HANDREIKING

Governance voor een verantwoorde toepassing van algoritmen

25 januari 2022

HANDREIKING

Governance voor een verantwoorde toepassing van algoritmen

Deze handreiking is opgesteld in opdracht van het programma publieke controle op algoritmes. Dit programma is een samenwerking van: de gemeenten Rotterdam (regie-voerder), Amsterdam, Utrecht, Den Haag en Haarlemmermeer, de provincies Zuid-Holland, Noord-Brabant en Limburg, het Kadaster, de Nationale Politie, Rijkswaterstaat, de VNG, de Unie van Waterschappen en het ministerie van Binnenlandse Zaken en Koninkrijksrelaties. De volgende adviseurs van Berenschot zijn de opstellers van de handreiking:

Wubbo Wierenga
Wouter Verbeek
Mariam Al-Saqaff
Arne van Rooijen

25 januari 2022

Inhoudsopgave

1. Inleiding

1.1 Doel van de handreiking

Deze handreiking heeft tot doel overheden te ondersteunen bij het inrichten van de governance op algoritmetoepassingen. Deze governance moet ertoe leiden dat algoritmetoepassingen op een verantwoorde manier worden ingezet. Het belang van verantwoorde toepassing van algoritmen door overheden is bijna niet te overschatten. In een democratische rechtsstaat moeten burgers, bedrijven, maatschappelijke organisaties en medeoverheden erop kunnen vertrouwen dat iedere overheidsorganisatie algoritmen op een goede manier inzet. Dit vertrouwen kan alleen ontstaan en doorgroeien als de governance op algoritmetoepassingen op orde is. Dit vereist een goed ingerichte governance. Met deze handreiking willen de overheidsorganisaties[1] in het programma *publieke controle op algoritmes* deze governance versterken.

1 De gemeentes Rotterdam, Amsterdam, Utrecht, Den Haag en Haarlemmermeer, de provincies Zuid-Holland, Noord-Brabant, Limburg, het Kadaster, de Nationale politie, Rijkswaterstaat, de VNG, de Unie van Waterschappen en het ministerie van Binnenlandse Zaken en Koninkrijksrelaties.

1.2 Voor wie is de handreiking bedoeld?

De handreiking is een referentiedocument voor alle overheden in Nederland. Dit biedt kansen en beperkingen. De grote winst van deze handreiking is dat het is ontwikkeld met hulp van uiteenlopende typen overheidsorganisaties. Gemeenten, provincies, uitvoeringsorganisaties op Rijksniveau en het ministerie van BZK waren betrokken. Daarmee is geborgd dat de handreiking breed gedragen wordt en inzetbaar is bij verschillende typen overheidsorganisaties. Bovendien, in een tijd waarin (digitale) interbestuurlijke samenwerking steeds belangrijker wordt, zijn gezamenlijk geformuleerde definities behulpzaam om samen aan grote opgaven te werken.

Tegelijkertijd leidt deze brede toepasbaarheid van het referentiedocument tot een beperking. De verschillende typen overheidsorganisaties staan voor verschillende opgaven en kennen uiteenlopende bestuurlijke- en organisatorische structuren. De handreiking biedt referenties om uniformiteit te bevorderen, maar leidt niet tot een blauwdruk voor alle overheidsorganisaties in Nederland. Dit is ook niet passend. Er zijn goede redenen waarom overheden verschillend zijn georganiseerd.

De doelgroep van deze handreiking is primair medewerkers binnen overheidsorganisaties die het beleid over digitalisering, datagebruik en algoritmetoepassingen vormgeven. Dit betekent dat deze handreiking veronderstelt dat de lezer kennis heeft over governance in overheidsorganisatie, algoritmetoepassingen of beide. Het heeft dus een wat gevorderd instapniveau en is bedoeld om deze professionals te ondersteunen in hun werk.

1.3 Opzet van de handreiking

De handreiking is een referentiedocument. Dit betekent dat het zowel geschikt is om van voor naar achter door te lezen, maar ook geschikt is om gericht informatie op te zoeken. Daarom is gekozen om de handreiking vorm te geven aan de hand van drie kernvragen over governance. Deze drie kernvragen zijn:

1. Wanneer is de toepassing van een algoritme door een overheidsorganisatie verantwoord? (hoofdstuk 2)
   Dit hoofdstuk bevat ***definities*** van de kernbegrippen in de governance van algoritmetoepassingen. Het gaat daarbij om de definities van algoritme, algoritmetoepassing en verantwoorde algoritmetoepassing.

2. Wat zijn de hoofdelementen van de governance voor de verantwoorde toepassing van algoritmen? (hoofdstuk 3)
   De definities in hoofdstuk 2 zijn breed. Dit betekent dat een grote variëteit aan algoritmetoepassingen die overheden inzetten, eronder vallen. Sommige van deze toepassingen zijn eenvoudig en risicoloos, andere zijn juist complex of ondoorzichtig. Daarnaast volgt ook uit hoofdstuk 2 dat een algoritmetoepassing veranderlijk is en een levenscyclus kent. Een governance voor de verantwoorde toepassing van algoritmen moet daarom rekening houden met:
   - de variëteit aan algoritmetoepassingen binnen de organisatie *en*
   - de verschillende fasen in de levenscyclus van de algoritmetoepassing.

   In hoofdstuk 3 worden de ***hoofdelementen*** van een dergelijke ***governance*** beschreven. Centraal staat een gelegitimeerd beheersplan, dat op sleutelmomenten in de levenscyclus van de algoritmetoepassing wordt gecheckt en herijkt. Het hoofdstuk begint met een overzicht van de sleutelmomenten in de levenscyclus van een algoritmetoepassing. Daarna wordt per sleutelmoment geschetst wat moet gebeuren met het beheersplan en de legitimering daarvan.

3. Welke praktische handvatten zijn er beschikbaar die kunnen bijdragen aan een duurzame governance op algoritmetoepassingen? (hoofdstuk 4)
   In hoofdstuk 4 wordt een aantal ***praktische handvatten*** gedeeld die tijdens het onderzoek naar voren kwamen. Het is geen uitputtende lijst, maar dient ter inspiratie voor overheidsorganisaties die met de inrichting van de governance aan de gang willen.

2. Definities voor een verantwoorde toepassing van algoritmen

2.1 Inleiding en centrale vraag

Zoals beschreven in de inleiding is een goed ingerichte governance een voorwaarde voor het vertrouwen van de samenleving in de toepassing van algoritmen door overheden. Een goed ingerichte governance vereist dat helder is wat de overheid verstaat onder 'verantwoorde toepassing van algoritmen'. In dit hoofdstuk wordt stapsgewijs een antwoord gegeven op deze vraag. Dit begint met definities die gebruikt kunnen worden als referentie voor verschillende overheidsorganisaties. In dit hoofdstuk worden achtereenvolgend definities gegeven van:

- algoritme en algoritmetoepassing (en het belang van dit onderscheid)
- de levenscyclus van een algoritmetoepassing
- de verantwoorde toepassing van algoritmetoepassing

2.2 De definitie van algoritme en algoritmetoepassing (en het belang van dit onderscheid)

Definitie algoritme:
Een algoritme is software die data-analyse, statistiek of logica kan uitvoeren. Dit kan variëren van een simpele regels tot niet-lineaire modellen die doen denken aan neurale netwerken.

Over de precieze definitie van algoritme en algoritmetoepassing is inmiddels veel literatuur beschikbaar en ook in beleidsstukken worden verschillende definities gehanteerd. De definities in deze handreiking zijn gebaseerd op de literatuur en deze beleidsstukken. Uiteindelijk was leidend dat de definities geschikt moeten zijn om de governance voor een verantwoorde toepassing van algoritmen binnen de overheid mogelijk te maken.

Definitie algoritmetoepassing:
Binnen deze handreiking is een algoritmetoepassing een proces dat op hoofdlijnen bestaat uit vier stappen:

1. Het gebruik van gegevens
2. De werking van een algoritme
3. Een resultaat (adviezen, voorspellingen, besluiten)
4. Impact op de eigen organisatie, burgers, bedrijven, maatschappelijke organisaties en medeoverheden

Om tot een verantwoorde inzet van algoritmen te komen, moet de governance niet alleen gericht zijn op de software (het algoritme zelf), maar ook op het proces waarin het algoritme een rol heeft (de algoritmetoepassing). Door te redeneren vanuit een algoritmetoepassing wordt het mogelijk om bij iedere processtap relevante aandachtspunten voor de governance inzichtelijk te maken. Om dit punt te illustreren, wordt hieronder per processtap één voorbeeld gegeven van een aandachtspunt voor de governance.

- De eerste processtap is het gebruik van gegevens. Dit kunnen bijvoorbeeld persoonsgegevens zijn, waardoor de algoritmetoepassing binnen het bereik van de AVG komt. Dit leidt tot eisen aan (en daarmee aandachtspunten voor) de governance van dit algoritme.

- De tweede stap is de werking van een algoritme. Een overheid moet zich kunnen verantwoorden over de inzet van een algoritme, ook als het algoritme niet een eigen product is, maar ingekocht wordt bij een externe leverancier. Daar moet in de governance rekening mee worden gehouden.

- De derde processtap is het resultaat van het algoritme. Een voorbeeld van een aandachtspunt is dat de status van het resultaat duidelijk moet zijn. Als het een besluit in de zin van de Algemene wet bestuursrecht is, moet ook aan de toepasselijke procedurele eisen uit de Awb worden voldaan. Deze moeten een plaats krijgen in de governance.

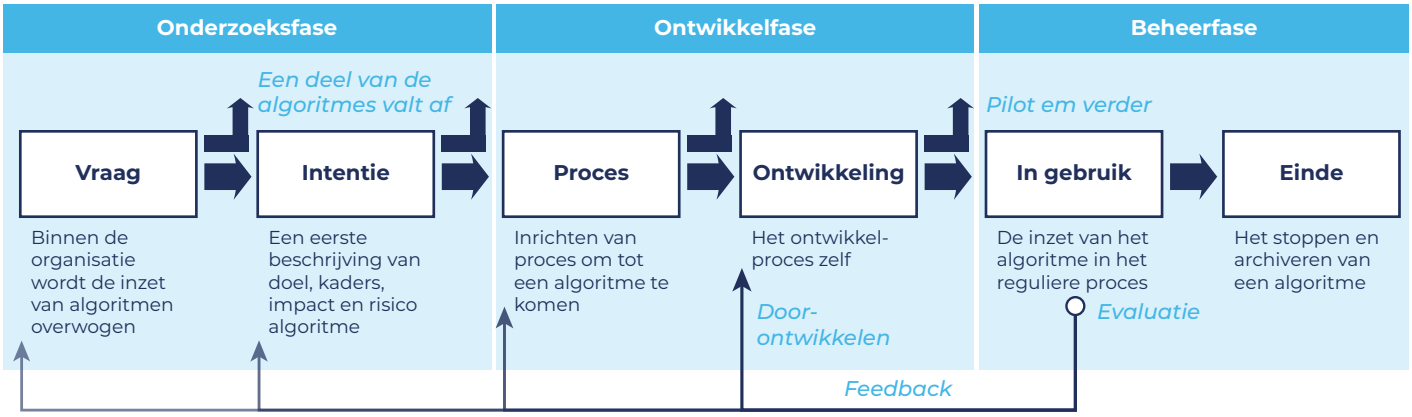
- De vierde processtap is de impact van de algoritmetoepassing. Het ligt voor de hand dat in de governance van een algoritmetoepassing de monitoring van de impact van de toepassing op bepaalde groepen burgers een plaats moet krijgen.

Kortom, deze definitie van algoritmetoepassing maakt het mogelijk om op een gestructureerde manier de eisen en aandachtspunten voor de governance van de inzet van algoritmen uit te werken. Daarom wordt in deze handreiking met deze brede definitie gewerkt.

2.3 De definitie van de levenscyclus van een algoritmetoepassing

Een centraal punt in de volgende paragraaf (2.4) is dat de governance van een algoritmetoepassing moet meebewegen met de levenscyclus van een algoritmetoepassing. Het is daarom belangrijk om eerst deze levenscyclus te definiëren. In deze handreiking wordt de levenscyclus van een algoritmetoepassing als volgt (op hoofdlijnen[2]) weergegeven.

Figuur 1. **Levenscyclus van algoritmetoepassing.**



Binnen de drie fasen zijn elk twee onderdelen geformuleerd.

- Binnen de ***onderzoeksfase*** wordt eerst de vraag geformuleerd, en wordt überhaupt de inzet van een algoritme overwogen. Vervolgens wordt de intentie uitgewerkt, hierbij wordt een eerste beschrijving van doel, kaders, acceptatiecriteria, de mogelijke impact en het risico van het algoritme vastgesteld.

- Tijdens de ***ontwikkelfase*** wordt eerst het proces ingericht, waarbij gekeken wordt naar bijvoorbeeld de samenstelling van het team. Tijdens het ontwikkelingsonderdeel vindt de ontwikkeling van het algoritme zelf plaats, dit kan intern gebeuren, maar natuurlijk ook door een commerciële partij. Ook wanneer een bestaand product wordt ingekocht – en de ontwikkelfase feitelijk voorbij is – is aan te raden de ontwikkelfase te onderscheiden van de beheerfase. De laatste stap in de ontwikkelfase is namelijk dat de algoritmetoepassing helemaal geschikt wordt gemaakt voor een specifiek regulier proces in een specifieke overheidsorganisatie en tijdens deze laatste stap kunnen keuzes worden gemaakt die van belang zijn voor de governance van de toepassing.

- De laatste fase is de ***beheerfase***, hierbij is het algoritme in gebruik. Uiteindelijk wordt het algoritme ook stopgezet en gearchiveerd.

Een belangrijk onderdeel van de levenscyclus zijn ***feedbackloops.*** Op basis van de ervaringen met de algoritmetoepassing worden algoritmetoepassingen doorontwikkeld of juist beëindigd. Feedback, doorontwikkeling en beëindiging zijn onderdeel van de gehele levenscyclus van een algoritmetoepassing. Wel vraagt feedback in de beheerfase extra aandacht. De praktijk leert namelijk dat algoritmetoepassingen in de beheerfase uit zicht kunnen raken.

De inrichting van de governance voor de verantwoorde toepassing van algoritmen is afhankelijk van het moment in de levenscyclus van een algoritme. De ontwikkelfase vraagt andere processen, mensen en middelen dan de beheerfase. Het is daarom belangrijk deze fases af te bakenen. De precieze afbakening kan per organisatie en per algoritmetoepassing verschillen. Vooral de afbakening tussen de onderzoeks- en ontwikkelfasen en de beheerfase is belangrijk.

2 Net als bij de definitie van algoritme, is voor de definitie van de levenscyclus gekozen voor een definitie die geschikt is om tot een governance voor verantwoorde toepassing van algoritmen te komen. Het is goed mogelijk dat in de wetenschap of bij ontwikkelaars van algoritmetoepassingen andere indelingen worden gebruikt.

2.4 De definitie van verantwoorde toepassing van algoritmen

Op basis van de definities van algoritme, algoritmetoepassing en de levenscyclus van algoritmen is het mogelijk om een verantwoorde toepassing van algoritmen te definiëren. Er is sprake van verantwoorde toepassing van algoritmen wanneer:

- de algoritmetoepassing juridisch, democratisch en maatschappelijk is gelegitimeerd ***(legitimiteit);***
- de governance van de algoritmetoepassing gebaseerd is op een beheersplan op basis van ethische waarden (bijvoorbeeld uit de CODIO) ***(waardengebaseerd beheersplan)***;
- de legitimering en het beheersplan meebewegen met de levenscyclus van een algoritmetoepassing ***(check en herijking).***

Deze gekozen definities van verantwoorde toepassing

2.4.1 Verantwoorde toepassing: het borgen van drie soorten legitimiteit

Voor al het overheidshandelen, dus ook de toepassing van algoritmen, geldt dat deze gelegitimeerd moet zijn. Vrij naar Stolk, Wesseling en Van de Beek in *de jacht op publieke waarde (2021)* worden in deze handreiking drie soorten legitimiteit onderscheiden:

- ***Juridische (rechtstatelijke) legitimiteit.*** De mate waarin de uitkomsten en het overheidshandelen voldoen aan de mensenrechtenverdragen, de Grondwet en bindende wet- en regelgeving.[3]
- ***Democratische legitimiteit.*** De mate waarin de uitkomsten en het overheidshandelen kunnen rekenen op steun van de bestuurlijke en politieke entiteiten.
- ***Maatschappelijke legitimiteit.*** De mate waarin de uitkomsten en het overheidshandelen kunnen rekenen op steun van burgers, bedrijven en maatschappelijke organisaties, die de effecten van het overheidshandelen ervaren.

De governance van algoritmetoepassing heeft (onder meer) als doel om ervoor te zorgen dat deze soorten legitimiteit tijdens de hele levenscyclus van de algoritmetoepassing zijn geborgd. In het vervolg van deze handreiking wordt beschreven hoe de governance hierop gericht kan worden.

2.4.2 Verantwoorde toepassing: een waardengebaseerd beheersplan

Een tweede element is het beheersplan. Dit plan bevat drie onderdelen:

- een ***risicoanalyse*** op basis van ethische waarden;
- een ***juridische check***, in ieder geval op de AI-verordening, de Algemene verordening gegevensbescherming (Avg) en de Algemene wet bestuursrecht (Awb);
- een beschrijving van ***beheersmaatregelen*** die volgen uit de risicoanalyse.

2.4.3 Verantwoorde toepassing: check en herijking van het beheersplan en legitimiteit

Een derde element van verantwoorde algoritmetoepassing is dat de legitimering en het beheersplan meebewegen met de levenscyclus van een algoritmetoepassing. In deze handreiking wordt dit vormgegeven door in de levenscyclus van een algoritmetoepassing checks in te bouwen. Tijdens deze checks wordt de waardenafweging in het beheersplan opnieuw bekeken en wordt beoordeeld of de juridische, democratische en maatschappelijke legitimiteit van de algoritmetoepassing nog op orde zijn. Als uit de check blijkt dat het beheersplan of de legitimiteit aandacht behoeven, worden deze herijkt. In hoofdstuk 3 wordt uitgewerkt hoe deze check en herijking in organisatieprocessen binnen de overheid een plek kunnen krijgen.

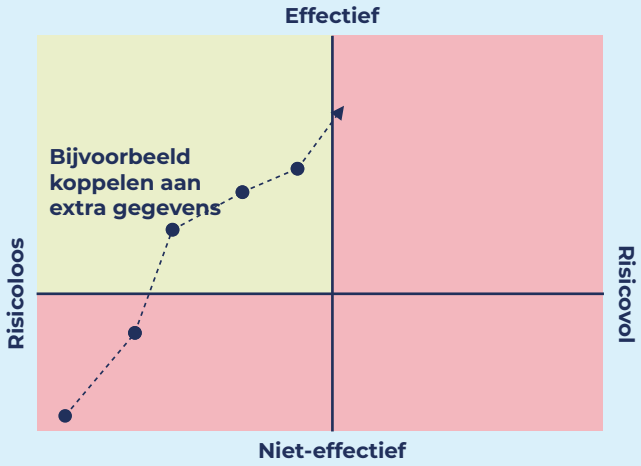
3 In het programma publieke controle op algoritmes heeft Pels Rijcken een praktisch toetsingskader opgesteld dat is gericht op de juridische legitimering van de toepassing van een algoritme.

Waarom zijn checks tijdens de levenscyclus van een algoritmetoepassing noodzakelijk?

In deze handreiking is de vooronderstelling dat checks tijdens de levenscyclus een voorwaarde zijn voor de verantwoorde toepassing van algoritmen. Deze vooronderstelling stoelt op – wat in de literatuur – het emergente karakter van (sommige) algoritmetoepassingen wordt genoemd. **Emergentie** is een proces waarbij de eigenschappen van een algoritmetoepassing geleidelijk veranderen.

Een goed voorbeeld hiervan is de trade-off tussen effectiviteit en risico's in algoritmetoepassing. Figuur 2 geeft hier een beeld van. Wanneer een algoritmetoepassing goed functioneert, bestaat de neiging de effectiviteit te optimaliseren door bijvoorbeeld verbeterde gegevens onderdeel te maken van de toepassing. Deze gegevens beïnvloeden de gehele toepassing en veranderen (mogelijk) daarmee de eigenschappen van de hele toepassing. Dit kan ertoe leiden dat de algoritmetoepassing risicovoller wordt.

Figuur 2. **Voorbeeld trade-off tussen effectiviteit en risico's in algoritmetoepassing en emergentie.**



Idealiter beweegt de legitimiteit en het beheersplan mee met de ontwikkeling van de algoritmetoepassing. Daarom volgt uit het emergente karakter van een algoritmetoepassing dat checks tijdens de hele levenscyclus nodig zijn.

3. De hoofdelementen van de governance: vijf sleutelmomenten en het gelegitimeerde beheersplan

3.1 Inleiding en centrale vraag

De definities van algoritme en algoritmetoepassing in hoofdstuk 2 zijn breed. Dit betekent dat een grote variëteit aan algoritmetoepassingen eronder vallen. Sommige van deze toepassingen zijn eenvoudig en risicoloos, andere zijn juist complex of ondoorzichtig. Bovendien, de risico's van een algoritmetoepassing volgen niet alleen uit de software of de gebruikte gegevens, maar ook uit het doel van de toepassing en de manier waarop de toepassing wordt gebruikt. Daarnaast volgt ook uit hoofdstuk 2 dat een algoritmetoepassing veranderlijk is en een levenscyclus kent.

Dit alles betekent dat een *one-size-fits-all* benadering voor alle algoritmetoepassingen niet leidt tot een doeltreffende of doelmatige governance. Laagrisico algoritmetoepassingen vergen een andere aanpak dan hoogrisico algoritmetoepassing en een algoritmetoepassing in de jeugdzorg vraagt wat anders dan een algoritmetoepassing in de dijkbewaking.

Een overheidsorganisatie staat daarom voor de uitdaging om tot een governance te komen die rekening houdt met:

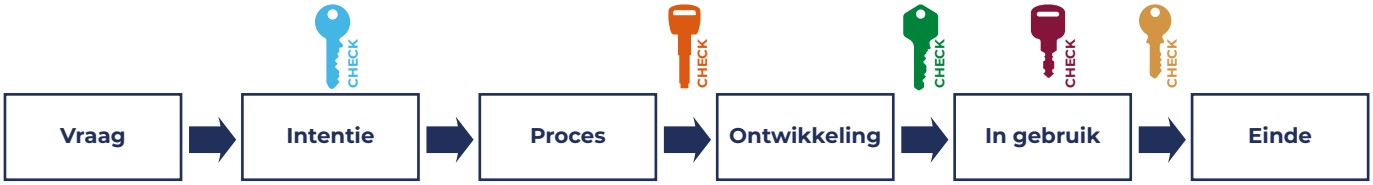
- de variëteit aan algoritmetoepassingen binnen de organisatie *en*
- de verschillende fasen in de levenscyclus van de algoritmetoepassing.

In dit hoofdstuk worden de **hoofdelementen** van een dergelijke **governance** beschreven. Centraal staat een gelegitimeerd beheersplan, dat op sleutelmomenten in de levenscyclus van de algoritmetoepassing wordt gecheckt en herijkt. Het hoofdstuk begint met een overzicht van de sleutelmomenten in de levenscyclus van een algoritmetoepassing. Daarna wordt per sleutelmoment geschetst wat moet gebeuren met het beheersplan en de legitimering daarvan.

3.2 Vijf sleutelmomenten in de levenscyclus van een algoritmetoepassing

Figuur 3 bevat vijf sleutelmomenten in de levenscyclus van een algoritmetoepassing. Deze vijf momenten staan als 'check' in de figuur.

Figuur 3.



De vijf sleutelmomenten zijn:

Sleutelmoment I. Kaderstelling en quickscan
Sleutelmoment II. Voor de start van het ontwikkelproces
Sleutelmoment III. Ingebruikname
Sleutelmoment IV. Wezenlijke wijziging
Sleutelmoment V. Periodieke controle.

Bij elk sleutelmoment staan twee onderwerpen centraal:

1. ***Legitimering.*** In hoeverre is op het betreffende sleutelmoment de algoritmetoepassing legitiem? Relevante vragen daarbij zijn:
   a. *Voor democratische legitimering.* Is de algoritmetoepassing (nog steeds) voldoende democratisch gelegitimeerd? Of is er aanleiding om de volksvertegenwoordiging, de bestuurder of een leidinggevende er (opnieuw) naar te laten kijken?
   b. *Voor juridische legitimering.* Is de algoritmetoepassing (nog steeds) in lijn met wet- en regelgeving? Zijn er veranderingen in wet- en regelgeving geweest die een scherpe blik vragen op de toepassing? Past de precieze toepassing van het algoritme (nog steeds) binnen de wettelijke kaders? Is de algoritmetoepassing nog steeds toegestaan op basis van de AI-verordening?
   c. *Voor maatschappelijk legitimering.* Zijn er onder burgers, bedrijven, maatschappelijke organisaties en medeoverheden (nieuwe) ideeën of opvattingen over de algoritmetoepassing ontstaan?

1. ***Het waardengebaseerde beheersplan.*** In hoeverre is het waardengebaseerde beheersplan (nog) in lijn met de algoritmetoepassing? Relevante vragen daarbij zijn:
   a. *Voor de waardengebaseerd risicoanalyse.* In hoeverre is de risicoanalyse (nog steeds) in lijn met de algoritmetoepassing?
   b. *Voor de juridische check.* In hoeverre zijn alle elementen van de juridische legitimering in de juridische check opgenomen?
   c. *Voor de beheersmaatregelen.* In hoeverre zijn de beheersmaatregelen in het beheersplan (nog) geschikt om de risico's uit de risicoanalyse te mitigeren?

In de bespreking van de sleutelmomenten worden op basis van deze vragen steeds aandachtspunten geformuleerd. Op die manier blijft het beheersplan up-to-date en beweegt de democratische, juridische en maatschappelijke legitimering mee met de levenscyclus van de algoritmetoepassing.

3.3 Sleutelmoment I: kaderstelling en quickscan

In de onderzoeksfase van de algoritmetoepassing wordt de intentie bepaald. Dit is een eerste beschrijving van doel, kaders, acceptatiecriteria, impact en risico van algoritmetoepassing. In deze beginfase van de algoritmetoepassing krijgen de elementen legitimering en beheersplan (voor uitleg: zie paragraaf 2.4) als volgt een plaats:

- *Legitimering.* Er wordt bepaald wie de algoritmetoepassing democratisch legitimeert (volksvertegenwoordiging, bestuurder, leiding ambtelijke organisatie), hoe de juridische legitimering geborgd wordt in het vervolgproces (wanneer en hoe welke juristen betrokken worden) en in hoeverre er aandacht moet zijn voor maatschappelijke legitimering.

- *Waardengebaseerd beheersplan.* In de onderzoeksfase is het (meestal) niet goed mogelijk om een volledig beheersplan op te stellen. De algoritmetoepassing is daarvoor nog te onbepaald. Wel is het mogelijk een eenvoudige ***quickscan*** uit te voeren. Op basis van deze quickscan wordt de inschatting gemaakt in hoeverre later in het ontwikkelproces (sleutelmoment II) een impact assessment moet worden uitgevoerd. Het is raadzaam de resultaten van de quickscan te documenteren zodat het beschikbaar is in de hele levenscyclus van de algoritmetoepassing. In het onderstaande kader staat een voorstel voor de inrichting van de quickscan.

Voorstel voor de inrichting van de quickscan

De quickscan wordt uitgevoerd aan het begin van de onderzoeksfase om te bepalen of de risico's van de algoritmetoepassing een impact assessment tijdens de ontwikkelfase vereisen. De quickscan bestaat uit twee onderdelen: een check op basis van de zes principes in de CODIO en een check op de lijst uit de AI-verordening. Daarmee wordt in de quickscan voorgesorteerd op de risicoanalyse en de juridische check in het beheersplan.

De check op basis van de zes principes in de CODIO

In dit tweede quickscan onderdeel worden op basis van de zes principes van de CODIO zes vragen gesteld. Het doel van deze vragen is dat in de onderzoeksfase alvast wordt stilgestaan bij mogelijke (ethische) risico's.

1. *Participatie:* geeft de te onderzoeken algoritmetoepassing reden om burgers, bedrijven, maatschappelijke organisaties of medeoverheden snel te betrekken?
2. *Maatschappelijke waarde:* is het realistisch dat de algoritmetoepassing wenselijke resultaten oplevert voor de samenleving?
3. *Procedurele rechtvaardigheid:* is er in deze fase al zicht op speciale aandachtspunten waar het gaat om discriminatie, uitlegbaarheid en gebruiksvriendelijkheid van de algoritmetoepassing?
4. *Mensenrechten:* is er in deze fase al zicht op speciale aandachtspunten waar het gaat om privacy, autonomie en de waardigheid van ieder mens?
5. *Bestuurskwaliteit:* in hoeverre vraagt de algoritmetoepassing iets extra's van de organisatie waar het gaat om wendbaarheid, risicobewustheid en veiligheid?
6. *Verantwoordelijkheid:* zijn er speciale aandachtspunten waar het gaat aanspreekbaarheid, controleerbaarheid en menselijke eindverantwoordelijkheid?

De check op de lijst uit de AI-verordening

Op basis van de AI-verordening wordt een quickscan gedaan op de verboden en hoog-risico algoritmetoepassing (de AI-verordening noemt dit AI-systemen). Deze check is op hoofdlijnen en moet, indien nodig, tijdens de ontwikkelfase (sleutelmoment II) worden uitgebreid en verbeterd. In dat geval moet de tekst van de AI-verordening gebruikt worden voor meer precisie.

***Verboden* zijn algoritmetoepassingen:**

1. die kwetsbaarheden bij specifieke groepen personen proberen uit te nutten,
2. social credit-achtige systemen en
3. bepaalde vormen van biometrische identificatie

***Hoog-risico* zijn de volgende algoritmetoepassingen:**

1. voor biometrische identificatie en categorisering die niet verboden zijn
2. voor beheer en exploitatie van kritieke infrastructuur
3. in situaties waarin een persoon in relatie tot de overheid kwetsbaar of afhankelijk is en het algoritme een rol speelt in de beoordeling van die kwetsbare of afhankelijke persoon
4. voor het voorbereiden van gerechtelijke uitspraken.

3.4 Sleutelmoment II: vóór de start van het ontwikkelproces

In de voorbereidende en verkennende fase wordt steeds duidelijker welke gegevens gebruikt gaan worden, hoe het algoritme *zelf* wordt vormgegeven, wat de beoogde resultaten zijn en welke impact het algoritme kan gaan hebben. Dit sleutelmoment bevindt zich vlak voor het moment dat de eerste regels code worden geschreven. Hierbij krijgen de elementen legitimering en beheersplan als volgt een plaats:

- *Legitimering.* Bij dit sleutelmoment ligt het zwaartepunt meestal niet bij democratische of maatschappelijke legitimering. Wel worden in deze fase relevante juridische vraagstukken beter inzichtelijk en ligt het voor de hand juridische kennis en expertise in het ontwikkelproces in te brengen. Daarbij hoort ook een check op de AI-verordening, de Avg en de Awb. Ook de (impliciete) eisen die vanuit beleidsterrein specifieke wetgeving[4] worden gesteld worden bestuurd. De resultaten van deze juridische checks worden onderdeel van het beheersplan.

4 Te denken valt bijvoorbeeld aan de Participatiewet of de Omgevingswet, afhankelijk van het toepassingsgebied van het algoritme.

- *Waardengebaseerd beheersplan.* In sleutelmoment II ligt het zwaartepunt bij het beheersplan. Omdat de algoritmetoepassing steeds meer vorm krijgt, is het mogelijk om een voldragen beheersplan te maken. Het beheersplan bestaat uit drie onderdelen:
  - een risicoanalyse op basis van ethische waarden. Om tot een risicoanalyse te komen, zijn impact assessments zoals IAMA[5] een nuttig hulpmiddel.
  - een juridische check, in ieder geval op de AI-verordening. Omdat de algoritmetoepassing steeds meer vorm krijgt, moet naast de risicoanalyse ook de lijst met verboden en hoogrisico algoritmetoepassingen uit de AI-verordening worden doorlopen. Het is belangrijk om daarbij de verordening zelf erbij te pakken. De lijst uit de verordening is (nog) aan wijzigingen onderhevig.
  - een beschrijving van beheersmaatregelen die volgen uit de risicoanalyse. Als de algoritmetoepassing na ingebruikname vooral wordt gebruikt door medewerkers zonder specifieke kennis over algoritmetoepassing, is extra alertheid vereist. In het beheersplan moet in dat geval speciaal aandacht zijn voor de kennis over data, model en architectuur, (informatie)beveiliging en wet- en regelgeving die aanvullend nodig is bij deze medewerkers.

In het onderstaande kader staat – ter illustratie – hoe op basis van een risicoanalyse op basis van de CODIO tot een beheersmaatregel kan worden gekomen.

3.5 Sleutelmoment III: ingebruikname

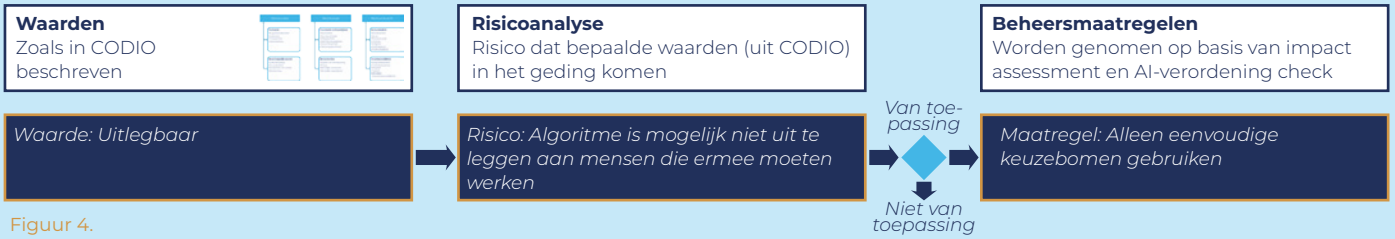
Een derde sleutelmoment is de ingebruikname van de algoritmetoepassing. In deze fase wordt de algoritmetoepassing onderdeel van het instrumentarium van de overheids-organisatie. Dit betekent in veel overheidsorganisaties dat de algoritmetoepassing vooral wordt gebruikt door medewerkers zonder specifieke kennis van algoritmetoepassingen. Ter voorbereiding van de ingebruikname krijgen de elementen legitimering en beheersplan als volgt een plaats:

- *Legitimering.* Bij de ingebruikname ligt het zwaartepunt bij de legitimering. Wie daarbij betrokken wordt (maatschappij, volksvertegenwoordiging, bestuurder, ambtelijke leiding), hangt af van de kaderstelling (sleutelmoment I).
  Het is bijvoorbeeld mogelijk dat is bepaald dat een algoritmetoepassing eerst langs de volksvertegenwoordiging moet, voordat het in gebruik genomen kan worden.
  Belangrijk is verder te bepalen of de kaders die destijds zijn bedacht, nog steeds passend zijn in het licht van de uiteindelijk ontwikkelde algoritmetoepassing. Ook is in deze fase een laatste juridische check op de algoritmetoepassing aan te bevelen.

- *Waardengebaseerd beheersplan.* Het beheersplan speelt een rol in de besluitvorming over de ingebruikname van de algoritmetoepassing. Het is belangrijk dat het beheersplan is opgesteld op een manier dat de beslisser (die de ingebruikname legitimeert; bijvoorbeeld de bestuurlijk of ambtelijk verantwoordelijke) voldoende begrijpt hoe de toepassing werkt en welk risico de overheidsorganisatie neemt bij de ingebruikname ervan.

Een voorbeeld: van risicoanalyse naar beheersmaatregelen

In beheerplan worden op basis van de risicoanalyse (onderdeel 1) de bijpassende beheersmaatregelen bepaald. Hieronder staat een voorbeeld aan de hand van de waarde 'uitlegbaarheid' uit de CODIO:



Figuur 4.

5 https://www.rijksoverheid.nl/documenten/rapporten/2021/02/25/impact-assessment-mensenrechten-en-algoritmes

3.6 Sleutelmoment IV: wezenlijke wijziging

Een vierde sleutelmoment ontstaat bij een wezenlijke wijziging van de algoritmetoepassing of een wezenlijke wijziging van de legitimerende omgeving waarbinnen de algoritmetoepassing wordt ingezet. De wezenlijke wijziging van de algoritmetoepassing kan ontstaat in iedere processtap in de algoritmetoepassing. Ter illustratie een voorbeeld per processtap:

- er worden wezenlijk andere ***gegevens*** gebruikt in de algoritmetoepassing;
- de code van het ***algoritme*** wordt complexer en wordt meer zelflerend;
- de status van de ***resultaten*** van de algoritmetoepassing wijzigt van een nuttige suggestie voor een besluit naar een geautomatiseerd besluit;
- het wordt duidelijk dat de algoritmetoepassing onverwachte neveneffecten heeft die de ***impact*** van de algoritmetoepassing in een ander licht stellen, bijvoorbeeld wanneer een algoritmetoepassing gericht op handhaving van regels, vooral bepaalde groepen binnen de samenleving raakt.

De wezenlijke wijziging van de legitimerende omgeving kan (op hoofdlijnen) gekoppeld worden aan de drie vormen van legitimering (zie paragraaf 2.4). Ter illustratie een voorbeeld per vorm van legitimiteit.

- er komt een nieuwe volksvertegenwoordiging of bestuur van de overheidsorganisatie die in meerderheid tegen de inzet van de algoritmetoepassing is. Dit wijzigt de ***politieke legitimering*** van de algoritmetoepassing en kan ertoe leiden dat de algoritmetoepassing moet worden beëindigd;
- de wettelijke grondslag voor het gebruik van gegevens wordt verruimd, waardoor nieuwe gegevens beschikbaar zijn voor de algoritmetoepassing en de voorwaarden voor de ***juridische legitimiteit*** van de algoritmetoepassing veranderen;
- er ontstaat onrust in de samenleving over de inzet van een algoritmetoepassing, waardoor de ***maatschappelijke legitimiteit*** van de algoritmetoepassing afbrokkelt.

Of deze wijzigingen van de algoritmetoepassing of de legitimerende omgeving wezenlijk zijn, moet per geval worden beoordeeld.

Voor de elementen legitimering en het beheersplan heeft dit het volgende gevolg:

- *Legitimering.* De wezenlijke wijziging kan effect hebben op de democratische legitimering (het valt niet meer binnen de kaders de bij de ingebruikname van de algoritmetoepassing zijn bepaald), de juridische legitimering (de toepassing past niet binnen de daarvoor bedoelde wet- en regelgeving) of de maatschappelijke legitimering (voor dit type toepassing is geen draagvlak in de samenleving).

- *Waardengebaseerd beheerplan.* Het beheersplan moet worden aangepast aan de wezenlijke wijziging. Ingeschat moet worden wat de effecten zijn op de waarden uit de CODIO en de lijst uit de AI-verordening, en of het moet leiden tot nieuwe beheersmaatregelen.

3.7 Sleutelmoment V: periodiek controlemoment

De verwachting is dat de komende jaren de inzet van algoritmetoepassingen binnen de overheid een grote vlucht gaat nemen. Veel experts maken zich daarbij zorgen dat sommige algoritmetoepassingen daarbij 'uit het zicht' verdwijnen, maar tegelijkertijd wel qua eigenschappen veranderen vanwege het emergente karakter van algoritmetoepassingen (zie paragraaf 2.4.3). Een periodiek controlemoment voor dit type veranderlijke algoritmetoepassingen is daarom het vijfde sleutelmoment. Bij een periodiek controlemoment krijgen de elementen legitimering en beheersplan als volgt een plaats:

- *Legitimering.* Tijdens eerdere sleutelmoment is bepaald wie de algoritmetoepassing democratisch legitimeert (volksvertegenwoordiging, bestuurder, leiding ambtelijke organisatie), hoe de juridische legitimering geborgd wordt in het vervolgproces (wanneer en hoe welke juristen betrokken worden) en in hoeverre er aandacht moet zijn voor maatschappelijke legitimering. Tijdens het periodieke controlemomenten moeten deze elementen van legitimiteit worden gecheckt op basis van de dan functionerende algoritmetoepassing. Indien nodig wordt actie ondernomen om de legitimiteit van de toepassing blijvend te garanderen.

- *Waardengebaseerd beheersplan.* Belangrijk is om eerst te controleren in welke mate de algoritmetoepassing gewijzigd is. Vervolgens moet, net als bij een wezenlijke wijziging (sleutelmoment IV), ingeschat worden wat de effecten van deze wijziging zijn op de waarden uit de CODIO en de lijst uit de AI-verordening, en of het moet leiden tot nieuwe beheersmaatregelen.

4. Best practices voor de inrichting van de governance

4.1 Inleiding en centrale vraag

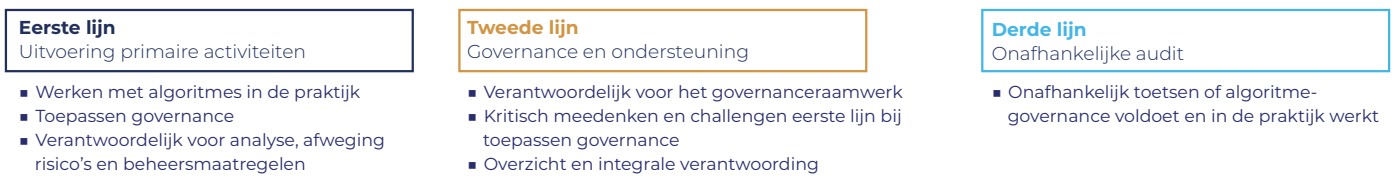
In hoofdstuk 3 stonden de twee hoofdelementen van de governance centraal: vijf sleutelmomenten een gelegitimeerd beheersplan. In aanvulling daarop wordt in dit hoofdstuk van de handreiking een aantal best practices voor het inrichten van de governance uitgelicht. De eerste twee best practices gaan over aansluiten bij bestaande governancestructuren. Dit versterkt de doelmatigheid van de governance. Het gaat om de koppeling van algoritmetoepassingen aan het three lines of defense model en het opstellen van een takenlijst via het RASCI-model. Daarna wordt een aantal best practices opgesomd die tijdens het onderzoek naar voren kwamen. Het is geen uitputtende lijst, maar dient ter inspiratie voor overheidsorganisaties die met de inrichting van de governance aan de gang willen.

4.2 Algoritmetoepassing en het three lines of defense model

Het three lines of defense model wordt in veel overheidsorganisaties gehanteerd. Om tot een doelmatige inrichting van de governance voor de verantwoorde toepassing van algoritmen te komen, valt daarom aansluiting bij dit organisatiemodel sterk aan te bevelen. Figuur 5 bevat een eenvoudige visualisering van zo'n risico gebaseerd model.

De verbinding met het gelegitimeerde beheersplan (zie hoofdstuk 3) is eenvoudig te maken. De verantwoordelijkheid voor het opstellen van het beheersplan ligt bij de eerste lijn. Binnen de tweede lijn wordt gewerkt met een framework (kader) waarmee gecontroleerd wordt of het gelegitimeerde beheersplan voldoet aan de standaarden van de overheidsorganisatie, en of de maatregelen in het beheersplan ook worden uitgevoerd. Het toetsingskader algoritmes van de Algemene Rekenkamer kan hiervoor de handvatten bieden.[6] De derde lijn voert regelmatig (interne of externe) audits uit op het hele proces. Daarbij kan gekozen voor een risico gebaseerde auditplanning.

Figuur 5. **Visualisering three lines of defense model.**



6 https://www.rekenkamer.nl/onderwerpen/algoritmes-digitaal-toetsingskader

4.3 Takenlijst governance in combinatie met RASCI

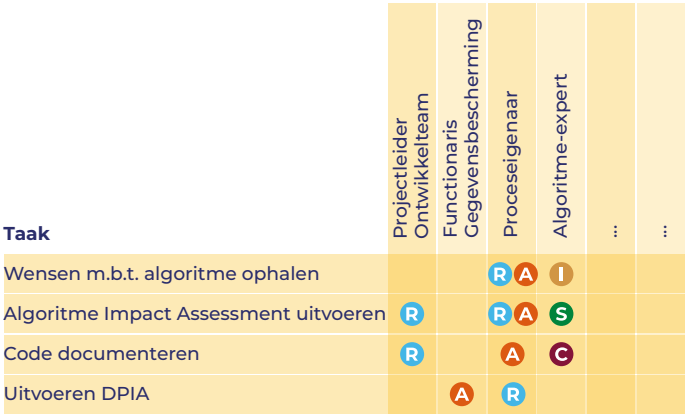
In veel organisaties wordt de governance uitgeschreven in takenlijsten. In deze handreiking hebben we een voorbeeld van een dergelijke takenlijst voor de governance van de verantwoorde toepassing van algoritmen opgenomen. Deze taken zijn erop gericht dat de toepassing van algoritmen politiek, juridisch en maatschappelijk gelegitimeerd zijn. De taken die nodig zijn voor het verrichten van een op waarden gebaseerde risicoanalyse zijn ook in dit takenoverzicht opgenomen.

De taken in het overzicht zijn gecategoriseerd per fase van de levenscyclus en zijn eveneens gekoppeld aan één van de zes principes van de CODIO. Wanneer uit het beheersplan blijkt dat er een vergroot risico is dat één van deze principes in het gedrang komt, kan er worden besloten extra mitigerende maatregelen te nemen die zien op dat principe. De takenindeling is eveneens congruent met de benadering van de vijf controlemomenten.

Om de taken in de takenlijst goed te verdelen, hanteren veel overheidsorganisaties (een variant van) het RASCI-model. Het RASCI-model bestaat uit een matrix die de rollen en verantwoordelijkheden van verschillende personen weergeeft bij een project. RASCI is een afkorting voor *Responsible* (Verantwoordelijk), *Accountable* (hoofdverantwoordelijk), *Support* (Ondersteunend), *Consulted* (Geraadpleegd), *Informed* (geïnformeerd). Deze zes rollen kunnen via een RASCI-model worden verdeeld, zo kan duidelijkheid worden geschept in wat er van iedereen wordt verwacht.

Er kan maar één iemand *accountable* (hoofdverantwoordelijk) zijn. Het ligt voor de hand de bestuurlijke verantwoordelijkheid te laten aansluiten bij de ambtelijke hoofdverantwoordelijkheid. Dit betekent dat de wethouder aan wie de hoofd-verantwoordelijke in bovenstaand schema rapporteert ook degene is die politiek en bestuurlijk verantwoordelijk is voor juiste uitvoering. Om een RASCI in te vullen is een heldere omschrijving van de taken en de rollen noodzakelijk. Per overheidsorganisatie verschillen de namen en inhoud van de rollen significant. Ter illustratie en inspiratie is in figuur 6 een grotendeels leeg RASCI-model weergegeven. De functieverdeling zoals die is weergegeven in de tabel is puur als voorbeeld bedoeld.

Figuur 6. Voorbeeld RASCI-model.



Een voorbeeld van taken die binnen een overheidsorganisatie kunnen worden belegd in het kader van de verantwoorde toepassing van algoritmen zijn:

Tabel 1. **Overzicht van taken die binnen een overheidsorganisatie belegd kunnen worden op het gebied van de verantwoorde toepassing van algoritmen.**

Hoofdfase	Deelfase	Taak	CODIO-principes
Onderzoeksfase	1 - Vraag	Het bepalen en vastleggen van het beoogde doel	Bestuurskwaliteit
Onderzoeksfase	1 - Vraag	Bepalen en vastleggen of een algoritme een juiste oplossing kan zijn	Bestuurskwaliteit
Onderzoeksfase	2 - Intentie	Quickscan Algoritme	Bestuurskwaliteit
Onderzoeksfase	2 - Intentie	Uitvoeren Data Protection Impact Assesment	Mensenrechten
Onderzoeksfase	2 - Intentie	Algoritme voorleggen aan politiek bestuurder	Verantwoordelijkheid
Onderzoeksfase	2 - Intentie	Inrichten Project Portfolio Management-Proces (of iets soortgelijks)	Bestuurskwaliteit
Onderzoeksfase	2 - Intentie	Make-or-buy besluit	Bestuurskwaliteit
Onderzoeksfase	2 - Intentie	Vaststellen opdrachtomschrijving	Bestuurskwaliteit
Onderzoeksfase	2 - Intentie	Bij inkoop: borgen onafhankelijkheid	Bestuurskwaliteit
Onderzoeksfase	2 - Intentie	Haalbaarheid aanpassingen onderzoeken	Bestuurskwaliteit
Onderzoeksfase	2 - Intentie	Vaststellen 'definition of done' (acceptatiecriteria)	Bestuurskwaliteit
Onderzoeksfase	2 - Intentie	Wensen ophalen bij eindgebruikers en stakeholders	Participatie
Onderzoeksfase	2 - Intentie	Toewijzen producteigenaar	Verantwoordelijkheid
Onderzoeksfase	2 - Intentie	Bij inkoop: borgen beheersing risico's	Verantwoordelijkheid
Onderzoeksfase	2 - Intentie	Afspraken maken met data-eigenaar over initiële levering	Verantwoordelijkheid
Onderzoeksfase	2 - Intentie	Vaststellen juridisch kader	Verantwoordelijkheid
Onderzoeksfase	2 - Intentie	Bepalen kwaliteit databronnen	Verantwoordelijkheid
Onderzoeksfase	2 - Intentie	Vastleggen documentatie in intern algoritmeregister	Verantwoordelijkheid

Hoofdfase	Deelfase	Taak	CODIO-principes
Onderzoeksfase	2 - Intentie	GO/NO GO Ontwikkeling algoritme	Bestuurskwaliteit
Ontwikkelfase	3 - Proces	Het inrichten van het team	Bestuurskwaliteit
Ontwikkelfase	3 - Proces	In kaart brengen wie kan worden geconsulteerd	Bestuurskwaliteit
Ontwikkelfase	3 - Proces	Opdracht verlenen	Bestuurskwaliteit
Ontwikkelfase	3 - Proces	Vastleggen opdrachtnemer/opdrachtgever	Verantwoordelijkheid
Ontwikkelfase	3 - Proces	Ophalen requirements	Bestuurskwaliteit
Ontwikkelfase	3 - Proces	Opstellen product requirement document	Bestuurskwaliteit
Ontwikkelfase	3 - Proces	Ophalen buisiness rules	Bestuurskwaliteit
Ontwikkelfase	3 - Proces	Vastleggen gemaakte keuzes	Verantwoordelijkheid
Ontwikkelfase	3 - Proces	In kaart brengen welke maatregelen nodig zijn voor privacy by design, security by design, archivering by design, ethics by design	Bestuurskwaliteit
Ontwikkelfase	4 - Ontwikkeling	Borgen maatregelen privacy by design, security by design, archivering by design, ethics by design	Maatschappelijke waarde
Ontwikkelfase	4 - Ontwikkeling	(Periodiek) Consulteren experts	Bestuurskwaliteit
Ontwikkelfase	4 - Ontwikkeling	Autorisatiematrix opstellen	Maatschappelijke waarde
Ontwikkelfase	4 - Ontwikkeling	Algoritme Impact Assesment uitvoeren	Bestuurskwaliteit
Ontwikkelfase	4 - Ontwikkeling	Opstellen risicoprofiel op basis van algoritme impact assesment	Bestuurskwaliteit
Ontwikkelfase	4 - Ontwikkeling	Toetsen aan acceptatiecriteria	Bestuurskwaliteit
Ontwikkelfase	4 - Ontwikkeling	Uitvoeren Data Protection Impact Assesment	Mensenrechten
Ontwikkelfase	4 - Ontwikkeling	Scheiden van data	Mensenrechten
Ontwikkelfase	4 - Ontwikkeling	(Periodiek) Consulteren eindgebruikers en stakeholders	Participatie
Ontwikkelfase	4 - Ontwikkeling	Uitvoeren bias analyse	Procedurele rechtvaardigheid
Ontwikkelfase	4 - Ontwikkeling	Uitvoeren analyse overfitting/underfitting	Maatschappelijke waarde
Ontwikkelfase	4 - Ontwikkeling	Afspraken maken met data-eigenaar over structurele levering	Verantwoordelijkheid
Ontwikkelfase	4 - Ontwikkeling	Controleren geleverde dataset	Maatschappelijke waarde
Ontwikkelfase	4 - Ontwikkeling	Anonimiseren/pseudonimiseren test en trainingsdata	Mensenrechten
Ontwikkelfase	4 - Ontwikkeling - Pilot	Pilotplan opstellen	Bestuurskwaliteit
Ontwikkelfase	4 - Ontwikkeling - Pilot	Go/no go pilot algoritme	Bestuurskwaliteit
Ontwikkelfase	4 - Ontwikkeling - Pilot	Uitvoeren pilot	Bestuurskwaliteit
Ontwikkelfase	4 - Ontwikkeling - Pilot	Evaluatie Pilot	Bestuurskwaliteit
Ontwikkelfase	4 - Ontwikkeling - Pilot	Inwinnen adviezen eindgebruikers en stakeholders naar aanleiding van pilot	Participatie
Ontwikkelfase	4 - Ontwikkeling	Go/No Go ingebruikname algoritme	Bestuurskwaliteit
Ontwikkelfase	4 - Ontwikkeling	Voorleggen go/no go algoritme aan politiek bestuurder	Verantwoordelijkheid
Beheersfase	5 - in gebruik	Persbericht versturen	Participatie
Beheersfase	5 - in gebruik	Publiekscommunicatie	Procedurele rechtvaardigheid
Beheersfase	5 - in gebruik	Inschrijven in extern algoritmeregister	Verantwoordelijkheid
Beheersfase	5 - in gebruik	Werkinstructies opstellen	Verantwoordelijkheid
Beheersfase	5 - in gebruik	Training voor eindgebruikers voorbereiden en geven	Verantwoordelijkheid
Beheersfase	5 - in gebruik	Communiceren waar werknemers melding kunnen maken van ongewenste impact	Verantwoordelijkheid
Beheersfase	5 - in gebruik	Jaarlijks evaluatiegesprek met eindgebruikers	Verantwoordelijkheid
Beheersfase	5 - in gebruik	Jaarlijkse evaluatie opstellen algoritme	Verantwoordelijkheid
Beheersfase	5 - in gebruik	Vastleggen wanneer en waarom van de uitkomst van het algoritme is afgeweken	Bestuurskwaliteit
Beheersfase	5 - in gebruik	continue monitoring algoritme	Verantwoordelijkheid
Beheersfase	5 - in gebruik	Periodieke bias analyse	Procedurele rechtvaardigheid
Beheersfase	5 - in gebruik	Bijwerken risicoprofiel	
Beheersfase	5 - in gebruik	Bij inkoop: Periodieke controle contractvoorwaarden, servicelevel en rapportages van de leverancier	Verantwoordelijkheid
Beheersfase	6 - Einde	Stoppen algoritme	Maatschappelijke waarde
Beheersfase	6 - Einde	Archiveren algoritme	Bestuurskwaliteit

4.4 Een overzicht van best practices

In deze paragraaf volgt een aantal best practices die tijdens het onderzoek naar voren kwamen. Het is geen uitputtende lijst, maar wordt ten zeerste aanbevolen voor overheidsorganisaties die met de inrichting van de governance aan de gang willen.

Beleg stelselverantwoordelijkheid voor de werking van de governance

Om op lange termijn de verantwoorde toepassing van algoritmen te garanderen, is van belang niet alleen gericht te zijn op de algoritmetoepassingen zelf, maar om ook de verantwoordelijkheid voor de werking van het stelsel te beleggen. Deze stelselverantwoordelijke ziet toe op de werking van de governance. Dit betekent onder meer dat de stelsel-verantwoordelijke regelmatig de werking van de governance evalueert en, indien nodig, de kaders voor de toepassing van algoritmen aanpast. Verder dragen stelselverantwoordelijken vaak zorg voor meer bewustwording over de mogelijkheden en risico's van algoritmetoepassingen.

Overweeg een bias analyse in aanvulling op de risicoanalyse

In aanvulling op de risicoanalyse in het beheersplan, kan overwogen worden om een bias analyse te doen. Deze analyse is vooral passend wanneer blijkt dat waarden als inclusiviteit, of procedurele rechtvaardigheid (vooral non-discriminatie) onder druk kunnen komen te staan door de algoritmetoepassing. De bias analyse (onbewuste vooringenomenheid) is een extra maatregel om te controleren of de onderliggende dataset en de ontwikkelde algoritmen voldoen aan de vereisten.

Het uitvoeren van een onderzoek naar onbewuste vooringenomenheid (bias) is een extra maatregel om te controleren of de onderliggende dataset en de ontwikkelde algoritmen voldoen aan bovenstaande vereisten. Het onderzoek gaat in het bijzonder in op de ethische aspecten en ondersteunt om te kunnen beoordelen of bepaalde groepen ongerechtvaardigd worden benadeeld binnen het werkproces (op basis van onderliggende data) en het ontwikkelde algoritme. Een methode om ethische kwesties bij dataprojecten te bespreken en documenteren is De Ethische Data Assistent (DEDA)[7]. Een methode om discriminatie en bias op te sporen in het model zelf (in geval van machinelearningmodellen) is de open source toolkit AI 360 Fairness[8].

Tot slot kan de handreiking met systeemprincipes voor non-discriminatie behulpzaam zijn.[9] Het helpt bij het ontwikkelen van algoritmische systemen om non-discriminatiewetgeving te doorgronden en vervolgens bias en discriminatie in de ontwikkeling van algoritmen vroegtijdig te onderkennen.

Aandacht voor rechtsbescherming in de bezwaarprocedure

Uit de risicoanalyse in het beheersplan kan blijken dat de algoritmetoepassing risico's kan opleveren voor waarden als uitlegbaarheid, aanvechtbaarheid, aanspreekbaarheid en controleerbaarheid. Dit zijn waarden die nauw samenhangen met de mogelijkheid van burgers om op te komen tegen overheidsbesluiten. Wanneer er een kans is dat op die manier de rechtsbescherming onder druk komt te staan, kan het aanpassen van bezwaarprocedures een goede oplossing zijn.

In het programma publieke controle op algoritmes is de 'Handreiking effectieve en efficiënte rechtsbescherming tegen het gebruik van algoritmen door de overheid' opgesteld. In deze handreiking staan aanwijzingen voor een toegankelijke, structurele en effectieve bezwaarprocedure. Het doel is dat het voor burgers duidelijker is hoe zij tegen een besluit dat – mede tot stand is gekomen door inzet van een algoritmetoepassing – bezwaar kunnen maken, waar zij informatie over de algoritme-toepassing kunnen vinden én hoe een zorgvuldige herover-weging plaatsvindt. Dit draagt bij aan de legitimering van de algoritmetoepassing.

Overweeg de aanstelling van een algoritme expert

De doorvertaling van ethische en juridische principes naar algoritmetoepassingen is complex. Dit vereist enerzijds het goed kennen en begrijpen van ethische principes en anderzijds diepgaande kennis van hoe algoritmen werken en waar risico's zitten. In een aantal overheidsorganisaties is daarom gekozen een algoritme expert aan te wijzen. Deze expert ondersteunt proceseigenaren en ontwikkelaars bij de verantwoorde toepassing van algoritmen en bevordert de bewustwording binnen de organisatie.

7 www.dataschool.nl/deda
8 www.aif360.mybluemix.net
9 https://www.rijksoverheid.nl/documenten/rapporten/2021/06/10/handreiking-non-discriminatie-by-design

Overweeg de instelling van een externe adviescommissie

Het is mogelijk om een externe adviescommissie in te stellen. Deze externe commissie kan op verschillende manieren de verantwoorde toepassing van algoritmen in een overheidsorganisatie bevorderen. In de onderstaande tabel, die is gebaseerd op de handreiking digitale ethiek van de VNG (2022), worden een aantal mogelijkheden geschetst aan de hand van drie dimensies.

Tabel 2.

Dimensie	Mogelijke invulling
Doel van de externe commissie	• Advisering over (ondersteuning bij) (het afwegen van) de verantwoorde toepassing van algoritmen. • Bijdrage aan de deskundigheidsbevordering over algoritmetoepassingen binnen de organisatie. • Signaleren wat er binnen de organisatie speelt op het gebied van algoritmetoepassingen en onder de aandacht brengen van het bestuur en het management van de organisatie. • Maatschappelijk klankbord.
Samenstelling van de externe commissie	• Op basis van professionele deskundigheid op het gebied van grondrechten, ICT-recht, technologie-ethiek, techniek (inclusief kunstmatige intelligentie), maatschappij en inwoners. • Op basis van een goede balans tussen mannen en vrouwen, jong en oud, overheid en bedrijfsleven, en achtergrond (wetenschap, medisch, onderwijs, journalistiek, juridisch, technisch, etc.).
Aan wie moet de externe commissie zich verantwoorden?	• Verantwoording aan ambtelijke leiding, bestuurlijke leiding of volksvertegenwoordiging • Zijn de vergaderingen van de commissie openbaar?

Richt een algoritmeregister in

In een algoritmeregister worden in gebruik zijnde algoritmetoepassingen vermeldt (ook pilots). Eventueel kunnen ook algoritmetoepassingen die ontwikkeld worden maar nog niet worden gebruikt, worden vermeld in het register. In elk geval is raadzaam om – in lijn met beheersplan – in het algoritmeregister op te nemen:

- een beschrijving van het doel van de algoritmetoepassing,
- de verantwoordelijke voor de algoritmetoepassing,
- de geïdentificeerde risico's,
- de check met de lijst van verboden en hoogrisico algoritmetoepassingen uit de AI-verordening,
- en de genomen beheersmaatregelen.

Leg de relatie met bestaande governance voor privacy en informatiebeveiliging

In alle overheidsorganisaties zijn structuren ingericht voor het naleven van de Avg en de regels rond informatiebeveiliging (bijvoorbeeld uit de Baseline Informatiebeveiliging Overheid (BIO)). Deze governance voor privacy en informatiebeveiliging kent veel raakvlakken met de governance voor de verantwoorde toepassing van algoritmen. Een voorbeeld is dat bij de uitvoering van een data protection impact assessment (DPIA) veel informatie wordt verzameld die ook relevant is voor de verantwoorde toepassing van algoritmen. Om dubbel werk te voorkomen, is het aan te bevelen om de governance voor de verantwoorde toepassing van algoritmen slim te koppelen aan deze verwante structuren.

Bijlagen

B1 Gesprekspartners

Tabel 3. **Geïnterviewde Informatie-adviseurs, ontwikkelaars, data scientists en gebruikers van algoritmen.**

Organisatie	Gesprekspartner
Gemeente Amsterdam	Bart de Visser Karen Schoonderwaldt Loek Foster
Gemeente Rotterdam	Yugesh Raghoenath Ivo van Rij René van der Toorn Vrijthoff David Brak Joop Polfliet
Kadaster	Erwin Folmer Tijn van Bussel
Ministerie van Binnenlandse Zaken	Marcel Hopman Martin Borhani
Provincie Noord-Brabant	Simone Daniels Paul Rakké Marcel Thaens
Rijkswaterstaat	John Steenbruggen

Tabel 4. **Geïnterviewde wetenschappers.**

Organisatie	Gesprekspartner
Tilburg University	Merel Noorman
TU Delft	Sem Nouws Roel Dobbe
Universiteit Utrecht	Iris Muis
VU Amsterdam	Sandjai Bhulai

B2 Bronnenlijst

- *N. Bonetje & N. Goedhart.* Pels Rijcken. Juridisch kader voor algoritmische toepassingen. 2021.
- *E. Dingemas et al.* Het PON & Telos. Informatiebehoeften van burgers over de inzet van algoritmes door overheden. 2021.
- *Europese Commissie.* Voorstel AI verordening. 2021.
- *Gemeente Amsterdam.* Onwikkelproces van de werkplaats. z.d.
- *Gemeente Amsterdam.* Template Kunstmatige Intelligentie Impact Assessment (inclusief BIAS-analyse). 2021.
- *Gemeente Amsterdam.* Handreiking effectieve en efficiënte rechtsbescherming tegen het gebruik van algoritmen door de overheid. z.d.
- *Gemeente Amsterdam.* Taken en verantwoordelijkheden bij het gebruik van Algoritmen door de Gemeente Amsterdam (conceptversie 0.12). z.d.
- *Gemeente Amsterdam.* Powerpoint: Taken en verantwoordelijkheden bij het gebruik van Algoritmen door de Gemeente Amsterdam. 2021.
- *Gemeente Rotterdam.* Governance voor de verantwoorde toepassing van hoog-risico algoritmen in de Gemeente Rotterdam. 2021.
- *Gemeente Rotterdam.* Kadernota Sturen en Verantwoorden 2020. z.d.
- *Gemeente Rotterdam.* Kadernota Sturen en Verantwoorden 2020. 2020.
- *J. Gerards et al.* Ministerie van Binnenlandse Zaken. Impact Assessment Mensenrechten en Algoritmes. 2021.
- *S. Kulk & S. van Deursen.* Montaigne Centrum voor Rechtsstaat en Rechtspleging. Juridische aspecten van algoritmen die besluiten nemen. Een verkennend onderzoek. 2020.
- *V. Peerenboom & S. van Weerdenburg.* Rijkswaterstaat. Handleiding voor Data Science Projecten. 2020.
- *Raad van State.* Digitalisering: Wetgeving en bestuursrechtsspraak. 2021.
- *Rekenkamer Rotterdam.* Gekleurde technologie. 2021.
- *B. van der Sloot et al.* Ministerie van Binnenlandse Zaken. Handreiking non-discriminatie by design. 2021.
- *A. Vankan et al.* Utrecht Data School. Aanbevelingsnotitie ethisch gedeelte algoritmekader. 2021.