De *AI-regulatory sandbox* in de Verordening artificiële intelligentie

Mathijs Tollerton en Damien van Vliet[1]

1. Inleiding
In de verordening artificiële intelligentie[2] (hierna: de AI-verordening of AIV) worden regels gesteld over het in de handel brengen en gebruiken van artificiële intelligentie (AI) en worden eisen gesteld aan AI-systemen met een hoog risico. De AI-verordening biedt de mogelijkheid om het ontwikkelen en testen van AI-systemen (deels) te laten plaatsvinden binnen een AI-testomgeving voor regelgeving (hierna: de *AI-regulatory sandbox*). In paragraaf 2 zal eerst kort worden ingegaan op de inhoud van de AI-verordening in het algemeen. In de paragrafen 3 tot en met 5 wordt ingegaan op de bepalingen over de *AI-regulatory sandbox* en hoe hier uitvoering aan wordt gegeven in Nederland (paragraaf 6). Ten slotte zal een korte analyse worden gegeven van deze *sandbox*.

2. De AI-verordening
AI wordt gekenmerkt door een snelle ontwikkeling en heeft een groeiende invloed op de samenleving en de economie. AI kan grote voordelen opleveren op het gebied van innovatie, efficiëntie en concurrentiekracht. Tegelijkertijd kan AI aanzienlijke risico's met zich meebrengen. Deze risico's betreffen onder meer het gebrek aan transparantie en controleerbaarheid van AI-systemen, het risico op discriminatie door bias in algoritmen of trainingsdata, de schending van het recht op privacy en gegevensbescherming en de mogelijke inzet van AI voor criminele doeleinden. Tegen deze achtergrond is de AI-verordening tot stand gekomen.

De AI-verordening heeft drie doelen (artikel 1, eerste lid, AIV):
1) de toepassing van mensgerichte en betrouwbare AI bevorderen;
2) een hoog niveau van bescherming van de gezondheid, veiligheid en grondrechten, met inbegrip van de democratie, rechtsstaat en bescherming van het milieu;
3) innovatie ondersteunen.

De AI-verordening harmoniseert de regels voor de ontwikkeling, het in de handel brengen, het in gebruik stellen en het gebruik van artificiële intelligentiesystemen (AI-systemen) in de Europese Unie. De AI-verordening kent een gelaagd systeem. Ten eerste zijn bepaalde onaanvaardbare praktijken op het gebied van AI verboden. Ten tweede worden er eisen gesteld voor AI-systemen met een hoog risico en worden er verplichtingen opgelegd aan de operatoren daarvan. Ten derde zijn er transparantieverplichtingen van toepassing op bepaalde AI-systemen. Ten slotte bevat de verordening bepalingen ten aanzien van AI-modellen voor algemene doeleinden.

De bepalingen over AI-systemen met een hoog risico volgen de vaste elementen uit het nieuwe wetgevingskader voor productwetgeving van de Unie:
1. Essentiële eisen voor AI-systemen waaraan moet worden voldaan ter bescherming van de veiligheid, gezondheid en grondrechten van personen, uitgedrukt in te behalen resultaten. De algemeen geformuleerde eisen kunnen worden uitgewerkt in specifiek geformuleerde normen (normalisatie).
2. Conformiteitsbeoordelingsprocedures die moeten worden doorlopen voordat een AI-systeem in de handel wordt gebracht.
3. Verplichtingen voor operatoren die onderdeel zijn van de toeleveringsketen van een AI-systeem (aanbieders, importeurs, distributeurs en gebruiksverantwoordelijken).
4. Markttoezicht door de lidstaten aangewezen markttoezichtautoriteiten.

2. De regulatory sandbox in de AI-verordening
Lidstaten zijn verplicht om ervoor te zorgen dat hun nationale bevoegde autoriteiten ten minste één *AI-regulatory sandbox* tot stand brengen (artikel 57, eerste lid, AIV). De *AI-regulatory sandbox* moet op nationaal niveau worden opgezet, maar dit kan ook in samenwerking met

[1] Beiden werkzaam als wetgevingsjurist bij de Directie wetgeving en juridische zaken van het Ministerie van Economische Zaken en Klimaat en het Ministerie van Landbouw, Visserij, Voedselzekerheid en Natuur.
[2] Verordening 2024/1689 van het Europees Parlement en de Raad van 13 juni 2024 tot vaststelling van geharmoniseerde regels betreffende artificiële intelligentie en tot wijziging van de van de Verordeningen (EG) nr. 300/2008, (EU) nr. 167/2013, (EU) nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 en (EU) 2019/2144, en de Richtlijnen 2014/90/EU, (EU) 2016/797 en (EU) 2020/1828 (verordening artificiële intelligentie).

bevoegde autoriteiten van andere lidstaten gebeuren of door aan te sluiten bij bestaande (nationale) *regulatory sandboxes*.

De *AI-regulatory sandbox* heeft de volgende doelstellingen (artikel 57, negende lid, AIV):

1. het verbeteren van de rechtszekerheid;
2. het ondersteunen van de uitwisseling van beste praktijken (*best practices*);
3. het bevorderen van innovatie en concurrentievermogen en het faciliteren van de ontwikkeling van een AI-ecosysteem;
4. het bijdragen aan kennis over de toepassing van regelgeving in de praktijk (*regulatory learning*);
5. het vergemakkelijken van de toegang tot de Europese markt voor AI-systemen.

De bepalingen in de AI-verordening over de *AI-regulatory sandbox* zijn behoorlijk algemeen en abstract geformuleerd. De *AI-regulatory sandbox* moet een gecontroleerde omgeving zijn ter bevordering van innovatie en ter vergemakkelijking van het ontwikkelen, trainen, testen en valideren van innovatieve AI-systemen (artikel 57, vijfde lid, AIV). De bevoegde autoriteiten moeten binnen deze gecontroleerde omgeving begeleiding, toezicht en ondersteuning verstrekken om risico's in kaart te brengen, met name betreffende grondrechten, gezondheid en veiligheid. Ook moeten zij begeleiding bieden bij het testen van systemen en het nemen van beperkende maatregelen door de potentiële aanbieders, teneinde de doeltreffendheid daarvan te bepalen met betrekking tot de verplichtingen en eisen van de AI-verordening en, indien van toepassing, ander Unie- en nationaal recht (artikel 57, zesde lid, AIV).

Significante risico's voor gezondheid, veiligheid of fundamentele rechten die in de *regulatory sandbox* ontstaan, moeten adequaat gemitigeerd worden (artikel 57, elfde lid, AIV). Als dat niet mogelijk is, hebben de bevoegde autoriteiten de bevoegdheid om het testen van het AI-systeem of deelname aan de regulatory sandbox tijdelijk of permanent te stoppen.

Om voorkomen dat de AI-regulatory sandboxes in verschillende lidstaten van elkaar verschillen zal de Commissie uitvoeringshandelingen vaststellen waarin gemeenschappelijke beginselen worden opgenomen met betrekking tot de toelatings- en selectiecriteria voor deelname, de procedures en de voor de deelnemers geldende voorwaarden (artikel 58 AIV).

3. Toegang tot de AI-regulatory sandbox
Om toegang te verkrijgen tot de *AI-regulatory sandbox* moeten potentiële aanbieders voldoen aan de door lidstaten gestelde toegangs- en selectiecriteria. De Commissie zal nadere voorwaarden hieraan stellen in de eerdergenoemde uitvoeringshandelingen. Aanvragen die voldoen aan de toelatingscriteria komen in principe in aanmerking voor toelating tot de *AI-regulatory sandbox*. Als de vraag tot toelating tot de regulatory sandbox de beschikbare capaciteit overschrijdt, zal er een selectie moeten worden gemaakt. Gelet hierop is het van belang dat deze selectie geschiedt op een eerlijke en transparante wijze met behulp van duidelijke en vooraf bekende criteria. De schriftelijke beslissing van een nationale bevoegde autoriteit om een potentiële aanbieder al dan niet toe te laten tot de AI-regulatory sandbox kan naar Nederlands recht gekwalificeerd worden als een besluit in de zin van artikel 1:3, eerste lid, van de Algemene wet bestuursrecht.

4. Rechtsgevolgen van deelname aan de AI-regulatory sandbox
De verordening stelt beperkte eisen aan de invulling van de *regulatory sandbox.* Deelname aan de *regulatory sandbox* is alleen mogelijk vóórdat het AI-systeem in de handel is gebracht of in gebruik is genomen. Op het moment van deelname aan de *regulatory sandbox* zijn de verplichtingen uit de AI-verordening daarom nog niet van toepassing, omdat die enkel zien op het op de markt brengen of in gebruik nemen van AI-systemen. Er is daarom niet voorzien in een mogelijkheid om af te wijken van de bepalingen uit de AI-verordening binnen de *AI-regulatory sandbox.*

De bepalingen van de AI-verordening inzake "het testen onder reële omstandigheden" staan wel afwijking van de bepalingen van de AI-verordening toe (artikel 60 AIV). Dit instrument kan door de lidstaten worden toegepast binnen de AI-regulatory sandbox (artikel 57, vijfde lid, AIV). In dit geval kan een AI-systeem onder strikte voorwaarden en tijdelijk in gebruik worden genomen, als dat nodig is om het systeem te testen. Het voor de eerste keer in gebruik nemen van een AI-

systeem valt onder de definitie van het begrip "in de handel brengen" van een AI-systeem en vereist daarmee een afwijking van de regels van de AI-verordening.

Andere regelgeving is tijdens de deelname aan een *regulatory sandbox* wel van toepassing. Hierbij kan bijvoorbeeld worden gedacht aan de Algemene verordening gegevensbescherming. De AI-verordening bevat geen grondslag voor afwijking van die regelgeving. Wel is bepaald dat als de bevoegde autoriteiten voor de toepassing van die andere nationale of Unie-regelgeving actief deelnemen aan een traject in de *AI-regulatory sandbox* en de aanbieder de door die autoriteit verstrekte richtsnoeren voor naleving van dat recht heeft opgevolgd, die autoriteiten met betrekking tot dat recht geen sancties mogen opleggen. De deelnemende aanbieders blijven wel aansprakelijk voor aan derde partijen toegebrachte schade die tijdens het experimenteren in de sandbox kan ontstaan (artikel 57, twaalfde lid, AIV).

In het geval dat de beoordeling van het innovatieve AI-systeem betrekking heeft op de verwerking van persoonsgegevens, moet de gegevensbeschermingsautoriteit betrokken worden bij de sandbox en bij het toezicht op aspecten die onder hun taken en bevoegdheden vallen (artikel 57, tiende lid, AIV). De verordening regelt niet wat deze betrokkenheid inhoudt. Deelname aan een AI-regulatory sandbox kan bij bepaalde AI-systemen, namelijk die ontwikkeld worden in het algemeen belang, een grondslag bieden om persoonsgegevens rechtmatig te verwerken voor een ander doel dan waarvoor ze zijn verkregen (artikel 59 AIV).

5. Uitkomst van een AI-regulatory sandbox
Nadat de activiteiten zijn uitgevoerd, kan de deelnemende aanbieder om een schriftelijk bewijs vragen waarin staat welke activiteiten met succes uitgevoerd zijn. Deelnemende aanbieders kunnen dergelijke documentatie gebruiken om bij het conformiteitsbeoordelingsproces of relevante markttoezichtactiviteiten aan te tonen dat zij aan de verordening voldoen. Daarnaast, of samen met het schriftelijke bewijs verstrekt de bevoegde autoriteit een eindverslag (*exit report*) waarin de in de *regulatory sandbox* uitgevoerde activiteiten en (leer)resultaten worden beschreven (artikel 57, zevende en achtste lid, AIV). In de uitvoeringshandelingen van de Commissie worden nadere algemene voorwaarden gesteld omtrent het schriftelijk bewijs en het *exit report*.

Via het *exit report* wordt nuttige kennis uit de uitgevoerde *sandbox*-trajecten vastgelegd en, met inachtneming van vertrouwelijkheidsbepalingen en toestemming van de deelnemende aanbieder, beschikbaar gesteld aan het AI-bureau en de AI-board (Europese autoriteiten die zijn ingesteld door de AI-verordening). Op die manier kunnen zij de leerresultaten gebruiken bij de uitoefening van hun taken, zoals het vaststellen van richtsnoeren. Met instemming van de deelnemende aanbieder en de nationale bevoegde autoriteit, kan het *exit report* ook openbaar worden gemaakt, zodat de leerresultaten direct aan het bredere AI-ecosysteem beschikbaar gesteld worden.

De verordening schrijft voor dat er op een positieve manier rekening moet worden gehouden met het schriftelijk bewijs en het *exit report* door de markttoezichtautoriteiten, aangemelde instanties, de Commissie en de AI-board bij de uitoefening van hun taken op grond van de verordening. Er is echter geen verplichting voor aanbieders om deze documenten te gebruiken om aan te tonen dat zij aan de verordening voldoen. Tevens is er geen verplichting voor de bevoegde autoriteiten om een eventueel oordeel in het exit report of het schriftelijk bewijs op te volgen. Daarom zijn het *exit report* en een schriftelijk bewijs naar Nederlands recht niet aan te merken als een besluit in artikel 1:3, eerste lid, van de Awb.

6. Uitvoeringswijze in Nederland en inhoud uitvoeringswetsvoorstel AI-verordening
In 2023 en 2024 is een verkenning en pilot voor de uitvoering van de *regulatory sandbox* in Nederland uitgevoerd, onder coördinatie van het ministerie van Economische Zaken en Klimaat, de AP en de RDI en in samenwerking met een brede groep toezichthouders. De uitkomsten van deze trajecten zijn vastgelegd in een 'vormvoorstel *regulatory sandbox',* die de basis heeft gevormd voor de uitwerking in wetgeving.[3] Relevante uitgangspunten uit het vormvoorstel zijn: Nederland krijgt één multisectorale *sandbox*, waarbij alle markttoezichtautoriteiten zijn betrokken, met één toegangspunt en de markttoezichtautoriteiten bieden zelf geen fysieke of technische infrastructuur aan voor het testen en ontwikkelen van AI.

[3] https://www.autoriteitpersoonsgegevens.nl/documenten/vormvoorstel-nederlandse-regulatory-sandbox.

In de verkenning en de pilot is verder gebleken dat er bij diverse organisaties een grote behoefte bestaat aan advies over de uitvoering van de AI-verordening, op een wijze waarvoor geen intensieve begeleiding, toezicht en ondersteuning in de *AI-regulatory sandbox* vereist is. Dit sluit ook aan bij de algemene taak die de markttoezichtautoriteiten hebben om advies over de AI-verordening te verstrekken (artikel 70, achtste lid, AIV). Voor dergelijke eenvoudige begeleiding is geen formele procedure vereist, zoals een besluit over de toegang op basis van toegangs- en selectiecriteria. Dit zal in de praktijk wel de vraag doen opkomen of het mogelijk is om een zuiver onderscheid te maken tussen het geven van regulier advies en een traject in het kader van de AI-regulatory sandbox, waarbij sprake is van begeleiding, toezicht en ondersteuning in een gecontroleerde omgeving, zeker als beide typen vragen op één plek worden ingediend.

In het ontwerp voor de uitvoeringswet verordening artificiële intelligentie, zoals dat is gepubliceerd voor internetconsultatie[4], wordt aan de aangewezen markttoezichtautoriteiten[5] in Nederland opdracht gegeven om gezamenlijk zorg te dragen voor het opzetten van één AI-testomgeving voor regelgeving, overeenkomstig het vormvoorstel. De markttoezichtautoriteit, wier markttoezicht de meeste raakvlakken heeft met het desbetreffende AI-systeem, verzorgt de begeleiding binnen de *AI-regulatory sandbox*. In het wetsvoorstel een grondslag opgenomen om bij of krachtens algemene maatregel van bestuur nadere regels te kunnen stellen. Deze regels zullen betrekking hebben op de toelatings- en selectiecriteria, de procedures en de geldende voorwaarden voor deelname aan en beëindiging van een *sandbox*-traject. Daarnaast kunnen regels worden gesteld over de verplichtingen van deelnemers vergoedingen voor deelname.

7. Analyse van de AI-regulatory sandbox
De eisen de AI-verordening stelt aan AI-systemen zijn erg abstract geformuleerd. De eisen zijn ook complex, bijvoorbeeld waar het gaat om het wegnemen van risico's voor de bescherming van (potentieel botsende) grondrechten. Dit zorgt ervoor dat deze eisen lastig te vertalen zijn naar de praktijk. Om die reden zullen de eisen worden vertaald naar concreet geformuleerde normen, toegesneden op specifieke soorten AI-systemen (normalisatie). Dit biedt echter geen soelaas voor nieuwe innovatieve AI-systemen, omdat daar nog geen ontwikkelde normen voor zullen zijn. Voor die groep nieuwe innovatieve AI-systemen is de *AI-regulatory sandbox* bedoeld. Dit helpt de betrokken ontwikkelaars, omdat zij vooraf geholpen worden door de toezichthouder, in plaats van achteraf geconfronteerd te worden met toezicht en handhaving. Dit dient de (rechts)zekerheid en bevordert zo innovatie. Markttoezichtautoriteiten krijgen op hun beurt beter inzicht in innovaties op dit terrein en kunnen de inzichten gebruiken in hun toezicht, maar bijvoorbeeld ook als input voor het herijken van normen (*regulatory learning*). Gezien de abstracte formulering van de eisen aan AI-systemen met een hoog risico, is niet te verwachten dat de uitkomsten van de *AI-regulatory sandboxes* snel zullen leiden tot aanpassing van die eisen in de verordening zelf.

Regulatory sandboxes bestaan in vele soorten en maten. Vaak bieden zij de mogelijkheid om (tijdelijk) af te wijken van bestaande wet- en regelgeving, met het oog op het testen van innovatieve producten of diensten in de praktijk. Bij de *AI-regulatory sandbox* is daar echter geen sprake van, omdat die geen afwijking van de verplichtingen van de AI-verordening mogelijk maakt. Dat is te verklaren uit het feit dat het ontwikkelen en testen van nieuwe AI-systemen plaatsvindt op een moment dat de verplichtingen uit de AI-verordening nog niet van toepassing zijn; de AI-systemen hoeven pas formeel aan de essentiële eisen uit de verordening te voldoen op het moment dat ze op de markt worden gebracht of voor het eerst in gebruik genomen. Dat kan echter alleen, als bij het ontwikkelen en testen toepassing wordt gegeven aan die essentiële eisen. Het instrument 'testen onder reële omstandigheden' staat wel afwijking toe, maar de verordening geeft geen concrete aanwijzingen over hoe dit instrument in de praktijk kan en moet worden toegepast binnen de *AI-regulatory sandbox*.

Gezien de algemene en abstracte bepalingen over de regulatory sandbox valt te verwachten dat er verschillen zullen bestaan tussen lidstaten bij de invulling van de *AI-regulatory sandboxes*. Die

[4] https://www.internetconsultatie.nl/uaiv/b1
[5] Hoewel artikel 57, eerste lid, van de verordening spreekt over "bevoegde autoriteiten", ligt het in de rede dat enkel de markttoezichtautoriteiten bij de *regulatory sandbox* zijn betrokken en dus niet de aanmeldende autoriteiten. Daarom wordt deze taak in het voorstel enkel opgedragen aan de markttoezichtautoriteiten, te weten: de RDI, AP, IGJ, ILT, NVWA, AFM, DNB en de NLA.

verschillen kunnen gevolgen hebben voor de praktijk, omdat een aanbieder zijn bewijs van de resultaten en het *exit report*, verkregen in de ene lidstaat, kan gebruiken in de andere lidstaten waar hij zijn AI-systeem op de markt brengt. Andere lidstaten kunnen bijvoorbeeld besluiten om wel fysieke of technische infrastructuur en om een grote rol te geven aan het testen onder reële omstandigheden. De juridische gevolgen zullen enigszins beperkt zijn, omdat het *exit report* maar beperkte waarde heeft, markttoezichtautoriteiten dienen er alleen op een positieve wijze rekening mee te houden. Maar dat zou tegelijkertijd ook de waarde van het instrument kunnen beperken.