-website waarin de effecten van verschillende voedingspatronen op een bepaalde ziekte bij vrouwen van middelbare leeftijd worden vergeleken.

— Door AI gemanipuleerde verslagen van beursgenoteerde ondernemingen die op de website van de betreffende onderneming zijn gepubliceerd en die informatie voor beleggers bevatten.

— Een door AI gegenereerd bericht op het socialemedia-account van een meteorologisch instituut waarin mensen worden gewaarschuwd voor een storm en waarin voorzorgsmaatregelen worden gegeven.

Voorbeelden van tekst die niet wordt gepubliceerd om het publiek te informeren over aangelegenheden van algemeen belang en buiten het toepassingsgebied van artikel 50, lid 4, van de AI-verordening vallen.

— Door AI gegenereerde fantasieromans.

— Door AI gemanipuleerde tekst die deel uitmaakt van de reclame of productbeschrijvingen van een bedrijf (met uitzondering van beweringen over bv. gezondheid, consumentenveiligheid of duurzaamheid).

— Nieuwssamenvatting gegenereerd door een chatbot die alleen beschikbaar is voor de gebruiker die de chatbot heeft aangestuurd.

— Door AI gemanipuleerde tekst door een consulent voor advies aan een klant over maatregelen die moeten worden genomen voor de naleving van de toepasselijke wetgeving.

([38]) Zie naar analogie C-89/04, Mediakabel, EU:C:2005:348, punt 33; C-192/04, Lagardère, EU:C:2005:475, punt 31; C-306/05, SGAE, EU:C:2006:764, punten 37-40; C-527/15, Stichting Brein, EU:C:2017:300, punten 43-44 en aldaar aangehaalde rechtspraak; zaak C-135/23, GEMA, EU:C:2024:526, punt 38 en aldaar aangehaalde rechtspraak.

6.2.2. *De verplichting tot openbaarmaking uit hoofde van artikel 50, lid 4, tweede alinea, van de AI-verordening*

(132) In artikel 50, lid 4, tweede alinea, van de AI-verordening wordt aan gebruiksverantwoordelijken van een AI-systeem dat tekst genereert of manipuleert die wordt gepubliceerd om het publiek te informeren over aangelegenheden van algemeen belang, de verplichting opgelegd bekend te maken dat de tekst kunstmatig is gegenereerd of gemanipuleerd. In overweging 134 van de AI-verordening wordt verduidelijkt dat, net als voor deepfakes, gebruiksverantwoordelijken van tekstpublicaties die binnen het toepassingsgebied van de bepaling vallen, klaar en duidelijk kenbaar moeten maken dat dergelijke content kunstmatig is gecreëerd of gemanipuleerd door de AI-output als zodanig aan te merken en de kunstmatige oorsprong ervan bekend te maken. Zoals vereist voor deepfakes, moeten de methoden voor labelen of openbaarmaking (waaronder disclaimers) die worden toegepast overeenkomstig artikel 50, lid 4, tweede alinea, van de AI-verordening ook duidelijk en waarneembaar zijn voor natuurlijke personen (bv. zicht- of hoorbare maatregelen), zonder dat ze een beroep hoeven te doen op bepaalde technische instrumenten of specifieke acties hoeven uit te voeren. In het licht van artikel 50, lid 7, van de AI-verordening kunnen gebruiksverantwoordelijken zich houden aan een praktijkcode die op grond van die bepaling als toereikend wordt beoordeeld, om de naleving van deze verplichting tot openbaarmaking aan te tonen. Voor nadere informatie over de effecten van de naleving van een dergelijke code en het toezicht op gebruiksverantwoordelijken die zich niet aan een dergelijke code houden, zie punt 8.1.

6.2.3. *Uitzondering op de transparantieverplichting voor tekst die een proces van menselijke toetsing of redactionele controle heeft ondergaan en onder redactionele verantwoordelijkheid valt*

(133) In artikel 50, lid 4, tweede alinea, van de AI-verordening is voorzien in een uitzondering op de in die bepaling vastgestelde transparantieverplichting indien aan twee cumulatieve voorwaarden is voldaan: i) de door AI gegenereerde of gemanipuleerde tekst moet een proces van menselijke toetsing of redactionele controle hebben ondergaan en ii) een natuurlijke of rechtspersoon moet redactionele verantwoordelijkheid dragen voor de publicatie. Indien nodig kunnen gebruiksverantwoordelijken gebruikmaken van relevante toepasselijke professionele of deontologische normen om de naleving van die vereisten aan te tonen, zoals hieronder uiteengezet.

i. Tekst die een proces van menselijke toetsing of redactionele controle heeft ondergaan

(134) De eerste voorwaarde voor de toepassing van de uitzondering op artikel 50, lid 4, tweede alinea, van de AI-verordening is dat de door AI gegenereerde of gemanipuleerde tekst een proces van menselijke toetsing of redactionele controle heeft ondergaan. Menselijke toetsing heeft betrekking op het doelbewust onderzoeken van de inhoud van de content door een of meer natuurlijke personen die beschikken over relevante kennis en professioneel oordeel met betrekking tot het onderzochte onderwerp (bv. academische collegiale toetsing of professionele validatieketens). Het uitvoeren van een feitenonderzoek om de juistheid van de content te verifiëren, is een minimumvereiste die deel moet uitmaken van die beoordeling. Redactionele controle heeft betrekking op de praktische controle die een verantwoordelijke redactionele instantie (bv. een hoofdredacteur) uitoefent over de content, waarbij die instantie de inhoud van de tekst op inhoudelijke gronden mag goedkeuren, wijzigen of afwijzen (waaronder het uitvoeren van feitenonderzoek en het waarborgen van de betrouwbaarheid van bronnen). Dit doet geen afbreuk aan de bestaande toetsings- en redactionele procedures en professionele normen die van toepassing zijn op aanbieders van mediadiensten.

(135) Oppervlakkige, louter formele of procedurele controles (bv. spellingscontrole of grammaticale correcties), het loutere bestaan van een redactioneel beleid, geautomatiseerde beoordelingsprocessen of terloopse redactionele goedkeuring zonder inhoudelijke betrokkenheid van de menselijke beoordelaar of de redactionele entiteit, voldoen niet aan de voorwaarden voor menselijke toetsing of redactionele controle voor de toepassing van deze uitzondering.

(136) Wanneer AI-systemen worden gebruikt om content aan te passen, aan te vullen of te herformuleren nadat die redactioneel is goedgekeurd, moet de resulterende content worden beschouwd als door AI gegenereerd of gemanipuleerd in de zin van artikel 50, lid 4, van de AI-verordening. Elke substantiële AI-interventie die plaatsvindt nadat het proces van menselijke toetsing of redactionele controle heeft plaatsgevonden, leidt dan ook tot de nietigheid van de uitzondering.

(137) In het licht van artikel 50, lid 7, van de AI-verordening kunnen gebruiksverantwoordelijken zich houden aan een praktijkcode die op grond van die bepaling als toereikend wordt beoordeeld, om de naleving aan te tonen van de in die bepaling vastgestelde voorwaarde van menselijke toetsing of redactionele controle (in het bijzonder door gebruiksverantwoordelijken die geen aanbieders van mediadiensten zijn en die onderworpen zijn aan bestaande redactionele professionele of deontologische normen). Zie punt 8.1 voor meer informatie over de effecten van de naleving van een dergelijke code en het toezicht op gebruiksverantwoordelijken die zich niet aan een dergelijke code houden.

ii. *Redactionele verantwoordelijkheid*

(138) De tweede voorwaarde voor de toepassing van de uitzondering op artikel 50, lid 4, tweede alinea, van de AI-verordening vereist dat een natuurlijke of rechtspersoon redactionele verantwoordelijkheid draagt voor de publicatie van de content. Dat houdt in dat die persoon de uiteindelijke wettelijke verantwoordelijkheid moet dragen voor de publicatie van de content, inclusief de menselijke toetsing of redactionele controle (bv. een persoon, redactieraad of de uitgeverij). Om publieke verantwoordingsplicht en het vertrouwen van het publiek te waarborgen, en in overeenstemming met de bestaande professionele medianormen, moeten de identiteit en contactgegevens van de natuurlijke of rechtspersoon of de functie met redactionele verantwoordelijkheid openbaar worden gemaakt op een gemakkelijk vindbare locatie (indien nog niet anderszins beschikbaar). Dat kan online gebeuren, bv. via de algemene voorwaarden van een website of andere juridische informatie die aan gebruikers wordt verstrekt. Offline kan dergelijke informatie worden opgenomen in het colofon van een publicatie.

Voorbeelden van tekst die menselijke toetsing of redactionele controle moet ondergaan door een natuurlijke of rechtspersoon die redactionele verantwoordelijkheid draagt

— Een door AI gemanipuleerd artikel in een krant of een door AI gegenereerde samenvatting van een artikel dat redactionele controle heeft ondergaan door de hoofdredacteur, waarbij de redactionele verantwoordelijkheid berust bij de rechtspersoon die de krant uitgeeft.

— Een door AI gemanipuleerde academische blog die een interne collegiale toetsing heeft ondergaan en waarbij de redactionele verantwoordelijkheid berust bij het onderzoekscentrum dat de blog beheert.

— Door AI gegenereerde waarschuwingen voor de openbare veiligheid die door een overheidsfunctionaris worden goedgekeurd alvorens ze onder de verantwoordelijkheid van de bevoegde overheidsinstantie voor civiele bescherming onder de mensen worden verspreid.

— Door AI gegenereerde duurzaamheidsverslagen die op de website van een beursgenoteerde onderneming worden gepubliceerd nadat ze door professionals in relevante functies (bv. naleving) een menselijke toetsing hebben ondergaan.

— Door AI ondersteunde vertaling van een door mensen geschreven artikel, waarbij de vertaling een menselijke toetsing heeft ondergaan.

Voorbeelden van tekst die niet voldoen aan de vereiste menselijke toetsing of redactionele controle om in aanmerking te komen voor de uitzondering

— Een website waarop door AI gegenereerde artikelen over het Uniebeleid worden geplaatst zonder doelbewuste menselijke toetsing of redactionele controle.

— Door AI gegenereerde artikelen die door een ander AI-systeem worden beoordeeld en bewerkt en waarbij een menselijke redacteur vóór publicatie slechts een oppervlakkige grammaticale controle uitvoert.

— Een door AI gegenereerd boek in zelfpublicatie over klimaatverandering, dat beschikbaar wordt gesteld op een platform voor e-commerce en dat geen toetsing heeft ondergaan door een bevoegde natuurlijke of rechtspersoon (noch door het platform).

6.2.4. *Uitzondering in verband met rechtshandhaving*

(139) Indien het gebruik van AI-systemen voor het genereren van tekst die wordt gepubliceerd om het publiek te informeren over aangelegenheden van algemeen belang zonder menselijke toetsing (bv. een door AI gegenereerde politiewaarschuwing over een terroristische dreiging die op sociale media is gepubliceerd) bij wet is toegestaan om strafbare feiten op te sporen, te voorkomen, te onderzoeken of te vervolgen, zijn de gebruiksverantwoordelijken vrijgesteld van de transparantieverplichting uit hoofde van artikel 50, lid 4, van de AI-verordening. De verduidelijkingen in de punten 46 tot en met 48 voor een soortgelijke uitzondering in artikel 50, lid 1, van de AI-verordening zijn ook relevant voor deze bepaling.

6.2.5. *Wisselwerking met andere rechtshandelingen van de Unie*

(140) Voor de mediawetgeving van de Unie is redactionele verantwoordelijkheid een cruciaal concept dat is gedefinieerd in artikel 2, punt 8, van Verordening (EU) 2024/1083 (de Europese verordening mediavrijheid) ([39]). In die context wordt verstaan onder "redactionele verantwoordelijkheid": "de uitoefening van effectieve controle over zowel de keuze van programma's of perspublicaties als over de organisatie ervan, met het oog op de levering van een mediadienst, ongeacht het bestaan van aansprakelijkheid krachtens het nationale recht voor de geleverde dienst". In bepaalde situaties kan deze definitie overlappen met het begrip "redactionele verantwoordelijkheid" dat wordt gebruikt in artikel 50, lid 4, tweede alinea, van de AI-verordening, vooral als de gebruiksverantwoordelijken van het AI-systeem ook kunnen worden aangemerkt als aanbieders van mediadiensten in het kader van de Europese verordening mediavrijheid. Redactionele verantwoordelijkheid in de zin van artikel 50, lid 4, van de AI-verordening moet daarom worden geïnterpreteerd volgens het bestaande media-acquis, ook al blijft het een apart begrip dat ook in bredere contexten en op andere gebruiksverantwoordelijken van toepassing kan zijn, mits ze de redactionele verantwoordelijkheid voor de publicatie van de tekst op zich nemen. Aanbieders van mediadiensten kunnen gebruikmaken van hun bestaande redactionele processen en normen om in aanmerking te komen voor de uitzondering van artikel 50, lid 4, tweede alinea, van de AI-verordening.

7. **HORIZONTALE EISEN DIE VAN TOEPASSING ZIJN OP DE INFORMATIE DIE WORDT VERSTREKT OP GROND VAN ARTIKEL 50, LID 5, VAN DE AI-VERORDENING**

(141) Op grond van artikel 50, lid 5, van de AI-verordening moet de informatie die overeenkomstig artikel 50, leden 1 tot en met 4, van de AI-verordening moet worden verstrekt, uiterlijk op het moment van de eerste interactie of blootstelling op een duidelijke en te onderscheiden wijze aan de betrokken natuurlijke personen worden verstrekt. De informatie moet ook aan de toepasselijke toegankelijkheidseisen voldoen. Tot de betrokken natuurlijke personen in deze context behoren de personen die rechtstreeks met het AI-systeem interageren (artikel 50, lid 1, van de AI-verordening), actieve of passieve gebruikers en andere personen die worden blootgesteld aan door AI gegenereerde of gemanipuleerde synthetische content (al dan niet met directe interactie met het AI-systeem (artikel 50, leden 2 en 4, van de AI-verordening), en personen die te maken krijgen met het gebruik van systemen voor het herkennen van emoties of biometrische categorisering (artikel 50, lid 3, van de AI-verordening).

7.1. **Het verstrekken van informatie op duidelijke en te onderscheiden wijze**

(142) Informatie wordt geacht op duidelijke wijze te zijn verstrekt als ze opvallend, gemakkelijk te begrijpen en toegankelijk is voor de betrokken natuurlijke persoon (inclusief personen met een handicap). Informatie wordt geacht op een te onderscheiden wijze te worden verstrekt als ze gemakkelijk te onderscheiden is van andere informatie en van de omgeving waarin de content wordt gepresenteerd. Ze moet tevens gemakkelijk te begrijpen zijn voor het publiek, ook specifieke groepen zoals kinderen of personen met een handicap als die deel uitmaken van het publiek dat naar verwachting aan de content zal worden blootgesteld of met het systeem zal interageren ([40]). Om informatie op een duidelijke en te onderscheiden wijze te verstrekken, kan de informatie onder meer worden gepresenteerd als onderdeel van de interactie (op grond van artikel 50, lid 1, van de AI-verordening) of de desbetreffende content (op grond van artikel 50, leden 2 en 4, van de AI-verordening). Informatie wordt niet geacht op duidelijke en te onderscheiden wijze te worden verstrekt als ze door natuurlijke personen onder normale omstandigheden van blootstelling of interactie gemakkelijk over het hoofd kan worden gezien of kan worden gemist (bv. alleen opgenomen als onderdeel van een handleiding of verborgen onder lagen menuopties op een online-interface, een deel van de gebruiksvoorwaarden die gebruikers vaak niet lezen).

([39]) Verordening (EU) 2024/1083 van het Europees Parlement en de Raad van 11 april 2024 tot vaststelling van een gemeenschappelijk kader voor mediadiensten op de interne markt en tot wijziging van Richtlijn 2010/13/EU (Europese verordening mediavrijheid) (PB L, 2024/1083, 17.4.2024, ELI: http://data.europa.eu/eli/reg/2024/1083/oj).

([40]) Wanneer kinderen deel uitmaken van de doelgroep van de content, moeten de kennisgevingen: i) kindvriendelijk, leeftijdsgeschikt, gemakkelijk te begrijpen en gemakkelijk toegankelijk zijn voor alle kinderen, ook kinderen met een handicap en/of extra toegankelijkheidsbehoeften; ii) duidelijk worden gepresenteerd op een manier die gemakkelijk te begrijpen is en zo eenvoudig en beknopt mogelijk is; iii) worden gepresenteerd op een manier die gemakkelijk te beoordelen is en die onmiddellijke en intuïtieve toegang biedt, op de punten waarop ze relevant zijn; iv) worden gepresenteerd in de officiële taal of talen van de lidstaat waar de dienst wordt verricht; v) aantrekkelijk zijn voor kinderen. Dit kan ook het gebruik van beelden, video's en/of personages of andere technieken omvatten; vi) geleidelijk en over een langere periode aan kinderen worden getoond, zodat de gebruiker ze het best kan onthouden.

7.2. **Eerste interactie of blootstelling**

(143) Op grond van artikel 50, lid 5, van de AI-verordening moet de informatie uiterlijk op het moment van de eerste interactie of blootstelling aan de betrokken natuurlijke personen worden verstrekt. De eerste interactie of blootstelling heeft niet alleen betrekking op de eerste natuurlijke persoon die interageert met of wordt blootgesteld aan (de output van) een AI-systeem, maar ook op elke daaropvolgende eerste interactie met of blootstelling aan (de output van) het AI-systeem door een andere natuurlijke persoon. Wat de eerste interactie met een AI-systeem betreft, moet de informatie ten minste eenmaal aan het begin van een interactieve sessie worden verstrekt, rekening houdend met de in de punten 36 en 40 vermelde specifieke contexten. Wat door AI gegenereerde of gemanipuleerde content betreft die binnen het toepassingsgebied van artikel 50, leden 2 en 4, van de AI-verordening valt, is de informatieverplichting van toepassing op elke output van een AI-systeem ten aanzien van elke natuurlijke persoon die aan de content wordt blootgesteld. Dat moet worden opgevat in de zin van elk moment waarop het redelijkerwijs aannemelijk is dat een natuurlijke persoon wordt blootgesteld aan de output van het AI-systeem en de openbaarmaking ontvangt, rekening houdend met de specifieke kenmerken van de openbaarmaking in het geval van kennelijk creatieve, artistieke en andere werken (zie punt 6.1.3.). Gebruiksverantwoordelijken mogen natuurlijke personen ook eerder informeren dan de eerste daadwerkelijke interactie of blootstelling (bv. aan het begin van content met deepfakes). Als het echter redelijkerwijs voorzienbaar is dat personen content niet vanaf het begin waarnemen, is openbaarmaking aan het begin van de content alleen niet voldoende om die personen toereikend te informeren en moet deze waar mogelijk worden aangevuld met openbaarmaking op latere momenten. Dat kan naargelang de situatie worden uitgevoerd, rekening houdend met de doelstelling van de transparantieverplichtingen om te zorgen voor een doeltreffende verstrekking van informatie en een daadwerkelijke bewustwording bij de gebruikers in steeds immersievere en realistischere AI-omgevingen, waarin gebruikers steeds meer moeite hebben om de oorsprong en waarheidsgetrouwheid van de content en de interactie te onderscheiden.

Voorbeelden van informatieverstrekking bij de eerste interactie of blootstelling op grond van artikel 50, lid 5, van de AI-verordening

— Eerste interactie: bij het starten van een gesprek met een chatbot of het starten van fysieke interactie met een AI-systeem.

— Eerste blootstelling: aan het begin van een video met een deepfake, aan het begin van een door AI gegenereerde of gemanipuleerde publicatie van tekst over aangelegenheden van algemeen belang of als er door AI gemanipuleerde deepfakes verschijnen bij het scrollen op sociale media.

7.3. **Naleving van de toepasselijke toegankelijkheidsvoorschriften**

(144) Aanbieders en gebruiksverantwoordelijken van AI-systemen die binnen het toepassingsgebied van de verschillende transparantieverplichtingen genoemd in artikel 50, leden 1 tot en met 4, van de AI-verordening vallen, moeten op de hoogte zijn van de toepasselijke toegankelijkheidsvoorschriften (bv. op grond van Richtlijn (EU) 2016/2102 ([41]) en Richtlijn (EU) 2019/882 ([42])) en beoordelen of hun product, dienst of content aan die regels moet voldoen. In dat geval moeten ze de toegankelijkheid van de overeenkomstig artikel 50 van de AI-verordening overgebrachte informatie waarborgen in overeenstemming met de toepasselijke toegankelijkheidsvoorschriften. Artikel 50 van de AI-verordening bevat geen afzonderlijke of extra toegankelijkheidsvoorschriften.

8. **HANDHAVING VAN ARTIKEL 50 VAN DE AI-VERORDENING**

(145) Dit punt heeft als doel het volgende te verduidelijken: i) de effecten van de naleving van een praktijkcode met betrekking tot de naleving van de transparantieverplichtingen voor door AI gegenereerde content; ii) de rol van de bevoegde toezichthoudende autoriteiten bij de handhaving van artikel 50 van de AI-verordening; iii) de toepasselijke sancties voor inbreuken op die verplichtingen, en iv) het moment waarop de verplichtingen van artikel 50 van de AI-verordening van toepassing worden.

([41]) Richtlijn (EU) 2016/2102 van het Europees Parlement en de Raad van 26 oktober 2016 inzake de toegankelijkheid van de websites en mobiele applicaties van overheidsinstanties (PB L 327 van 2.12.2016, blz. 1).

([42]) Richtlijn (EU) 2019/882 van het Europees Parlement en de Raad van 17 april 2019 betreffende de toegankelijkheidsvoorschriften voor producten en diensten (PB L 151 van 7.6.2019, blz. 70).

8.1. **Effecten van de naleving van een als toereikend beoordeelde praktijkcode**

(146) Aanbieders en gebruiksverantwoordelijken van AI-systemen die binnen het toepassingsgebied van artikel 50, leden 2 en 4, van de AI-verordening vallen, kunnen aantonen dat ze voldoen aan hun transparantieverplichtingen met betrekking tot door AI gegenereerde content op grond van die bepalingen door een praktijkcode na te leven die overeenkomstig artikel 50, lid 7, van de AI-verordening als toereikend is beoordeeld ([43]). Een dergelijke code komt niet in de plaats van de AI-verordening of deze richtsnoeren. Ze vormt er een aanvulling op door het enige in de hele Unie erkende praktische kader te bieden waarmee aanbieders en gebruiksverantwoordelijken van generatieve AI-systemen kunnen aantonen dat ze aan die verplichtingen voldoen, ongeacht hun plaats van vestiging, hun activiteiten en de bevoegde markttoezichtautoriteiten.

(147) Hoewel aanbieders en gebruiksverantwoordelijken de naleving van die verplichtingen ook met toereikende alternatieve middelen kunnen aantonen, is de naleving van een praktijkcode die toereikend wordt geacht, een eenvoudige, voorspelbare en rechtszekere manier om de naleving aan te tonen. Voor ondertekenaars van een praktijkcode die op grond van artikel 50, lid 7, van de AI-verordening als toereikend is beoordeeld ([44]), zullen de Commissie en de bevoegde markttoezichtautoriteiten hun toezichtactiviteiten toespitsen op de beoordeling of die ondertekenaars de praktijkcode hebben nageleefd en de daarin vervatte maatregelen hebben uitgevoerd. Naarmate aanbieders en gebruiksverantwoordelijke die zich aan de regels houden transparant zijn over de maatregelen die ze nemen om te voldoen aan artikel 50, leden 2, 4 en 5, van de AI-verordening, zullen ze een groter vertrouwen genieten bij de Commissie, de andere bevoegde markttoezichtautoriteiten en andere belanghebbenden (waaronder het grote publiek). Elke opt-out van bepaalde onderdelen door ondertekenaars van een praktijkcode die als toereikend wordt beschouwd, leidt ertoe dat die aanbieders en gebruiksverantwoordelijken het voordeel verliezen dat ze in dat opzicht het aantonen van naleving kunnen vergemakkelijken.

(148) Van aanbieders en gebruiksverantwoordelijken die geen praktijkcode hebben ondertekend die krachtens artikel 50, lid 7, van de AI-verordening toereikend wordt geacht, wordt verwacht dat ze met andere toereikende middelen aantonen hoe ze aan hun verplichtingen uit hoofde van artikel 50, leden 2, 4 en 5, van de AI-verordening hebben voldaan. Bovendien wordt van dergelijke aanbieders en gebruiksverantwoordelijken verwacht dat ze toelichten hoe de maatregelen die ze nemen de naleving van hun verplichtingen uit hoofde van de AI-verordening waarborgen. Ze moeten bijvoorbeeld een kloofanalyse uitvoeren waarin hun maatregelen worden vergeleken met de maatregelen die zijn vastgelegd in een praktijkcode die als toereikend is beoordeeld. Aangezien de bevoegde autoriteiten (waaronder het AI-bureau) minder inzicht zullen hebben in de manier waarop aanbieders en gebruiksverantwoordelijken die geen praktijkcode hebben ondertekend, de naleving van hun verplichtingen van artikel 50, leden 2, 4 en 5, van de AI-verordening waarborgen, zullen ze bij het toezicht op de naleving waarschijnlijk meer gedetailleerde informatie nodig hebben. Aanbieders die zich niet aan een praktijkcode houden, kunnen bijvoorbeeld ook worden onderworpen aan een groter aantal verzoeken om informatie en verzoeken om toegang om de doeltreffendheid, interoperabiliteit, robuustheid en betrouwbaarheid te beoordelen van de technische oplossingen om te voldoen aan artikel 50, lid 2, van de AI-verordening. Gebruiksverantwoordelijken kunnen eveneens worden onderworpen aan dergelijke verzoeken met betrekking tot hun labelingspraktijken op grond van artikel 50, lid 4, van de AI-verordening.

(149) Afhankelijk van de specifieke omstandigheden kunnen de bevoegde autoriteiten bij de vaststelling van het bedrag van boetes als verzachtende factor rekening houden met toezeggingen die zijn nagekomen in overeenstemming met een praktijkcode die als toereikend wordt beschouwd ([45]).

(150) Indien een praktijkcode op grond van artikel 50, lid 7, van de AI-verordening niet als toereikend wordt beschouwd, kan de Commissie een uitvoeringshandeling vaststellen waarin gemeenschappelijke regels voor de uitvoering van de verplichtingen van artikel 50, leden 2, 4 en 5, van de AI-verordening worden gespecificeerd, die van toepassing zouden zijn op alle relevante aanbieders en gebruiksverantwoordelijken van AI-systemen die onder het toepassingsgebied van die bepalingen vallen ([46]).

([43]) Zie de informatie over de praktijkcode over door AI gegenereerde content, gepubliceerd op 10 juni 2026.
([44]) Zie het advies van de Commissie over de beoordeling van de praktijkcode over de transparantie van door AI gegenereerde content.
([45]) Artikel 99, lid 7, punt e), van de AI-verordening.
([46]) Artikel 50, lid 7, van de AI-verordening.

8.2. **Markttoezichtautoriteiten**

(151) De door de lidstaten aangewezen markttoezichtautoriteiten, het AI-bureau ([47]) en de Europese Toezichthouder voor gegevensbescherming ([48]) zijn verantwoordelijk voor het toezicht op en de handhaving van de regels voor AI-systemen die onder hun bevoegdheid vallen, inclusief de transparantieverplichtingen van artikel 50 van de AI-verordening. Een dergelijke handhaving vindt plaats binnen het systeem voor markttoezicht en conformiteit van producten dat is vastgesteld bij Verordening (EU) 2019/1020 en de AI-verordening. De handhavingsbevoegdheden van markttoezichtautoriteiten met betrekking tot AI-systemen zijn vastgelegd in die verordeningen. Die autoriteiten kunnen op eigen initiatief of naar aanleiding van een klacht maatregelen nemen voor de handhaving van de in artikel 50 van de AI-verordening genoemde verplichtingen. Elke getroffen persoon en elke andere natuurlijke of rechtspersoon die redenen heeft om van mening te zijn dat de bepalingen geschonden zijn, heeft het recht om een klacht in te dienen ([49]). De verantwoordelijkheden van de markttoezichtautoriteiten op grond van de AI-verordening doen geen afbreuk aan de bevoegdheden en taken van andere toezichthoudende autoriteiten of autoriteiten voor de bescherming van de grondrechten op grond van hun eigen rechtskader in situaties waarin zowel artikel 50 van de AI-verordening als dat kader van toepassing zijn (bv. gegevensbescherming, consumentenbescherming).

8.3. **Sancties**

(152) Aanbieders en gebruiksverantwoordelijken die niet voldoen aan de toepasselijke transparantieverplichtingen van artikel 50 van de AI-verordening, kunnen een geldboete krijgen tot 15 miljoen EUR of, indien de overtreder een onderneming is, tot 3 % van haar totale wereldwijde jaaromzet voor het voorgaande boekjaar, indien dat hoger is. Aan instellingen, organen en instanties van de EU die de transparantieverplichtingen schenden, kunnen administratieve boetes worden opgelegd van maximaal 750 000 EUR. In het geval van kmo's, waaronder start-ups, bedraagt elke geldboete niet meer dan de bovengenoemde percentages of het bovengenoemde bedrag, naargelang welk het laagst is. De boetes moeten doeltreffend, evenredig en ontmoedigend van aard zijn, waarbij rekening wordt gehouden met de belangen van kmo's en kleine midcapondernemingen (waaronder start-ups) en met hun economische levensvatbaarheid. Bij het besluiten om al dan niet een administratieve geldboete op te leggen en bij het bepalen van het bedrag van de administratieve geldboete in elk individueel geval, moeten de bevoegde autoriteiten alle relevante omstandigheden van de specifieke situatie beoordelen en, waar nodig, een reeks factoren die van invloed zijn op de ernst van de inbreuk (bv. de aard, ernst en duur van de inbreuk en van de gevolgen daarvan, het nalatige of opzettelijke karakter, andere verzwarende of verzachtende omstandigheden, de mate van samenwerking met markttoezichtautoriteiten enz.).

8.4. **Inwerkingtreding**

(153) Overeenkomstig artikel 113 van de AI-verordening is artikel 50 van de AI-verordening van toepassing met ingang van 2 augustus 2026. Dat betekent dat alle onder het toepassingsgebied vallende AI-systemen die in de Unie in de handel worden gebracht of in gebruik worden genomen, op die datum aan die bepaling moeten voldoen, ongeacht de datum waarop ze in de handel zijn gebracht of in gebruik zijn genomen ([50]). De verordening tot wijziging van de AI-verordening (de AI-omnibus), die onlangs door de Uniewetgever is vastgesteld, voorziet alleen in een gerichte grandfatheringregel met betrekking tot de markerings- en detectieverplichtingen van artikel 50, lid 2, van de AI-verordening voor generatieve AI-systemen die vóór 2 augustus 2026 in de handel zijn gebracht of in gebruik zijn genomen. Dat geeft de aanbieders van die bestaande systemen een overgangsperiode om hun systemen uiterlijk op 2 december 2026 in overeenstemming te brengen met de voorschriften. Systemen die gedeeltelijk interactief en gedeeltelijk generatief zijn, komen alleen in aanmerking voor die overgangsperiode met betrekking tot de markeringsverplichting van artikel 50, lid 2, van de AI-verordening; de naleving van de verplichting tot openbaarmaking voor AI-systemen die direct interageren met natuurlijke personen moet vanaf 2 augustus 2026 worden gewaarborgd.

([47]) Als de markttoezichtautoriteit voor AI-systemen die gebaseerd zijn op GPAI-modellen die door dezelfde aanbieder worden aangeboden (artikel 75, lid 1, van de AI-verordening). Dat heeft betrekking op de verplichting die krachtens artikel 50, leden 1 en 2, van de AI-verordening op aanbieders rust. Wat betreft artikel 50, leden 3 en 4, van de AI-verordening, is het AI-bureau alleen bevoegd indien de betreffende aanbieder van het GPAI-model en het systeem tevens gebruiksverantwoordelijke van het systeem is.

([48]) Als de markttoezichtautoriteit voor de instellingen, agentschappen en organen van de EU.

([49]) Artikel 85 van de AI-verordening.

([50]) De bijzondere grandfatheringregel van artikel 111, lid 2, van de AI-verordening is alleen van toepassing op AI-systemen met een hoog risico die vóór de datum van toepassing van de regels voor AI-systemen met een hoog risico in de handel zijn gebracht met betrekking tot de naleving van die systemen van de voorschriften en verplichtingen voor AI-systemen met een hoog risico. Indien het AI-systeem met een hoog risico tevens onderworpen is aan een of meer transparantieverplichtingen van artikel 50 van de AI-verordening, is die bijzondere grandfatheringregel niet van toepassing, aangezien de toepassing van de verschillende verplichtingen voor AI-systemen met een hoog risico en betreffende transparantie van artikel 50 cumulatief zijn en de grandfathering uitsluitend gerechtvaardigd en beperkt is met betrekking tot de naleving van de bepalingen inzake systemen met een hoog risico.

(154) Door AI gegenereerde of gemanipuleerde output die binnen het toepassingsgebied van artikel 50, lid 2, van de AI-verordening valt en deepfakes die binnen het toepassingsgebied van artikel 50, lid 4, eerste alinea, van de AI-verordening vallen en die vóór 2 augustus 2026 zijn gegenereerd of gemanipuleerd, hoeven niet met terugwerkende kracht te worden gemarkeerd of gelabeld. Hetzelfde geldt voor teksten over aangelegenheden van algemeen belang die binnen het toepassingsgebied van artikel 50, lid 4, tweede alinea, van de AI-verordening vallen en die vóór 2 augustus 2026 door AI zijn gegenereerd of gemanipuleerd en gepubliceerd. Omgekeerd moeten teksten die vóór 2 augustus 2026 zijn gegenereerd of gemanipuleerd en die op of na die datum worden gepubliceerd, wel worden gemarkeerd. Wegens het door de AI-verordening nagestreefde doel van transparantie en meer vertrouwen in en integriteit van het informatie-ecosysteem, worden gebruiksverantwoordelijken van AI-systemen en andere actoren die in het bezit zijn van bestaande niet-gelabelde deepfakes of die verspreiden, niettemin aangemoedigd dat wel te doen, zonder dat van het wordt verwacht dat ze daartoe onevenredige inspanningen leveren (zoals het controleren van bestaande contentdatabanken of het aanpassen van reeds gedrukte productverpakkingen).

9. **HERZIENING EN ACTUALISERING VAN DE RICHTSNOEREN**

(155) Deze richtsnoeren vormen een eerste interpretatie met praktische voorbeelden van de transparantieverplichtingen van artikel 50 van de AI-verordening. De Commissie zal de richtsnoeren evalueren zodra dit nodig wordt geacht op grond van de praktische ervaringen die zijn opgedaan bij de uitvoering van de transparantieverplichtingen en het tempo van de technologische, maatschappelijke en wettelijke ontwikkelingen op dit gebied. Daaronder vallen ook ervaringen met handhavingsmaatregelen op het gebied van markttoezicht en interpretaties van artikel 50 van de AI-verordening door het HvJ-EU. Tijdens een dergelijke evaluatie kan de Commissie besluiten deze richtsnoeren in te trekken of te wijzigen. De Commissie moedigt aanbieders en gebruiksverantwoordelijken van AI-systemen, nationale markttoezichtautoriteiten (via de AI-board), het AI-adviesforum, onderzoekers en maatschappelijke organisaties aan om een bijdrage te leveren aan dit proces door deel te nemen aan toekomstige openbare raadplegingen.