

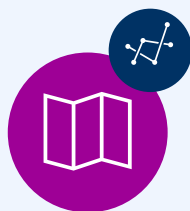


# Werkagenda coördinerend algoritmetoezicht in 2024

Algoritmes en AI raken steeds sterker verweven met onze maatschappij. Dit heeft economische en maatschappelijke waarde, maar brengt ook gevaren met zich mee. Als coördinerend algoritmetoezichthouder ambieert de Autoriteit Persoonsgegevens (AP) bij te dragen aan een stevige bescherming van publieke waarden en grondrechten bij de ontwikkeling en inzet van algoritmes en AI. In 2024 richten we ons op vier belangrijke aandachtsgebieden: transparante algoritmes, auditering, governance en het voorkomen van discriminerende algoritmes. Ook bereiden we ons voor op de AI Act en onderzoeken we de risico's van generatieve AI, zowel voor de korte als lange termijn.

De AP is sinds 2023 coördinerend toezichthouder op algoritmes en AI en waakt voor de risico's van algoritmes en AI voor publieke waarden en grondrechten. Hiervoor onderneemt de Directie Coördinatie Algoritmes (DCA) activiteiten om risico's en effecten te signaleren, te analyseren en periodiek te rapporteren. Ook wordt gewerkt aan het versterken van de samenwerking tussen toezichthouders en de (gezamenlijke) uitleg over, en inzichten in bestaande en nieuwe wetgeving over algoritmes en AI.

In 2024 stelt de DCA een aantal activiteiten centraal. Met (discussie)documenten en bijeenkomsten wordt ervoor gezorgd dat het AI- & algoritmetoezicht in Nederland zich ontwikkelt en aansluit op ontwikkelingen op het gebied van internationale verdragen, wetgeving, normen, en standaarden die gaan gelden voor de ontwikkeling of het gebruik van algoritmes en AI. De activiteiten komen samen in vijf thema's:



**Algoritmerisico's  
in kaart**



**Versterken van  
samenwerking**



**Beleid en  
guidance**



**Coördinatie  
voorbereiding AI Act**



**Internationale  
afspraken**



In 2024 publiceren we de RAN twee keer en betrekken we de gehele samenleving voor input en dialoog over AI- en algoritmerisico's

## 1. Risico's van algoritmes en AI in kaart

- De AP signaleert, analyseert en rapporteert AI- & algoritmerisico's. De uitkomsten worden tweemaal per jaar publiek gedeeld in de **Rapportage AI- & algoritmerisico's Nederland (RAN)**. Door middel van de RAN ontstaat er vroeg zicht op risico's en effecten door thema's, incidenten, ontwikkelingen en het overkoepelend risicobeeld te onderzoeken, beschrijven en publiceren.
- Een vast team verdiept zich doorlopend in risico's, incidenten, technische ontwikkelingen en ontwikkelingen in wetgeving of beheersing van algoritmes en AI. In 2024 versterken we mogelijkheden voor **dialoog met de samenleving** zodat bijvoorbeeld input mogelijk is op onderwerpen voor de RAN.



In 2024 vergroten we de activiteiten van de **Algoritme & AI-Kamer (AAK)**, dragen we bij aan een stelsel van expertise en verkennen we hoe de samenwerking tussen meldpunten versterkt kan worden

## 2. Versterken van samenwerking tussen toezichthouders en meldpunten

- Algoritmes en AI spelen in vrijwel alle sectoren en domeinen, daarom is samenwerking tussen toezichthoudende en controlerende organisaties van belang. Dit doet de AP als regisseur van de **Algoritme & AI-Kamer (AAK)** van het Samenwerkingsplatform Digitale Toezichthouders (SDT) en door bij te dragen aan de Brede Werkgroep Toezicht op AI.
- De AP draagt bij aan een **stelsel van expertise** over algoritmes en AI bij o.a. overheidsorganisaties, auditing-bureau's, toezichthouders, wetenschap en het maatschappelijk middenveld. Zo ontstaat zicht op 'best practices'. In 2024 organiseert de DCA rondetafelgesprekken, workshops, voorlichting en presentaties aan (overheids)organisaties en burgers. Voor specialisten organiseren we het jaarlijkse algoritmeseminar. Door middel van deze bijeenkomsten krijgen we dieper inzicht in de knelpunten, uitdagingen en negatieve effecten in de praktijk.
- De AP ambieert de **samenwerking tussen meldpunten** te versterken op het terrein van incidenten die samenhangen met algoritmes en AI. In 2024 onderzoeken we hoe de samenwerking en uitwisseling tussen meldpunten versterkt kan worden om meldingen van burgers waarbij een algoritme of AI-systeem betrokken is op de beste manier te kunnen opvolgen.



Transparantie-eisen, auditerings-raamwerken, het voorkomen van discriminatie en de benodigde governance van algoritmes zijn in 2024 onze aandachtspunten voor beleid en guidance

## 3. Werken aan beleid en guidance voor verantwoorde inzet van AI en algoritmes

- De AP werkt aan **domeinoverstijgende beleidsanalyse en guidance**, gericht op een versterkte bescherming van publieke waarden en grondrechten bij de ontwikkeling en inzet van algoritmes.
- In 2024 richt de DCA zich op de volgende **vier aandachtsgebieden**: uitgangspunten voor transparantie over algoritmes, raamwerken voor periodieke auditering van algoritmes, principes voor het voorkomen van discriminatie door algoritmes (bias-toetsingen) en governance van organisaties voor de benodigde risicobeheersing voor algoritmes.
- In 2024 gaan we samen met andere toezichthouders een verkenning doen naar de risico's van **generatieve AI**. Er zal bekeken worden op welke wijze deze risico's voor de korte en lange termijn beheerst kunnen worden.



We werken samen met andere toezichthouders aan voorstellen over de inrichting van het nationale toezicht op de AI Act en hebben bijzondere aandacht voor het toezicht op verboden AI per jaareinde 2024

## 4. Coördinatie voorbereiden van het toezicht op de AI Act

- Er is in de EU een akkoord over de **AI Act**. Deze verordening geeft regels waaraan met name ontwikkelaars van AI-systemen moeten voldoen. Het doel van de AI Act is dat mensen en bedrijven in de EU kunnen rekenen op veilige, transparante, traceerbare, niet-discriminerende en duurzame AI-systemen die onder toezicht staan.
- De AP coördineert samen met de RDI de **toezichtvoorbereiding** op de AI Act. Dit komt samen in inrichtingsadviezen aan de relevante ministeries. Hierbij zijn meer dan tien toezichthouders betrokken. Een eerste tussenadvies is in december 2023 gedeeld. Een volgend tussenadvies en een eindadvies volgen later dit jaar.
- De AI Act **bevat ook verboden AI**, bijvoorbeeld op het gebied van social credit-scoring, manipulatie van mensen, en bepaalde vormen van biometrische identificatie voor rechtshandhaving. Deze verboden zullen al snel na de inwerkingtreding van de verordening van toepassing zijn, rond jaareinde 2024. De DCA treft hiertoe specifieke voorbereidingen.

## Tijdslijn infasering verplichtingen onder de AI Act

- **Zomer 2024**
  - AI Act van kracht en start overgangperiode
- **Eind 2024**
  - Verbod op bepaalde AI-systemen is van toepassing
- **Eerste helft 2025**
  - Praktijkcode voor AI-modellen voor algemene doeleinden, op te stellen door het Europese AI Office, moet gereed zijn
- **Medio 2025**
  - (Nationale) markttoezichthouders en aanmeldende autoriteiten zijn aangewezen
  - Europese AI Board is opgericht
  - Regelgeving voor AI-modellen voor algemene doeleinden is van toepassing
  - Bepalingen over sancties, waaronder boeteregels voor verboden AI, zijn van toepassing
- **Medio 2026**
  - Bepalingen over hoog risico-AI in Bijlage III (AI-systemen die niet onder bestaand markttoezicht vallen) zijn van toepassing
  - De Nederlandse regulatory sandbox moet 'open' zijn
  - Alle overige onderdelen van de AI Act zijn van toepassing, met uitzondering van hoog-risico AI-toepassingen in producten in Bijlage I
- **Medio 2027**
  - Bepalingen over hoog risico-AI in Bijlage I zijn van toepassing (na 36 maanden)



In 2024 investeren we in internationale samenwerking en afspraken op het gebied ten behoeve van kennisuitwisseling, incidentmonitoring, best practices en het gezamenlijk uitwerken en implementeren van Europese regels

## 5. Bouwen aan internationale samenwerking en afspraken

- De veilige inzet van AI en algoritmes is mede afhankelijk van stevige internationale standaarden en krachtig en tijdig grensoverschrijdend toezicht. Nederland loopt voorop in de inrichting van het AI- en algoritmetoezicht. Daarom zullen we ook in 2024 verder bouwen aan **internationale samenwerking**. De focus ligt op het tot stand brengen van kennisuitwisseling, gezamenlijke incidentmonitoring, ontwikkeling van best practices en het gezamenlijk uitwerken en implementeren van Europese regelgeving en standaarden. Dit realiseren we door kennis te delen en samen te werken met (nieuwe) internationale partners
- zoals het Europese AI Office, de nog op te richten Europese AI Board, UNESCO en de OECD.
- Naast deelname aan werkgroepen en congressen, investeert de DCA ook in de **bilaterale samenwerking** met specifieke internationale partners. Hiermee wordt een internationaal netwerk opgebouwd dat nu en in de nabije toekomst kan bijdragen aan een beter inzicht in risico's en een eenduidige uitleg van wetgeving en standaarden.

# Het coördinerend algoritmetoezicht: een taak in opbouw

## Taak in opbouw

De AP is sinds 2023 coördinerend toezichthouder op algoritmes en AI en waakt voor de risico's voor publieke waarden en grondrechten. Deze activiteiten worden uitgevoerd binnen de DCA. We richten ons op vroegtijdige risicosignalering, het versterken van bestaande samenwerkingen bij AI- en algoritmetoezicht, en guidance over de beheersing van AI- en algoritmerisico's. Ook gaan we in 2024 verder aan de slag met het versterken van de publieksvoorlichting over

de risico's van algoritmes en AI. Aanvullend wordt gewerkt aan de uitwerking van en het toezicht op specifieke bindende AI- en algoritmeregelgeving zoals de AI Act. Uitgangspunt bij de werkagenda en het beoordelingskader dat de DCA volgt voor het analyseren van AI- en algoritmerisico's is dat er voor veilige algoritmische processen drie verdedigingslijnes nodig zijn. Onze visie op AI- en algoritmerisico's werken we in 2024 verder uit.

## Drie verdedigingslijnes voor veilige algoritmische processen



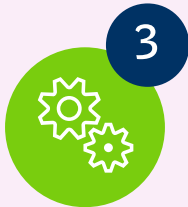
### Continue beheersen van (kans op) mogelijke fouten en schade in algoritmische processen

Hiervoor is het bijvoorbeeld wenselijk dat, alvorens een algoritme in gebruik wordt genomen, er een bias-toetsing plaatsvindt om te helpen voorkomen dat een algoritme discrimineert. Periodieke auditing van een AI-systeem of algoritme helpt om te waarborgen dat ook gedurende het gebruik ervan geen inbreuk op grondrechten plaatsvindt. De eisen die de AI Act stelt aan hoog-risico systemen ondersteunen de eerste verdedigingslinie.



### Snel corrigeren door signalen over fouten op te volgen

Dit vereist een organisatie-structuur en werkprocessen waar individuele algoritmefouten gecorrigeerd kunnen worden en eventueel het gehele systeem snel aangepast kan worden. Transparantie en uitlegbaarheid van algoritmes en AI is een voorwaarde voor deze tweede verdedigingslinie om personen weerbaar te maken. Een samenwerkings-stelsel tussen (publieke) meldpunten draagt ook bij aan deze verdedigingslinie.



### Zodanige stelselinrichting dat impact van fouten in algoritmische processen beperkt is

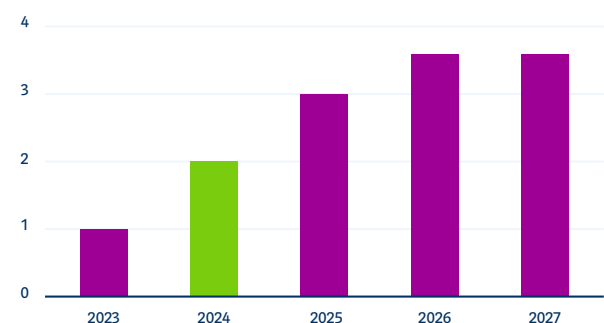
Ondanks het minimaliseren en beheersen van de kans op fouten door de eerste twee verdedigingslijnes, kan het toch misgaan. Door het werken aan de derde verdedigingslinie wordt daar van tevoren rekening mee gehouden. Overal waar algoritmes en AI worden ingezet kunnen deze systemen, of de mensen die ze gebruiken, fouten maken. Bij het werken aan deze derde verdedigingslinie moet ook rekening gehouden worden met algoritmevervorming: de mogelijkheid dat algoritmes bewust of onbewust de omgeving veranderen waarin ze worden ingezet.

## Financiering en afspraken

- De AP ontvangt voor het oppakken van de nieuwe taak als coördinerend algoritmetoezichthouder vanaf 2023 1 mln. euro, oplopend naar structureel 3,6 mln. euro in 2026. In 2024 is het budget 2 mln. euro.
- Vanwege de coördinerende rol op het digitaliseringsdomein is het ministerie van BZK het beleidsmatig aanspreekpunt voor het coördinerend algoritmetoezicht. De minister voor Rechtsbescherming houdt toezicht op de financiële middelen die de AP via de begroting van het ministerie van JenV ontvangt.
- Op 21 december 2023 heeft staatssecretaris Van Huffelen (Koninkrijksrelaties en Digitalisering), mede namens de minister van EZK en de minister van Rechtsbescherming, de Tweede Kamer geïnformeerd over de voortgang en activiteiten van de DCA in 2024 (Kamerstuk 26643, nr. 1110).

## Budget Coördinerend Algoritmetoezicht (DCA)\*

IN EUR MLN. (PRIJSPEL 2022)



\*) EXCLUSIEF EVENTUELE AANVULLENDE VOORBEREIDENDE MIDDELEN TEN BEHOEVE VAN COÖRDINEREND EN MARKTTOEZICHT ONDER DE AI ACT