2021

Aandacht voor algoritmes

Algemene Rekenkamer

Vooraf

Wij zijn begin 2020 met dit onderzoek gestart. In Nederland werden toen net de corona-maatregelen afgekondigd en het kabinet moest al zijn aandacht richten op het beheersen van de crisis.
Onze onderzoeksperiode viel precies samen met deze eerste ‘golf van besmettingen met het nieuwe coronavirus’. Tijdens de ambtelijke afstemming van ons onderzoek was er sprake van een tweede golf van besmettingen met het nieuwe coronavirus. Ondanks de impact die het coronavirus had (en heeft) op het dagelijks leven, hebben de medewerkers bij de ministeries en hun organisatieonderdelen alle informatie opgeleverd waar wij om vroegen. Daarnaast hebben ze tijd gemaakt voor interviews om onze vragen te beantwoorden. Mede dankzij hun inspanningen konden wij, onder deze uitzonderlijke omstandigheden, ons onderzoek voortzetten.

De tekst in dit document is vastgesteld op 14 januari 2021.
Dit document is op 26 januari 2021 aangeboden aan de Tweede Kamer.

Inhoud

1. Samenvatting | 5
1.1 Conclusies | 6
1.2. Aanbevelingen | 8

2. Over dit onderzoek | 9
2.1 Waarom dit onderzoek? | 9
2.2 Wat hebben we onderzocht en hoe? | 12
2.3 Leeswijzer | 14

3. Inzicht in algoritmes | 15
3.1 Totaalbeeld algoritmes | 15
3.2 Voor welke activiteiten en processen worden algoritmes toegepast bij de rijksoverheid en bij organisaties die aan de overheid zijn verbonden, welke typen/categorieën zijn er te onderscheiden en wat zijn de effecten en risico's? | 17
3.3 Hoe is de besturing en kwaliteitsbeheersing van algoritmes vormgegeven? | 20

4. Toetsingskader algoritmes | 22
4.1 5 perspectieven | 23
4.2 Denksessie: begrippen en definities | 25

5. Praktijktoets: 3 algoritmes | 26
5.1 Selectie algoritmes | 26
5.2 Belangrijkste inhoudelijke observaties | 30

6. Conclusies en aanbevelingen | 35

6.1 Een algoritme hoeft geen black box te zijn | 36

6.2 Centraal inzicht ontbreekt, behoefte aan concrete instrumenten | 37

6.3 Voorspellende en voorschrijvende algoritmes volop in ontwikkeling met nu nog beperkte impact burger | 38

6.4 De burger staat onvoldoende centraal | 39

6.5 Verbeterpunten voor een verantwoorde inzet en doorontwikkeling van algoritmes | 39

7. Reactie en nawoord | 43

7.1 Reactie staatssecretaris van BZK | 43

7.2 Nawoord Algemene Rekenkamer | 46

Bijlagen | 49

Bijlage 1 Methodologische verantwoording | 49

Bijlage 2 Literatuurlijst en bronnen toetsingskader | 53

Bijlage 3 Toetsingskader algoritmes | 55

Bijlage 4 Eindnoten | 63

1.
Samenvatting

Bij de uitvoering van het beleid maakt de rijksoverheid gebruik van algoritmes. Algoritmes zijn sets van regels en instructies die een computer geautomatiseerd volgt bij het maken van berekeningen om een probleem op te lossen of een vraag te beantwoorden.[1]
We wilden onderzoeken wat die algoritmes nu precies wel en niet doen, demystificeren dus. We wilden vragen beantwoorden als: hoe voorkomt de rijksoverheid dat er vooroordelen in algoritmes sluipen? Overziet de rijksoverheid wat de inzet van algoritmes voor gevolgen heeft voor personen en bedrijven die met het overheidsbeleid te maken krijgen?

Neem bijvoorbeeld SyRI (Systeem Risico Indicatie), een systeem dat binnen de overheid (zoals het Uitvoeringsinstituut Werknemersverzekeringen (UWV) en de Belastingdienst) werd gebruikt om fraude op te sporen met algoritmes. In februari 2020 oordeelde de rechter dat de wetgeving die de inzet van SyRI regelt een te grote inbreuk op de privacy van burgers vormde.[2]

Ook in de Tweede Kamer worden regelmatig zorgen geuit door Kamerleden over zaken als discriminatie en vooringenomenheid die bij de inzet van algoritmes op de loer zouden liggen. En medio 2020 kwam daar de maatschappelijke discussie over de corona-app bij. Naast zorgen over de herkomst, de verzameling en het gebruik van data ging deze discussie ook over een transparante en controleerbare werking van de gebruikte algoritmes.

Algoritmes maken een steeds groter onderdeel uit van het functioneren en handelen van de rijksoverheid en zijn daarmee onderdeel van de dienstverlening naar burgers en bedrijven. Wij inventariseerden voor welke activiteiten en processen algoritmes worden toegepast bij de rijksoverheid en de daaraan verbonden organisaties, welke categorieën er te onderscheiden zijn en wat de risico's van het gebruik van algoritmes zijn. Daarnaast onderzochten we hoe de besturing en kwaliteits-beheersing van algoritmes bij de rijksoverheid en de daaraan verbonden organisaties is ingericht.

1.1 Conclusies

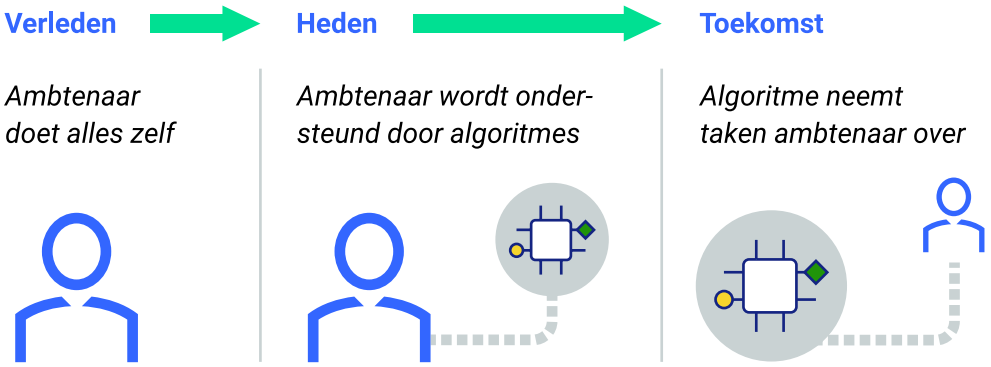
Wij hebben geconstateerd dat binnen de rijksoverheid met name relatief eenvoudige algoritmes worden ingezet. De effecten van eenvoudige algoritmes op de burgers zijn beperkt omdat deze relatief eenvoudige algoritmes automatische besluiten nemen. Het gaat dan vaak om het automatiseren van een administratieve handeling, bijvoorbeeld het automatisch versturen van brieven zoals een ontvangstbevestiging. We hebben geen volledig zelflerende algoritmes aangetroffen binnen de rijksoverheid; alleen lerende algoritmes. Er is altijd een mens betrokken bij het leren door het algoritme. Anders gezegd: er is sprake van "*human in the loop*".

Uit ons onderzoek blijkt dat algoritmes voor ons als onafhankelijk controleur geen *black box* zijn: wij hebben de geïnventariseerde algoritmes kunnen bekijken en beoordelen. Ook hebben wij op basis van de geïnventariseerde voorspellende en voorschrijvende[3] algoritmes geconstateerd dat er bij de ontwikkeling en het gebruik van algoritmes veel aandacht wordt besteed aan het beperken van de privacy risico's. Wij hebben ook vastgesteld dat de geïnventariseerde algoritmes niet zelf besluiten nemen, maar de uitvoerende functionarissen nadrukkelijk betrokken zijn bij het gebruik van deze algoritmes. Deze algoritmes ondersteunen de uitvoerende functionarissen bij het maken van analyses en bij het nemen van besluiten.

Dat neemt niet weg dat er – anno 2021 – ruimte is voor verbetering, omdat het gebruik van algoritmes de komende jaren alleen maar zal toenemen. Wanneer algoritmes zelflerend[4] en daardoor complexer worden kan aan snelheid, kwaliteit en objectiviteit van besluitvorming worden gewonnen. De uitvoerend functionaris komt daarmee ook op meer afstand te staan van besluiten die door de rijksoverheid over burgers en bedrijven worden genomen. Er moeten dan meer eisen worden gesteld aan de kwaliteit van de algoritmes. Daarom is het belangrijk dat het kabinet – in de

eerste plaats de minister van Binnenlandse Zaken – nu aan de slag gaat met de aandachts- en verbeterpunten die wij in dit rapport toelichten. Daarnaast wijzen wij erop dat cybersecurity- en informatiebeveiligingseisen een belangrijke randvoorwaarde zijn om algoritmes verantwoord in te zetten. Uitdagingen hierbij zijn het voorkomen en signaleren van cyberaanvallen, zoals digitale sabotage, spionage en criminaliteit.[5]

Ambtenaren worden steeds meer ondersteund door algoritmes



Ondanks de grote maatschappelijke aandacht voor algoritmes zijn er tot op heden nog geen concrete instrumenten om algoritmes te toetsen of te analyseren. Daarom hebben wij een toetsingskader ontwikkeld. In ons toetsingskader geven wij de bestaande normen aan die van toepassing zijn op algoritmes zodat mogelijke risico's worden beperkt. Wij koppelen de te toetsen aspecten en onderzoeksvragen aan die risico's. Hoe groot de kans is dat risico's zich voor een specifiek algoritme voordoen en wat de schade kan zijn, hangt af van de mate waarin geavanceerde technieken worden toegepast, herkomst, manier van verzamelen en de kwaliteit van de data en de impact van het algoritme op de burger. Ons toetsingskader is bedoeld om te helpen algoritmes transparanter te maken en potentiële risico's over algoritmes bespreekbaar te maken. Met dit toetsingskader kunnen controleurs en auditors in de toekomst algoritmes eenduidig en uniform beoordelen.

1.2 Aanbevelingen

Om centraal inzicht te verkrijgen in de mate waarin en op welke manier algoritmes worden gebruikt en om concrete handvatten te bieden bevelen wij het kabinet het volgende aan:

- Draag zorg voor eenduidigheid en uniformiteit van begrippen en kwaliteitseisen voor algoritmes.

Om algoritmes op een verantwoorde wijze in te zetten en door te ontwikkelen bevelen wij het kabinet het volgende aan:

- Draag zorg voor een vertaling van het toetsingskader naar hanteerbare kwaliteitseisen van algoritmes;
- Zorg ervoor dat alle relevante disciplines worden betrokken bij de ontwikkeling van algoritmes;
- Draag zorg voor het inzicht hebben en houden in het functioneren van de *IT General Controls*;[6]
- Leg afspraken omtrent de inzet van algoritmes vast en richt de continue monitoring op het nakomen van deze afspraken goed in.

Verder constateren wij dat de burger niet centraal staat bij de inzet van algoritmes. Daarom bevelen wij het kabinet aan:

- Geef de burgers inzicht in het algoritme en geef aan waar zij terechtkunnen als ze vragen hebben over algoritmes.

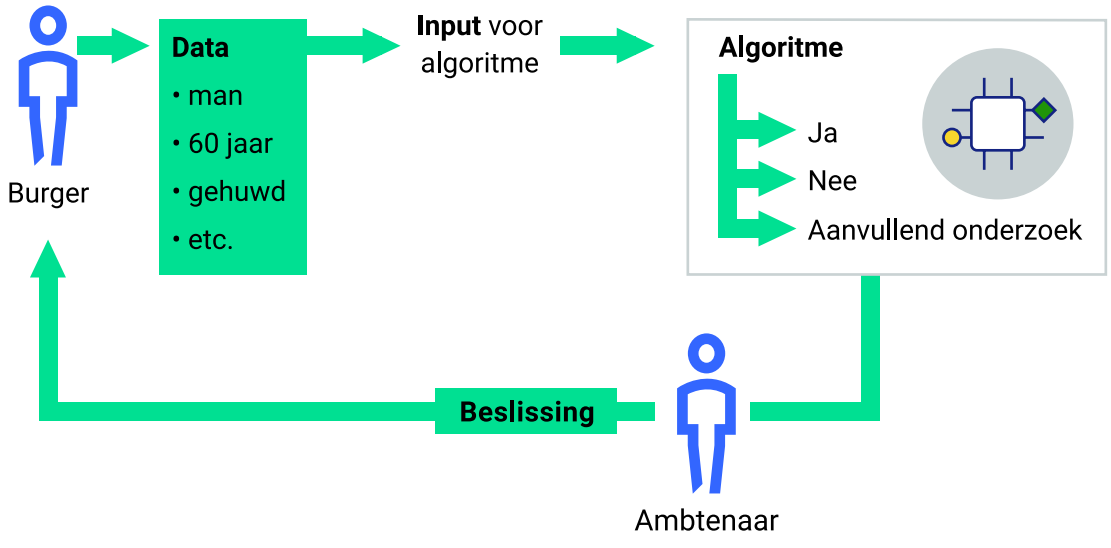
2.
Over dit onderzoek

2.1 Waarom dit onderzoek?

De rijksoverheid functioneert al tientallen jaren met behulp van algoritmes. Met een algoritme bedoelen we: een set van regels en instructies die een computer geautomatiseerd volgt bij het maken van berekeningen om een probleem op te lossen of een vraag te beantwoorden.[7] Algoritmes kennen zeer uiteenlopende verschijningsvormen van rekenmodellen, beslisbomen en andere statistische analyses tot complexe dataverwerkingsmodellen en 'zelflerende' toepassingen.

Een algoritme is een set van regels en instructies die een computer uitvoert om een probleem op te lossen of een vraag te beantwoorden

Voorbeeld: Iemand vraagt een uitkering aan. Heeft hij daar recht op?



Het gebruik van algoritmes neemt zeer snel toe door de steeds verdergaande automatisering en digitalisering. Sociale media, navigatiesystemen en applicaties als buienradar: ze werken allemaal met algoritmes. Vragen over algoritmes – wat kunnen ze betekenen voor de samenleving en welke risico's brengt het gebruik van algoritmes met zich mee – kunnen rekenen op (soms extreem) positieve én negatieve reacties.

Soms ontstaat de indruk alsof algoritmes zich steeds intelligenter gedragen. Dat komt omdat steeds meer data en betere hardware beschikbaar komen, waardoor algoritmes steeds sneller meer data kunnen verwerken, dus innovatiever worden en diverser van aard worden. Ook kunnen ze in meer toepassingen gebruikt worden (zoals bij robotisering) en bezitten in de meest geavanceerde vorm "het vermogen [..] om externe gegevens correct te interpreteren, om te leren van deze gegevens en om deze lessen te gebruiken om specifieke doelen en taken te verwezenlijken via flexibele aanpassing".[8] Dit wordt vaak beschreven als kunstmatige of artificiële intelligentie (AI). Thema's als AI en algoritmes staan enorm in de belangstelling bij zowel burgers als de rijksoverheid en de verwachtingen erover zijn hooggespannen.

Met dit onderzoek willen we een feitelijke bijdrage leveren aan het gesprek over kansen en risico´s voor algoritmes en AI binnen de rijksoverheid. Het toetsingskader dat wij hebben ontwikkeld, kan een basis vormen om algoritmes verantwoord in te zetten én kan het uitgangspunt zijn voor discussies over de controle van en het toezicht op algoritmes.

2.1.1 Kansen voor algoritmes

AI is volgens het kabinet een sleuteltechnologie. Het kader hiervoor is het Strategisch Actieplan voor Artificiële Intelligentie, dat het kabinet op 8 oktober 2019 aan de Tweede Kamer heeft aangeboden.[9] Een publieke-private alliantie op dit vlak, de Nederlandse AI-coalitie, zal in 2021 een investering van € 23,5 miljoen ontvangen om onderzoek te doen naar kunstmatige intelligentie en het ontwikkelen van toepassingen.

Uit de reacties van de ministeries op onze onderzoeksvragen komt naar voren dat er ook binnen de rijksoverheid brede overeenstemming is over het feit dat AI veel en nieuwe kansen biedt. Vrijwel elk departement ontwikkelt toepassingen of zet deze al in. Soms zijn dit zeer innovatieve algoritmes waarbij kunstmatige intelligentie gebruikt wordt. Algoritmes ondersteunen en verbeteren vaak de bedrijfsvoerings- en dienstverleningsprocessen van organisaties. Ze zorgen er bijvoorbeeld voor dat organisaties gericht mensen en middelen kunnen inzetten bij controles of inspecties.

Daarnaast bieden algoritmes ook kansen om besluitvormingsprocessen juist transparanter en gemakkelijker controleerbaar te maken. De techniek achter een algoritme, de data waar een algoritme gebruik van maakt en de omgang met die data, liggen namelijk vast in instructies; instructies die nogal eens ontbreken bij menselijke besluitvormingsprocessen.

2.1.2 Bedreigingen door algoritmes

Toepassing van algoritmes door overheidsorganisaties levert ook een aantal bedreigingen op. We noemen er 4:

1. Allereerst kan de werking van het algoritme bij de rijksoverheid en de invloed daarvan op overheidshandelen niet begrijpelijk genoeg zijn of niet goed genoeg worden uitgelegd aan burgers. Dit kan zowel te maken hebben met de gebruikte technologie (zoals neurale netwerken), als met de complexiteit (teveel variabelen of componenten).
2. Daarnaast bestaat het risico dat het algoritme of de dataverzameling die het algoritme gebruikt, vooroordelen bevat die tot discriminatie kunnen leiden. De mens heeft ook vooroordelen, maar bij de inzet van algoritmes bestaat het risico dat het algoritme vooral afhankelijk is van de afwegingen (bijvoorbeeld de te hanteren data) van de programmeur of datascientist. De programmeur of datascientist kan specifieke kennis en ervaring met de context ontberen, bijvoorbeeld inhoudelijke kennis over een subsidieproces. Deze kennis is essentieel om een goede afweging te kunnen maken.
3. Een derde bedreiging bij algoritmes die leren van data is dat vaak niet van tevoren bekend of voorspelbaar is wat het algoritme leert, in hoeverre er ongewenste leereffecten zijn. Bepaalde correlaties in de gebruikte data kunnen bijvoorbeeld een algoritme opleveren dat discrimineert.
4. Ten vierde zijn veel algoritmes bij de rijksoverheid afkomstig van externe leveranciers. Dit geldt ook voor (ICT-)systemen waar algoritmes deel van uitmaken. De exacte data en de werking van het algoritme horen vaak tot het eigendom van de betreffende externe leverancier. Soms schermt de leverancier die informatie af. Als het gaat om aansprakelijkheid van het algoritme of om aspecten zoals de verwerking van persoonsgegevens, kan of wil de overheid niet zomaar vertrouwen op de toelichting van deze leverancier. Dit maakt risicoanalyse en -beheersing van het algoritme door de rijksoverheid moeilijker.

2.1.3 Demystificatie

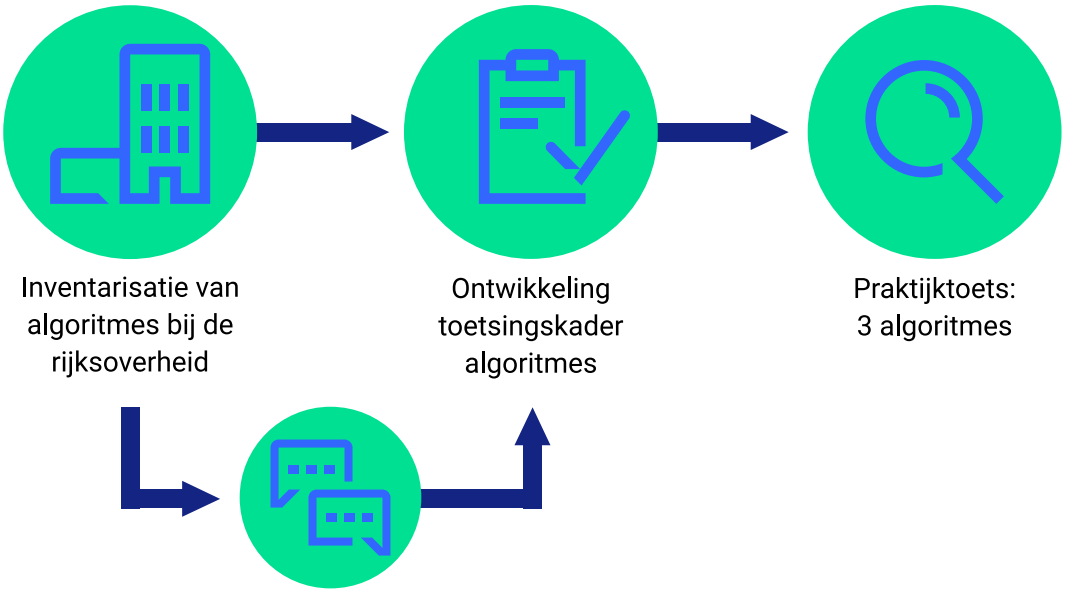
Naast kansen en bedreigingen zijn er mythes en hypes rondom algoritmes. Algoritmes worden vergeleken met menselijke intelligentie en er zijn algoritmes die bij specifieke besluiten beter presteren dan mensen. Het idee kan postvatten dat de overheid de grip op de besluiten die ze neemt kwijt is; dat kan begrijpelijkerwijs leiden tot grote onrust. Een algoritme kan bij interactie met de omgeving heel 'intelligent' overkomen. Van intelligentie is bij algoritmes echter geen sprake. Algoritmes hebben geen bewustzijn en geen beeld van de werkelijkheid.

Uitgangspunt is dat als de rijksoverheid algoritmes toepast; dit moet leiden tot een efficiëntere bedrijfsvoering of betere dienstverlening naar de burgers. Algoritmes zijn een middel om een doel te bereiken, geen doel op zich.
Op dit moment zijn algoritmes in de meeste gevallen instructies die een computer met behulp van data uitvoert om tot een beslissing te komen. De complexiteit en snelheid van algoritmes neemt echter toe. Deze toename in combinatie met de potentiele maatschappelijke onrust zorgt ervoor dat bij de controleurs en toezicht-houders steeds meer behoefte is aan concrete richtlijnen of toetsingskaders waarmee zij algoritmes kunnen analyseren en beoordelen.

2.2 Wat hebben we onderzocht en hoe?

Wij hebben een verkennend en beoordelend onderzoek gedaan naar voorspellende en voorschrijvende algoritmes die een relevante impact hebben op werkprocessen en/of dienstverlening van de rijksoverheid en van organisaties die aan de overheid zijn verbonden.
Een *voorspellend* algoritme wordt ingezet voor een analyse van 'Wat zal er gebeuren?', een *voorschrijvend* algoritme voor een analyse van 'Wat moet er gebeuren?' In ons onderzoek hebben we voortgebouwd op de indeling/typering die is beschreven in de bijlage bij de Kamerbrief over waarborgen tegen risico's van data-analyses door de overheid.[10] In bijlage 1 staat een uitgebreide toelichting over de methodologische verantwoording van ons onderzoek. We hebben in dit onderzoek nadrukkelijk niet gestreefd naar een volledige inventarisatie van alle algoritmes bij de rijksoverheid.

De aanpak van het onderzoek naar algoritmes bestaat uit drie onderdelen



Inventarisatie

Het onderzoek is gestart met ons verzoek aan de kerndepartementen om te inventariseren welke relevante toepassingen van voorspellende en voorschrijvende algoritmes ze gebruiken. We hebben aangegeven dat we voor dit onderzoek graag informatie willen hebben over de algoritmes die:

- een *voorspellende* of een *voorschrijvende* functie hebben, én
- substantiële impact hebben op overheidshandelen of op beslissingen over een concrete casus, burger of een bedrijf.

We hebben de doelen onderzocht waarvoor deze algoritmes worden ingezet, wat de impact daarvan is op burgers en op welke manier sturing en verantwoording plaatsvindt. Ons onderzoek is gericht op de beantwoording van de volgende onderzoeksvragen:

1. Voor welke activiteiten en processen worden algoritmes toegepast bij de rijksoverheid en bij organisaties die aan de rijksoverheid zijn verbonden, welke typen/categorieën zijn er te onderscheiden en wat zijn de effecten en risico's? (§ 3.2)
2. Hoe is de besturing en kwaliteitsbeheersing van algoritmes vormgegeven bij de rijksoverheid en bij organisaties die aan de rijksoverheid zijn verbonden? (§ 3.3)

Denksessie september 2020

Tijdens onze inventarisatie hebben wij opgemerkt dat de uitvoerende

functionarissen die zich bezighouden met het ontwerp, de implementatie of het beheer van algoritmes, behoefte hebben aan meer interdepartementale samenwerking en aan praktische handvatten om op een verantwoorde manier met algoritmes te werken. Om aan deze behoeften tegemoet te komen, organiseerden wij op 22 september 2020 een denksessie in samenwerking met het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK), het Ministerie van Veiligheid en Justitie (VenJ) en Agentschap Telecom van het Ministerie van Economische Zaken en Klimaat (EZK). Deze partijen vervullen een voortrekkersrol binnen de rijksoverheid als het gaat om algoritmes. Aan de sessie namen 30 experts[11] van binnen en buiten de overheid deel. De opbrengsten van deze sessie zijn opgenomen in hoofdstuk 4.

Toetsingskader
Het toetsingskader dat de Algemene Rekenkamer binnen dit onderzoek heeft ontwikkeld, maakt gebruik van diverse al beschikbare informatie, kaders en raamwerken. Het toetsingskader is een praktisch handvat dat wij bij onze toekomstige onderzoeken zullen inzetten. Maar ook andere (overheids)organisaties kunnen het kader gebruiken om te toetsen of hun algoritmes aan bepaalde kwaliteitscriteria voldoen én of de risico's voldoende in beeld of beperkt in beeld zijn. Het toetsingskader maakt deel uit van de rapportage van dit onderzoek en is voor iedereen toegankelijk: www.rekenkamer.nl/algoritmes-toetsingskader.

Praktijktoets: 3 algoritmes
Vervolgens hebben wij 3 algoritmes uit onze inventarisatie geselecteerd en beoordeeld met behulp van het toetsingskader dat wij hebben ontwikkeld. Het doel is om op deze manier ons toetsingskader aan de praktijk te toetsen op bruikbaarheid en om het kader verder aan te scherpen. De beoordeling van algoritmes stelt ons tenslotte in staat noodzakelijke verbeterpunten te adresseren over de beheersing van de risico's bij de inzet van algoritmes binnen de rijksoverheid.

2.3 Leeswijzer

Dit onderzoek bestaat uit 3 onderdelen. In hoofdstuk 3 beschrijven we wat onze inventarisatie binnen de rijksoverheid en binnen organisaties die aan de overheid zijn verbonden, heeft opgeleverd. Hoofdstuk 4 bevat een toelichting op de totstandkoming van het toetsingskader, de 5 perspectieven waar het toetsingskader uit bestaat en de denksessie die als onderdeel van dit onderzoek heeft plaatsgevonden op 22 september 2020. In hoofdstuk 5 staan de belangrijkste observaties en aandachtspunten die naar voren kwamen uit de praktijktest van ons toetsingskader. Daarna volgt hoofdstuk 6 met onze conclusies en aanbevelingen.

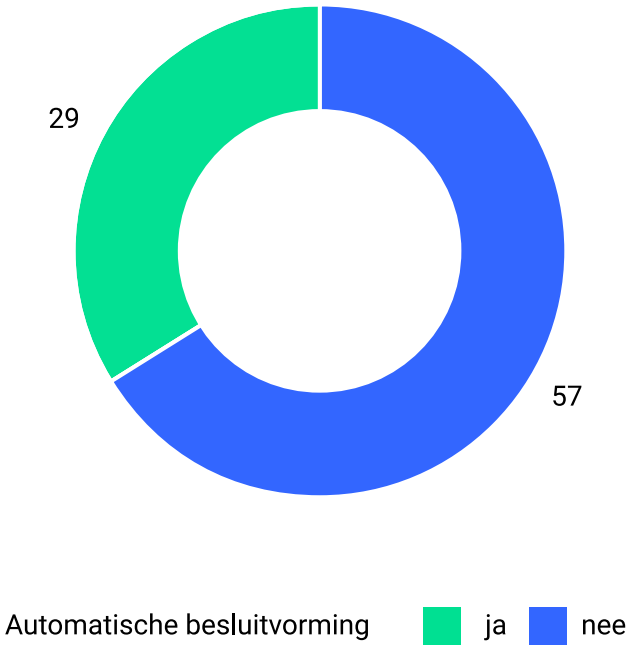
3.
Inzicht in algoritmes

3.1 Totaalbeeld algoritmes

Wij hebben geïnventariseerd welke voorspellende en voorschrijvende algoritmes in gebruik zijn binnen de rijksoverheid. Op basis hiervan hebben wij een eerste indruk gekregen van algoritmes die ingezet worden bij beslissingen die van belang zijn voor burgers en bedrijven. Het betreft een zelfrapportage van alle ministeries, waarbij onze focus op voorspellende en voorschrijvende algoritmes ligt. Dit geeft wel een goed beeld, maar niet een uitputtend beeld van alle algoritmes die binnen de rijksoverheid gebruikt worden.

Uit de aangeleverde algoritmes blijkt dat ongeveer een derde van deze geïnventariseerde voorspellende en voorschrijvende algoritmes, opereren op basis van automatische besluitvorming. We hebben geen volledig zelflerende algoritmes aangetroffen in deze inventarisatie binnen de rijksoverheid; alleen lerende algoritmes. Automatische besluitvorming vindt alleen plaats bij algoritmes die eenvoudige administratieve handelingen uitvoeren, zonder enige impact voor de burger.

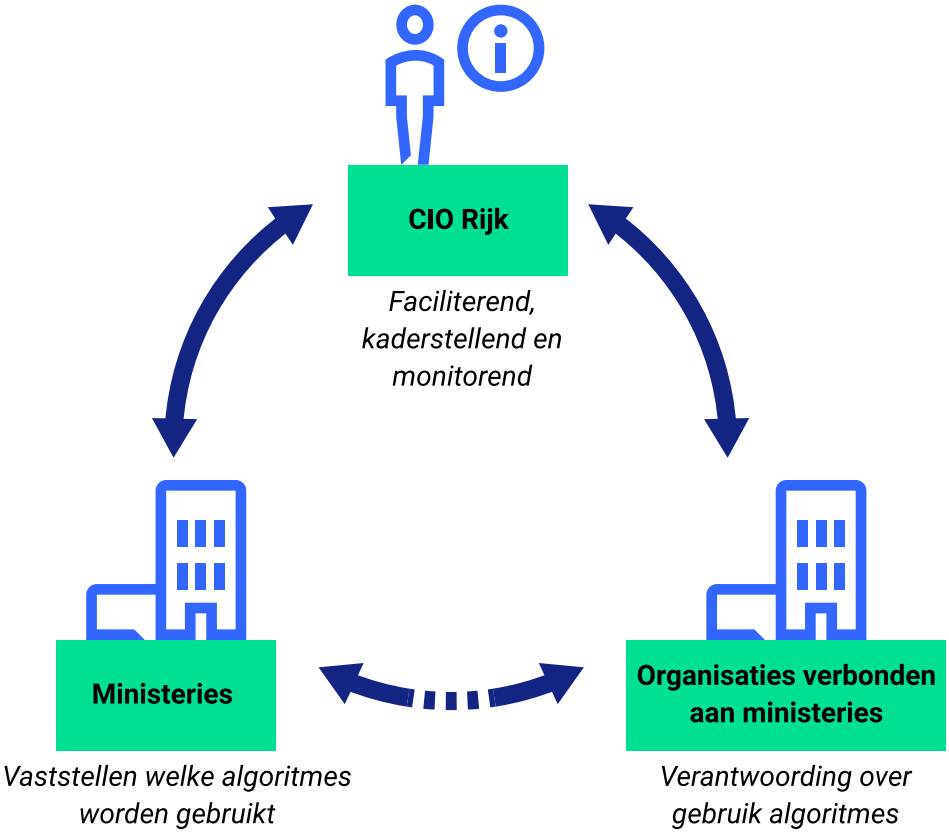
Meerderheid aangeleverde algoritmes heeft geen automatische besluitvorming



Uit de reacties van de ministeries blijkt dat alle ministeries, met uitzondering van het Ministerie van Algemene Zaken die de gevraagde algoritmes niet gebruikt, zowel voorspellende als voorschrijvende algoritmes gebruiken in hun dienstverlening. Van de aangeleverde algoritmes is de verhouding voorspellend – voorschrijvend bijna gelijk, 60% van de algoritmes is voorspellend.

Het aantal voorspellende en voorschrijvende algoritmes dat is aangeleverd voor dit onderzoek, verschilt per overheidsorganisatie. Grote overheidsorganisaties als UWV en SVB (Sociale Verzekeringsbank) verstrekken gelden, uitkeringen en subsidies die direct op wetten zijn gebaseerd. Deze instanties worden gekenmerkt door het gebruik van voorschrijvende algoritmes.[12] Het aantal zegt niet automatisch iets over de expertise van de instanties over algoritmes, omdat de algoritmes verschillen in hun complexiteit en mogelijke impact. Daarnaast stellen we vast dat er geen eenduidige definitie en geen eenduidige categorisering van algoritmes is binnen de rijksoverheid, waardoor we interpretatieverschillen zagen tussen de ministeries bij de aanlevering van de algoritmes.

Ministeries en CIO Rijk hebben beperkt zicht op de algoritmes die worden ingezet



Bijna alle ministeries en de CIO Rijk geven aan dat zij centraal (vanuit het kerndepartement) geen goed beeld hebben van de algoritmes die het ministerie zelf gebruikt. Het gevolg hiervan is dat de ministers de risico's en mogelijke (negatieve) impact van algoritmes op dienstverlening door de rijksoverheid niet tijdig kunnen beperken. Dat gebrek aan inzicht geldt ook voor organisaties die zijn verbonden aan het ministerie, zie bovenstaande figuur. Een aantal ministeries en de CIO Rijk geeft aan dat ons onderzoek een eerste stap is geweest om een realistisch beeld te krijgen van het gebruik van algoritmes binnen hun organisatie.

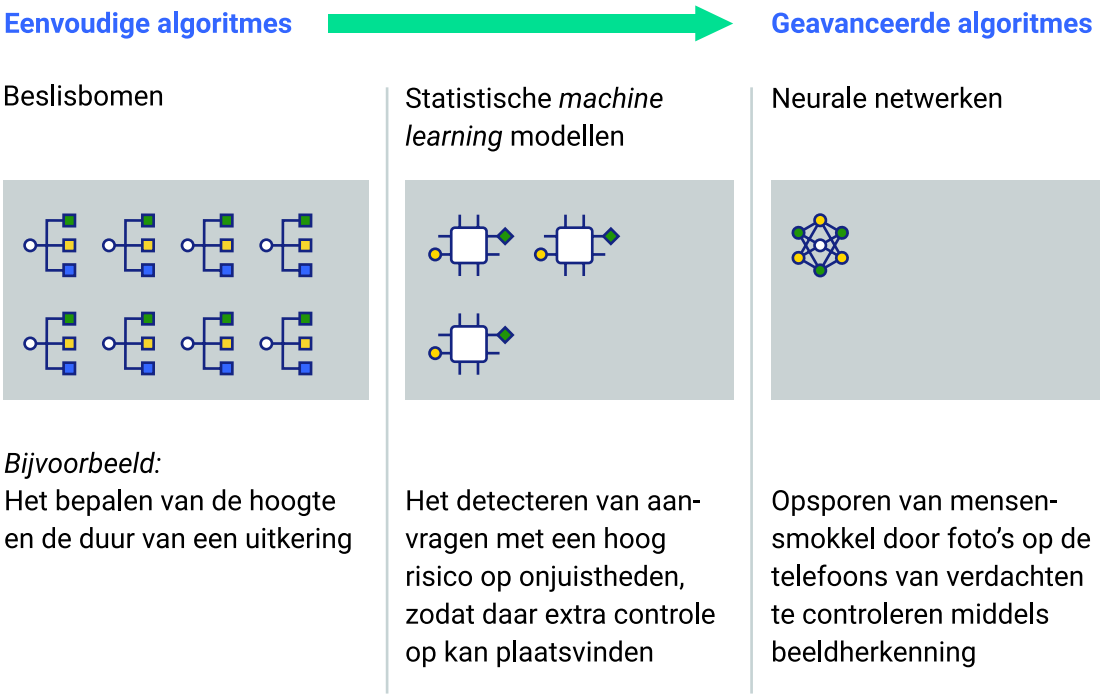
3.2 Voor welke activiteiten en processen worden algoritmes toegepast bij de rijksoverheid en bij organisaties die aan de overheid zijn verbonden, welke typen/categorieën zijn er te onderscheiden en wat zijn de effecten en risico's?

Voor de nadere indeling van algoritmes gebruiken wij de input van de bijlage bij de brief van de minister van Rechtsbescherming aan de Kamer van 8 oktober 2019.[13]

We maken onderscheid tussen algoritmes op basis van hun complexiteit, van eenvoudig tot complex. Een beslisboom is een voorbeeld van een eenvoudig algoritme. Keuzes van deze algoritmes zijn goed uit te leggen. Een voorbeeld hiervan is het bepalen van de hoogte van een uitkering.

Een *deep learning*[14]-algoritme is daarentegen een complex algoritme. Voorspellingen die dit soort algoritme maakt, zijn moeilijk te doorgronden, omdat het voor de beoordelaar niet zonder meer duidelijk is welke kenmerken van de data het algoritme zwaar laat meewegen. Een voorbeeld hiervan is Siri (het spraakherkenningsprogramma van Apple) en AlphaGo. Het laatste programma is een computerprogramma dat Google heeft ontwikkeld en dat in 2016 de menselijke wereldkampioen in het bordspel Go heeft verslagen.

De rijksoverheid gebruikt vooral eenvoudige algoritmes, nauwelijks geavanceerde algoritmes



Tussen deze 2 uitersten van categorieën zijn algoritmes in verschillende gradaties van complexiteit en mate van uitlegbaarheid mogelijk. Uit onze inventarisatie en bovenstaande figuur blijkt dat de overheid eenvoudige en geavanceerde, voorspellende en voorschrijvende algoritmes inzet. De aangeleverde algoritmes betreffen vooral eenvoudige algoritmes en de midden categorie algoritmes. Van de aangeleverde algoritmes betreft maar maximaal 10% de categorie geavanceerde

algoritmes. De algoritmes raken verschillende processen en onderdelen van de rijksoverheid. Een groot deel van de algoritmes biedt ondersteuning in het werkproces, waardoor zaken efficiënter af te handelen zijn. We onderscheiden 3 doelen in de toepassingen van algoritmes bij de rijksoverheid, met verschillende effecten en risico's. De helft van de aangeleverde algoritmes betreft doel 1, de andere helft van de algoritmes is gelijkmatig verspreid over doel 2 en 3.

Doel 1: Automatiseren van administratie en uitvoeren van eenvoudige wetgeving
Een deel van de algoritmes wordt ingezet om eenvoudige menselijke handelingen te automatiseren. De rijksoverheid maakt hier op grote schaal gebruik van. Dit kan veel efficiencywinst opleveren, vooral door veel snellere verwerking van grote datavolumes. Het gaat bij deze algoritmes vaak om het geautomatiseerd uitvoeren van wetgeving.

Een voorbeeld hiervan is het algoritme 'subsidie woonhuismonumenten' van de Rijksdienst voor het Cultureel Erfgoed. Een beslisboom (met eenvoudige 'als-dan'-regels) bepaalt of particuliere eigenaren van rijksmonumenten recht hebben op een subsidie. Kenmerkend aan dit soort algoritmes is dat ze vaak voorschrijvend zijn en automatisch een handeling uitvoeren, zonder tussenkomst van een mens. Het gaat om administratieve en financiële activiteiten. Het risico op fouten met impact op de burger hierbij is laag, omdat het een eenvoudige handeling is die uitgevoerd wordt door een eenvoudig algoritme, met een hoge technische transparantie en lage foutkans.

Doel 2: Verbeteren en faciliteren van bedrijfsvoering
Algoritmes met als doel om overheidsprocessen efficiënter en beter te maken, gebruiken complexere data. De uitkomst is niet altijd een-op-een over te nemen door de betrokken expert. Ze doen een voorspelling of analyse, waar een expert vervolgens nog mee aan de slag moet gaan.
Een voorbeeld is de Object Detectie Sonar van Rijkswaterstaat. Dit algoritme geeft aan waar objecten zich in de zee bevinden, op basis van metingen van de zeebodem, zodat een expert weet of een waterbouwkundig project kan worden opgestart. Een ander voorbeeld is de voorspelling van het aantal telefoontjes bij een klantcontactcentrum, zodat het benodigd aantal in te plannen medewerkers bepaald kan worden. Het zijn vaak voorspellende algoritmes zonder automatische besluitvorming. Het risico op fouten naar de burger of op een omvangrijke financiële stroom is aanwezig maar is beperkt. Het algoritme doet namelijk alleen voorbereidend 'werk'; een analyse, waarna een expert hiermee verder gaat en het uiteindelijke besluit neemt.

Doel 3: Gerichte inzet van capaciteit en middelen op basis van risicovoorspelling
Dit zijn algoritmes die functionarissen ondersteunen bij het selecteren van casussen voor nader onderzoek. Deze algoritmes helpen de beschikbare hoeveelheid mensen en middelen efficiënt in te zetten. Een voorbeeld is het visumaanvraagproces. Het Ministerie van Buitenlandse Zaken (BuZa) zet een algoritme in dat helpt bij het indelen van alle aanvragen van visa in verschillende tracks. De aanvragen worden door het algoritme ingedeeld naar kansrijke en complexe/risicovolle aanvragen, waarna een medewerker de aanvraag controleert. Het algoritme informeert de medewerker aan welke aanvragen waarschijnlijk meer tijd besteed moet worden, zonder automatisch te beslissen over de aanvraag.

Uit eerder onderzoek bleek al dat risicogericht controleren op brede schaal binnen de rijksoverheid plaatsvindt. In onze inventarisatie wordt dit bevestigd. De Belastingdienst[15] maakt er veel gebruik van, bijvoorbeeld voor het doelgericht controleren van belastingaangiftes. Kenmerkend is dat het algoritme een advies geeft. De betrokken functionaris kan in zijn professionele oordeelsvorming van dit advies afwijken. Er is dus geen sprake van automatische besluitvorming.

Bij deze ondersteunende algoritmes voor een risicovoorspelling bestaat het risico dat de uitgangspunten van het risicoprofiel strijdig zijn met de geldende wet- en regelgeving danwel een (ongewenste) afwijking gaan vertonen op basis van de verborgen beperkingen van de inputdata. Denk aan discriminatie of het gebruik van bijzondere persoonsgegevens. Ook bestaat de kans dat het advies van het algoritme de uiteindelijke beslissing van de medewerker beïnvloedt.

3.3 Hoe is de besturing en kwaliteitsbeheersing van algoritmes vormgegeven?

Uit de inventarisatie blijkt dat voor de besturing en het kwaliteitsbeheer van algoritmes met algemene normenkaders en richtlijnen wordt gewerkt waarbij de focus vooral ligt op de AVG (Algemene Verordening Gegevensbescherming) en de BIO (Baseline Informatiebeveiliging Overheid). Vanuit de ministeries en organisaties betrokken medewerkers zijn op zoek naar een geheel van relevante normenkaders of richtlijnen specifiek voor algoritmes die recht doen aan de bredere politieke en maatschappelijke discussie over algoritmes. De medewerkers worstelen met de manier waarop de besturing en beheersing in de praktijk moet worden vormgegeven. Binnen veel ministeries en met het Rijk verbonden organen is behoefte aan een toetsingskader om meer grip te krijgen op algoritmes, vooral omdat de specifieke risico's van algoritmes niet altijd bekend of duidelijk zijn.

Er is in de vragenlijsten door bijna alle betrokkenen werkzaam bij ministeries aangegeven dat er ook behoefte is aan centrale kaderstelling binnen de rijksoverheid over het gebruik en de risicobeheersing van algoritmes. De verantwoordelijke ministers kunnen rust creëren door hierover een standpunt in te nemen dat tegemoet komt aan de behoeften van zowel interne als externe betrokkenen. Dit kan via één gezamenlijk toetsingskader binnen de rijksoverheid. In de inventarisatie geven medewerkers van een drietal ministeries aan dat een toetsingskader niet generiek kan zijn. Bij het beoordelen van de 3 algoritmes in de praktijk hebben wij gezien dat de risico's van algoritmes die beoordeeld moeten worden vrij generiek zijn. Uit ons onderzoek concluderen wij dat een generiek toetsingskader kan worden toegepast op rijksniveau.

4. Toetsingskader algoritmes

De grote maatschappelijke aandacht voor algoritmes heeft geleid tot vele initiatieven, normen en toetsingskaders van verschillende partijen en vanuit verschillende invalshoeken. Er zijn tot nu toe echter geen integrale en concrete instrumenten voor toetsing of analyse van algoritmes. Met integraal doelen wij op het gebrek aan het samenbrengen van alle relevante normenkaders en richtlijnen voor algoritmes in één geheel. Met de concretisering bedoelen wij een vertaling van de normenkaders en richtlijnen naar te toetsen aspecten, de bijbehorende risico's en de onderzoeksvragen die aan bod moeten komen.

Binnen het Rijk zijn bijna alle ministeries actief op dit onderwerp. Ook buiten de rijksoverheid zijn organisaties actief op dit onderwerp, zoals NOREA, de beroeps-organisatie voor IT-auditors in Nederland, en grote accountantskantoren. Het toetsingskader dat de Algemene Rekenkamer in dit onderzoek heeft ontwikkeld, maakt zo veel mogelijk gebruik van beschikbare informatie, kaders en raamwerken. Het toetsingskader is een praktisch handvat dat wij bij onze toekomstige onder-zoeken willen inzetten. Maar ook andere overheidsorganisaties kunnen het kader gebruiken om te toetsen of hun algoritmes aan bepaalde kwaliteitscriteria voldoen én of de risico's voldoende in beeld zijn en/of worden beperkt. We beogen hiermee duidelijk en transparant te zijn over de vragen die bij toekomstige onderzoeken naar algoritmes aan de orde zullen komen. Ministeries krijgen via dit toetsingskader nu al zicht op de risico's die wij onderkennen, en kunnen dus nu al maatregelen nemen om die risico's te beperken. Het toetsingskader maakt deel uit van de rapportage van dit onderzoek en is voor iedereen toegankelijk: www.rekenkamer.nl/algoritmes-toetsingskader.

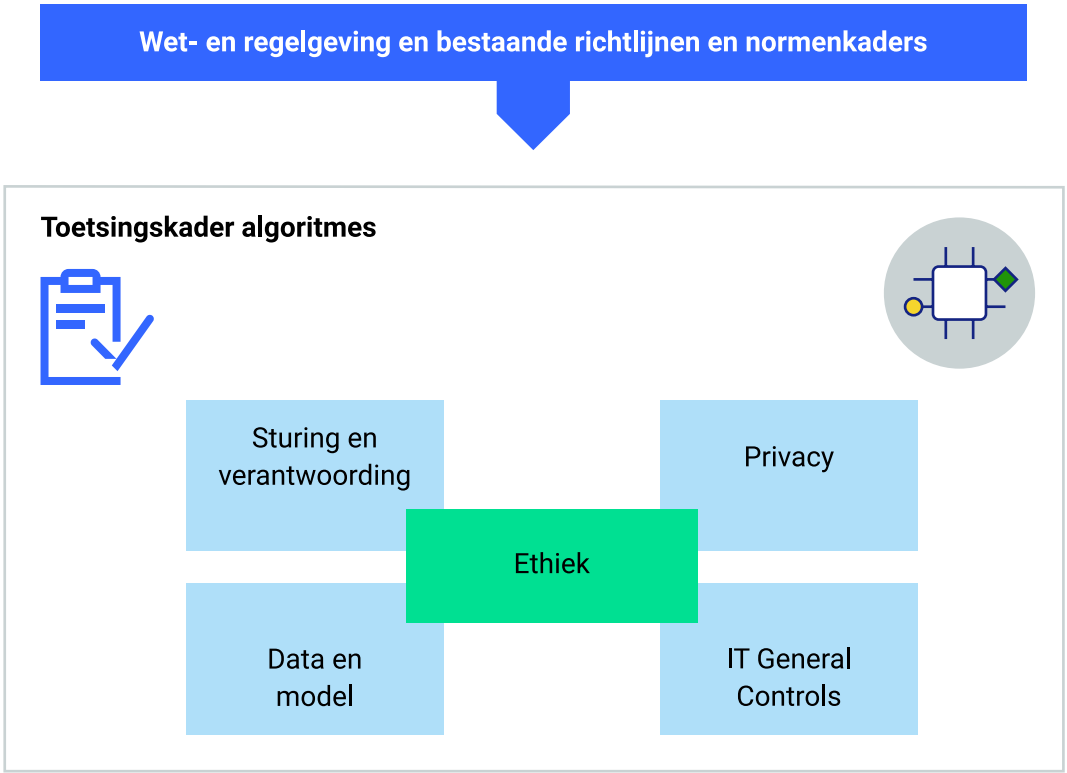
4.1 Vijf perspectieven

Het toetsingskader bestaat uit 5 perspectieven:

1. sturing en verantwoording;
2. model en data;
3. privacy;
4. *IT General Controls* (ITGC);
5. ethiek.

Ethiek is geen separaat onderdeel in het toetsingskader maar verweven in alle 4 de perspectieven. Het toetsingskader is gebaseerd op bestaande normenkaders en richtlijnen (zie bijlage 2), waarbij wij een concretisering hebben gemaakt naar de aspecten die moeten worden getoetst, de bijbehorende risico's en de onderzoeks-vragen die tijdens de toetsing aan bod moeten komen.

Het toetsingskader algoritmes bestaat uit 5 perspectieven



Sturing en verantwoording

Bij het onderdeel sturing en verantwoording in het toetsingskader gaat het om het vastleggen van verschillende elementen: de rollen, verantwoordelijkheden en deskundigheid, het *lifecycle management* van het algoritme, risico-afwegingen bij het gebruik van het algoritme en afspraken met externe partijen over bijvoorbeeld

aansprakelijkheid. Wij hebben gebruikgemaakt van bestaande kaders voor *IT-governance* om de toetsing vorm te geven van de algoritmes die we hebben onderzocht op sturing en verantwoording. COBIT[16] (*Control Objectives for Information and related Technology*) is de basis van de elementen die we in het toetsingskader hebben opgenomen over sturing en verantwoording.

Model en data

In het onderdeel model en data komen vragen aan bod over de kwaliteit van de data en over de ontwikkeling, het gebruik en het onderhoud van het model onderliggend aan het algoritme. Er worden vragen gesteld over eventuele vooroordelen (op basis van het ethisch perspectief) in de data, dataminimalisatie en of de output van het model wordt getoetst. We hebben hiervoor geput uit de wetenschappelijke literatuur en de *machine learning*-praktijk. Het zwaartepunt van de eisen in ons toetsingskader ligt bij de ontwikkeling van het model, maar we besteden ook aandacht aan de werking, het gebruik én het onderhoud in de praktijk. We hebben het toetsingskader toepasbaar gemaakt voor het hele spectrum aan algoritmes: van eenvoudige beslismodellen tot machine learning-modellen. Dat kan ertoe leiden dat een onderdeel van het toetsingskader niet van toepassing is op een specifiek algoritme.

Privacy

Bij het gebruik van algoritmes wordt in een aantal gevallen gebruikgemaakt van (bijzondere) persoonsgegevens.[17] Het is van belang dat algoritmes voldoen aan de wettelijke verplichtingen die gelden voor het verwerken van persoonsgegevens. De Algemene Verordening Gegevensbescherming (AVG) is een belangrijke bron voor ons kader.

***IT General Controls* (ITGC)**

IT General Controls (ITGC) zijn de beheersmaatregelen die een organisatie heeft getroffen om ervoor te zorgen dat de IT-systemen betrouwbaar en integer zijn. Het zijn traditionele ICT-maatregelen, zoals het beheer van toegangsrechten, continuïteit en change management. Bij het onderdeel ITGC wordt gekeken naar de loggings-informatie, de toegangsrechten en het wachtwoordbeheer van het algoritme. Daarbij gaat het hier om de inbedding in de applicatie en de onderliggende componenten, zoals de database en het besturingssysteem. Belangrijkste normenkaders voor de IT General Controls zijn de internationale norm ISO/IEC 27002 en de BIO.

4.2 Denksessie: begrippen en definities

Toen tijdens het onderzoek bleek dat bij alle betrokkenen in de uitvoering sprake is van interpretatieverschillen van begrippen en definities over algoritmes, hebben wij een denksessie georganiseerd op 22 september 2020. Dit hebben wij gedaan in samenwerking met het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK), het Ministerie van Veiligheid en Justitie (VenJ) en Agentschap Telecom van het Ministerie van Economische Zaken en Klimaat. Het doel van deze sessie was het identificeren, bespreken en zo mogelijk overbruggen van de verschillen in terminologieën die worden gebruikt bij algoritmes. In de denksessie zijn 5 thema's besproken:

1. datagedreven;
2. datakwaliteit;
3. AI en algoritmes;
4. AI bij de overheid;
5. transparantie.

Het verslag van de denksessie is opgenomen in bijlage 1.

5. Praktijktoets: 3 algoritmes

5.1 Selectie algoritmes

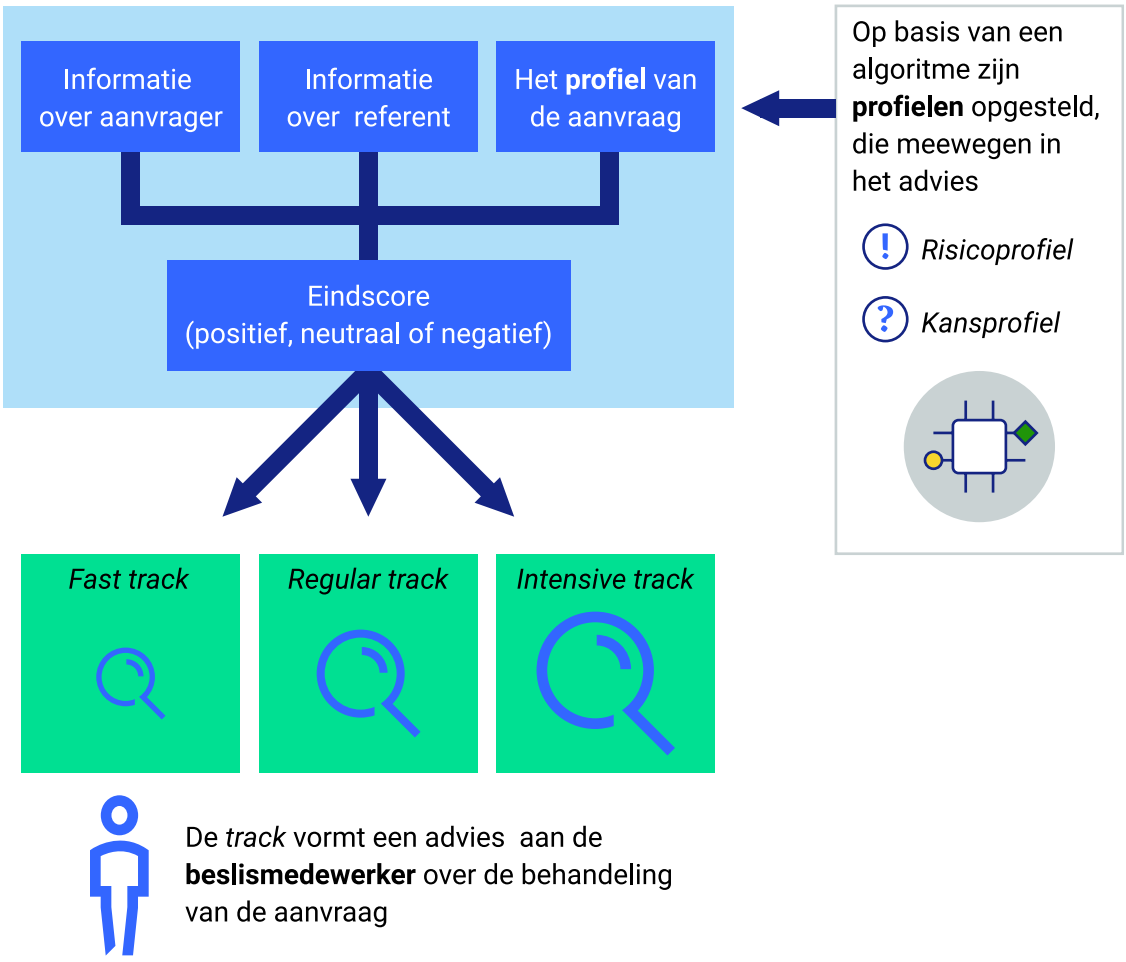
Met de toetsing van 3 algoritmes beoogden we ons toetsingskader aan de praktijk te toetsen op bruikbaarheid. Bovendien wilden we het kader verder aanscherpen. Wij hebben geen beoordeling per casus uitgevoerd. De bevindingen bij de casussen zijn om deze reden ook veralgemeniseerd. Het doel was daarnaast om aanvullend op de inventarisatie meer informatie op te halen over de risico's die samenhangen met algoritmes. Op basis hiervan kunnen wij de noodzakelijke verbeterpunten vaststellen die nodig zijn voor een doorontwikkeling van algoritmes binnen de rijksoverheid.

Het toetsingskader dat we hebben ontwikkeld, hebben we toegepast op 3 concrete algoritmes:

1. een beslisboom die advies geeft voor (extra) controle op aanvragen van personen;

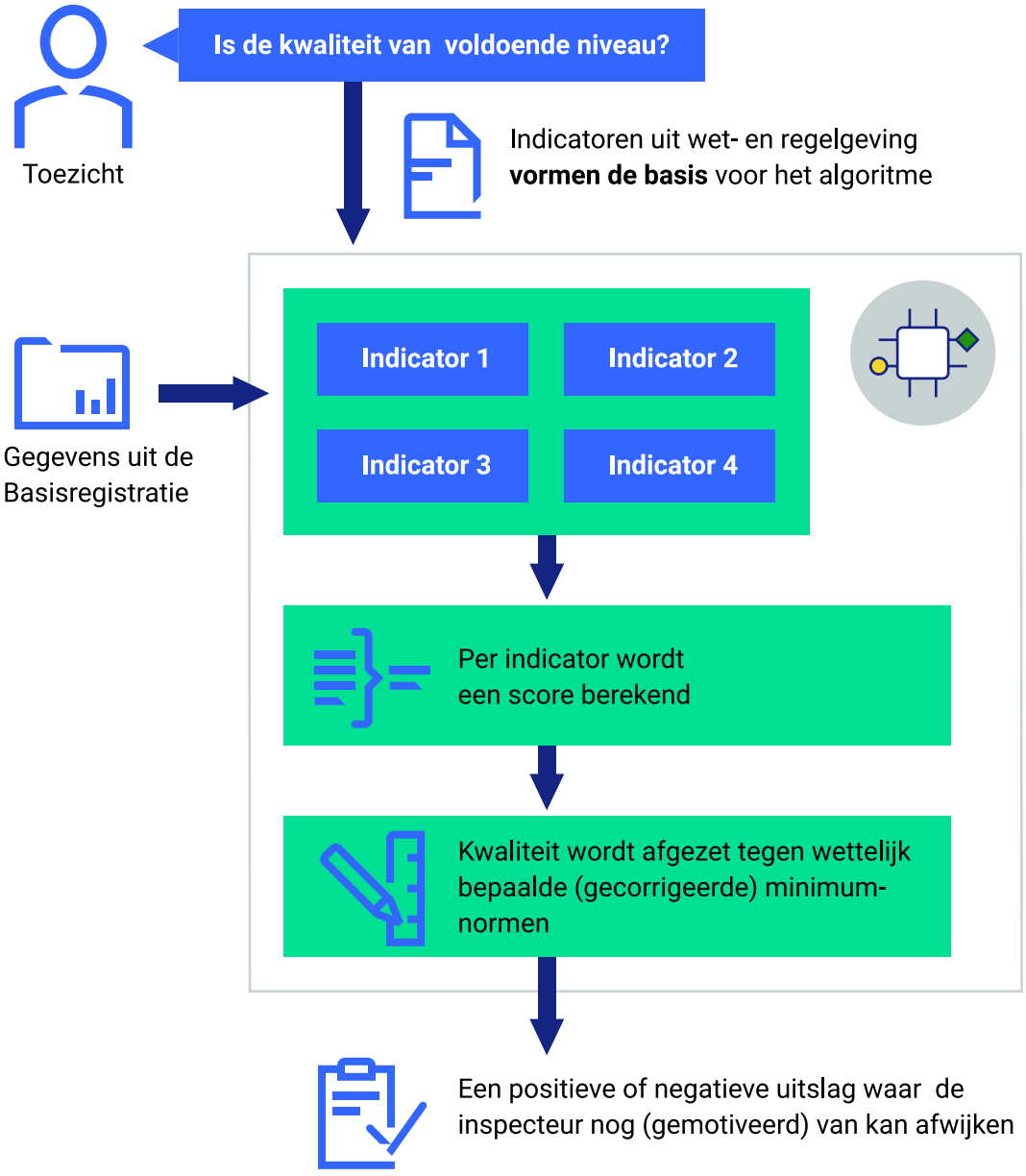
Een algoritme op basis van simpele beslisregels kan ambtenaren helpen om aanvragen efficiënter te controleren

Een persoon of bedrijf doet een aanvraag voor bijvoorbeeld een subsidie, een (reis)-document, of een uitkering. Moet deze aanvraag (extra) gecontroleerd worden?



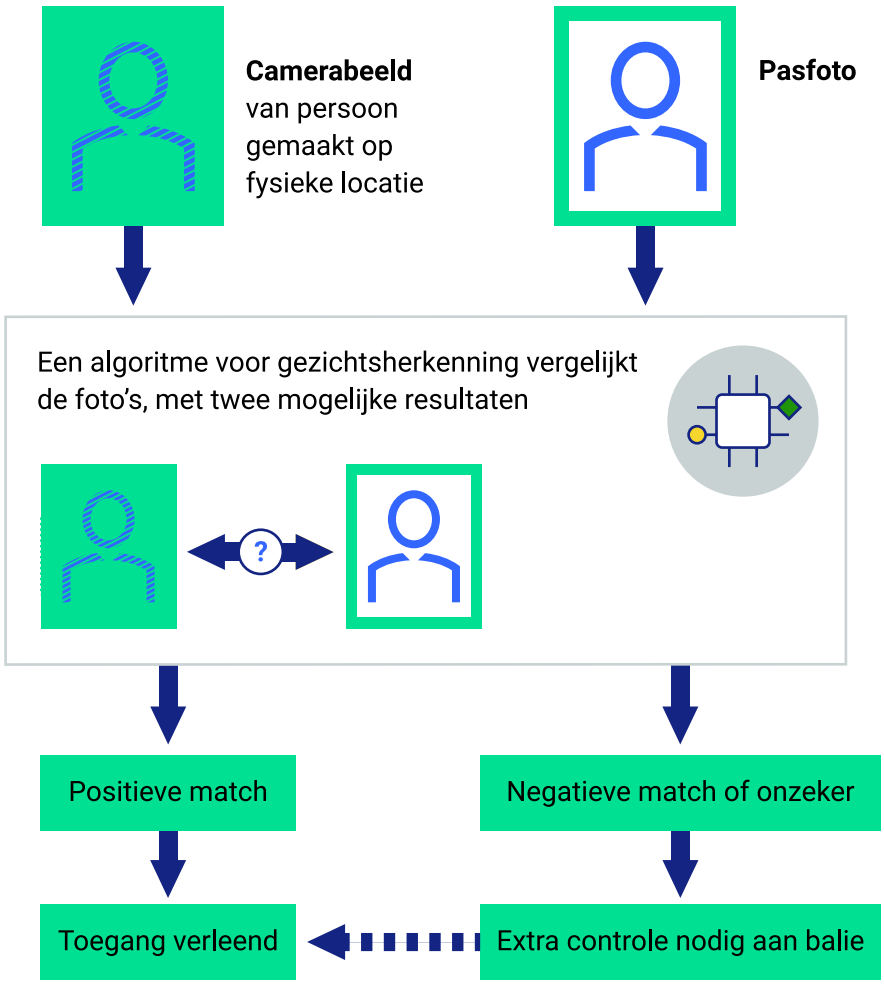
2. een beoordelingssysteem om afwijkend presterende objecten te detecteren als input voor toezichthouders en inspecties;

Toezichthouders en inspecties worden bij hun beoordelingen ondersteund door algoritmes



3. beeldherkenning voor het toekennen van fysieke toegang tot een terrein of gebouw.

Algoritmes kunnen, op basis van beeldherkenning, helpen bij het verlenen van fysieke toegang aan personen



We hebben deze algoritmes geselecteerd om de volgende redenen:

1. het zijn voorspellende en/of voorschrijvende algoritmes die in de praktijk worden toegepast;
2. de algoritmes hebben substantiële impact op burgers en bedrijven;
3. de algoritmes maken gebruik van verschillende technieken.

5.2 Belangrijkste inhoudelijke observaties

Sturing en verantwoording

De mate waarin de onderzochte algoritmes voldoen aan de gestelde eisen voor sturing en verantwoording verschilt. Voor een algoritme troffen we onderbouwing en vastlegging van de afgelopen jaren aan van de uitgangspunten en eisen die op het algoritme van toepassing waren. Bij een ander algoritme gaven de documenten geen duidelijkheid. Dit laatste betekent overigens niet dat het inzicht van het ministerie in het doel en functioneren van het algoritme volledig ontbreekt. Betrokken medewerkers van het ministerie hebben dit minimaal op hoofdlijnen in beeld. In alle casussen zien we dat er een periodieke toetsing en herijking[18] van het algoritme plaatsvindt.

Wij signaleren in de 3 casussen dat de afspraken, rollen, taken en verantwoordelijkheden van partijen die bij de rijksoverheid betrokken zijn bij het gebruik van algoritmes, belegd en verduidelijkt moeten worden. Dit is nodig om als per departement dan wel uitvoerende instantie, onder verantwoordelijkheid van de CIO, systematisch inzichtelijk te hebben in de vraag of het algoritme doet wat het moet doen. Ook stelden wij vast dat er veelal geen *lifecycle management*[19] is ingericht voor algoritmes. Er gaat veel aandacht naar het ontwerp en de ingebruikname van het algoritme, maar niet naar de instandhouding en het onderhoud ervan. Dat heeft niet alleen technisch inhoudelijke consequenties maar ook budgettair. Tekort aan onderhoudsbudget, onvoldoende onderhoud of niet adequate bemensing kan ertoe leiden dat het algoritme in de loop van de tijd niet meer voldoet aan de veranderde ethische principes of wet- en regelgeving.

Model en data

Het begrip uitlegbaarheid wordt niet eenduidig toegepast. In één van de 3 casussen is geprobeerd de uitkomsten van het model te verduidelijken. In een andere casus wordt juist bewust niet gestreefd naar transparantie. In deze casus geeft het algoritme alleen een signaal dat iets niet in orde is met een aanvraag van een burger, maar niet de reden waarom. Op deze manier is door de betrokken uitvoeringsorganisatie getracht om de beoordelaar te stimuleren zelfstandig onderzoek te doen. Dit voorkomt dat een automatisch besluit met minimale menselijke tussenkomst wordt genomen.

De aandachtspunten bij model en data gaan zowel over de techniek achter het ontwikkelen van modellen voor het algoritme, als over de kwaliteit van de data. Als het gaat om techniek is de deskundigheid van de medewerkers vaak in orde. Wij zien 2 mogelijke risico's bij het datamanagement:

1. Het eerste risico is dat gebruikgemaakt wordt van historische data en dat daarmee maatschappelijke vernieuwing wordt afgeremd. Wat in het verleden gold, wordt op het heden toegepast. Bijvoorbeeld: welke competenties heeft een goede leidinggevende? Dat verandert als gevolg van maatschappelijke ontwikkelingen. Wanneer er nog geen data beschikbaar is als gevolg van nieuwe wet- en regelgeving, is de inzet van een algoritme niet mogelijk.
2. Vooroordelen in de data zijn een tweede risico. Wanneer een bepaalde bevolkingsgroep afwijkend is behandeld in het verleden, dan zal een algoritme deze discriminatie overnemen.

Uit de inventarisatie en de 3 casussen blijkt dat al bij de ontwikkeling van algoritmes niet alle relevante kennisdisciplines betrokken zijn. Vaak zijn privacy-deskundigen, programmeurs/dataspecialisten wel betrokken bij de ontwikkeling, maar juristen en beleidsadviseurs niet. Dat kan ertoe leiden dat een algoritme niet voldoet aan alle wet- en regelgeving, ethische principes of dat het algoritme niet bijdraagt aan het beleidsdoel. Ook ontbreken vaak maatregelen om ethische risico's zoals vooroordelen in de gekozen data te beperken.

Privacy

Het belangrijkste kader voor privacy en gegevensbescherming is de AVG.
We hebben de 3 casussen getoetst aan ons toetsingskader. In het onderdeel privacy komen elementen aan de orde zoals het verwerkingsregister van persoonsgegevens, het uitvoeren van privacy impact assessments, een wettelijke grondslag voor gebruik van gegevens en data minimalisatie. De onderzochte casussen voldoen vrijwel geheel aan de eisen die wij relevant achten voor algoritmes op het gebied van privacy. Bij 1 casus was het privacybeleid, de gebruikte data en de algoritmes niet in voldoende detail openbaar beschikbaar. Dit is belangrijk voor derden zoals burgers om te weten welke data worden gebruikt, hoe het algoritme functioneert en welke impact dat op hen heeft. Zeker als in de toekomst de hoeveelheid gebruikte data toeneemt en algoritmes complexer worden blijft dit een belangrijk aandachtspunt.

Het valt op bij de onderzochte casussen op dat de burger niet op een eenvoudige manier kennis kan nemen van de algoritmes die de overheid toepast en welke data de overheid daarbij gebruikt. En hoe kan de burger vaststellen welke impact deze

algoritmes hebben? Het volstaat niet om alleen aan de formele eisen van de AVG te voldoen. Persoonsgegevens en data die door de burger worden aangeleverd zijn van de burger, die moet weten wat ermee wordt gedaan.

Gegevensverwerkingsregisters zijn niet bij alle casussen openbaar en *privacy-statements* die zijn gekoppeld aan de onderzochte algoritmes, zijn niet altijd duidelijk en toegankelijk genoeg. Soms is de werking van algoritmes en de gebruikte variabelen expliciet vastgelegd in wet- en/of regelgeving maar die is vaak niet eenvoudig te lezen of te begrijpen. Dat heeft tot gevolg dat de burger beperkt zicht heeft op algoritmes. In één praktijkcasus zien we dat door de betrokken functionarissen een extra inspanning wordt gedaan om de gebruikte variabelen eenvoudig te kunnen toelichten. Dit doen deze functionarissen door de wet- en regelgeving te vertalen naar onder andere een lijst met veel gestelde vragen en een video.

In het verlengde van het programma Regie op Gegevens[20] en MijnOverheid[21] is het belangrijk dat burgers weten waar zij terechtkunnen als ze vragen hebben over algoritmes, waar ze fouten in data kunnen aangeven of hoe ze bezwaar kunnen maken tegen gebruik van data of uitkomsten van algoritmes. Op dit moment zijn Data Protection Impact Assessments[22] (DPIA's), *privacy statements* en verwerkings-registers onvoldoende toegankelijk en onvoldoende begrijpelijk voor niet-deskundigen en leken.

IT General Controls (ITGC)

Van de 4 perspectieven in ons toetsingskader, is er voor ITGC het minst aandacht. Dit blijkt uit de (beperkte) documentatie die wij hebben ontvangen van de getoetste partijen. Het gaat hier met name om (beheer) van toegangsrechten en het maken van *back-ups*. In 2 van de 3 gevallen was er weinig of geen informatie beschikbaar over de manier waarop voldaan wordt aan ITGC-normen[23], en kon deze informatie ook niet (snel) worden aangeleverd. In het derde geval hebben wij na enige toelichting wél de benodigde documentatie ontvangen. Wij concluderen daarom dat 2 van de 3 algoritme-eigenaren niet in voldoende mate kunnen aantonen dat zij de risico´s voldoende beheersen. Wij zien een tweetal oorzaken:

- Het algoritme is in beheer bij een externe partij. De betrokken functionarissen vertrouwen erop dat het daar goed geregeld is, maar heeft er zelf weinig zicht op. Als wij om bewijs vragen, kunnen de functionarissen van het onderzochte ministerie hier niet (snel) aan komen.
- ITGC-nomen zijn op een hoger/ander niveau in de organisatie vastgelegd, maar dit is onvoldoende gespecificeerd voor het algoritme.

Onze rijksbrede inventarisatie van algoritmes onderschrijft de oorzaak van uitbesteding van het beheer aan een externe partij. Dit geldt ook voor 2 van de 3 algoritmes uit onze praktijktoets. In één geval is het beheer belegd bij een zogeheten *Shared Service Organisatie* (SSO) van het Rijk. In het andere geval doet een private partij het beheer.

Dit heeft ertoe geleid dat wij voor veel ITGC-normen niet kunnen vaststellen in hoeverre het algoritme daaraan voldoet. De functionarissen van het ministerie dat het algoritme in eigen beheer heeft, waren in staat om alle onderdelen te onderbouwen.

Ethiek

Ethiek is geen losstaand onderdeel in het toetsen van algoritmes. Ethiek is verweven in de 4 verschillende perspectieven die we hiervoor hebben beschreven. Dit betekent dat ethiek van toepassing is op alle vier de perspectieven. Vanuit het perspectief ethiek hebben we 4 onderwerpen onderscheiden, op basis van bestaande bronnen (bijlage 2) en normen:

1. respect voor menselijke autonomie;
2. voorkomen van schade;
3. *fairness* (een eerlijk algoritme);
4. verklaarbaarheid en transparantie.

Respect voor menselijke autonomie

Uit ons onderzoek blijkt dat de 3 algoritmes functioneren als hulpmiddel; ze nemen (nog) geen automatische besluiten. Bij één casus zorgt de technische applicatie (het algoritme) ervoor dat de medewerkers meerdere verschillende bronnen kunnen raadplegen, zodat ze efficiënte beslissingen kunnen nemen. Het algoritme ondersteunt op deze manier de medewerkers.

Voorkomen van schade

Bij het voorkomen van schade is het met name van belang dat het algoritme altijd doet waar het voor gemaakt is. Daarnaast moet de privacy van mensen worden gewaarborgd en de bijbehorende data worden beschermd. Ongeautoriseerde toegang kan leiden tot wijziging, beschadiging en/of verlies van data. Onze bevindingen hierover hebben we bij het onderdeel ITGC toegelicht.

Fairness

Fairness betekent dat het algoritme rekening houdt met diversiteit in de populatie en niet discrimineert. Zonder doeltreffende maatregelen kan er een onwenselijke systematische afwijking voor specifieke personen, groepen of andere eenheden ontstaan. Bij één van de 3 casussen heeft een externe partij het algoritme getoetst op afwijkende prestaties. Bij een andere casus toetst een externe partij alle data *vooraf* kritisch om te beoordelen of deze gegevens onmisbaar zijn voor het doel van het algoritme.

Verklaarbaarheid en transparantie

Eigenaren van een algoritme moeten verantwoording afleggen over de gevolgde procedure bij de totstandkoming van het algoritme, en de werking van het algoritme uitleggen. Bij alle 3 onderzochte casussen is het algoritme uitlegbaar en heeft er een afweging plaatsgevonden tussen de uitlegbaarheid van het model en de prestatie van het model. In geen van de 3 casussen gaat het overigens om een zelflerend algoritme, wat bijdraagt aan de grote mate van uitlegbaarheid van het algoritme.

Om verantwoording af te kunnen leggen over de gevolgde procedure is het van belang dat de procedures zijn gedocumenteerd. Wij constateren dat dit zowel bij de algoritmes in eigen beheer als in volledig beheer van een (externe) partij een aandachtspunt is. Het algoritme in eigen beheer had geen vastlegging van het ontwerp van het model, wel de parameters.

Tot slot

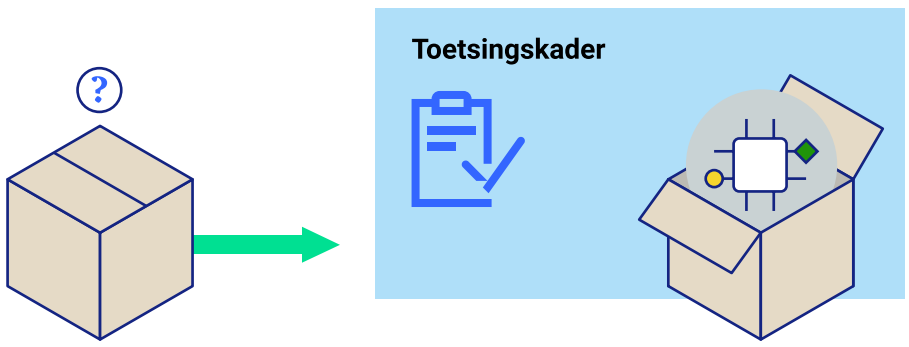
Voor de toetsing op de naleving van de ethische principes fairness en verklaarbaarheid en transparantie moet de onafhankelijk controleur kunnen vaststellen welke data gehanteerd is en deze data kunnen testen. In één casus zijn de gegevens niet bewaard om te kunnen voldoen aan de privacywetgeving. De data zijn dan voor ons als onafhankelijk controleur achteraf niet controleerbaar. Een externe partij heeft deze vooraf getoetst. De casus voldoet daarmee aan de privacywetgeving maar wij kunnen zelf niet meer vaststellen of de ethische principes worden nageleefd.

6. Conclusies en aanbevelingen

Wij hebben onderzocht hoe algoritmes bij de rijksoverheid in de praktijk functioneren en welke verbeteringen mogelijk zijn. Vragen over algoritmes – wat kunnen ze doen en welke risico's brengt het gebruik van algoritmes met zich mee – kunnen rekenen op (soms) uiteenlopende reacties, van extreem negatief tot extreem positief – en alles daar tussen in. Het toetsingskader dat wij hebben ontwikkeld, kan de basis zijn om algoritmes verantwoord in te zetten én kan het uitgangspunt voor discussies over de controle van en het toezicht op algoritmes.

Wij willen hiermee transparantie stimuleren en potentiële risico's over algoritmes bespreekbaar te maken. Transparantie en grip op algoritmes moeten de regel zijn, niet de uitzondering.

Een algoritme is geen black box



Onze hoofdconclusie op basis van de geïnventariseerde algoritmes is dat binnen de rijksoverheid veel aandacht is voor het beperken van de privacy risico's die een rol spelen bij algoritmes. Wij hebben vastgesteld dat automatische besluitvorming alleen plaatsvindt bij algoritmes die eenvoudige administratieve handelingen uitvoeren, zonder enige impact voor de burger. Wij hebben ook vastgesteld dat de complexe geïnventariseerde algoritmes niet zelf besluiten nemen, maar de uitvoerende functionarissen nadrukkelijk betrokken zijn bij het gebruik van deze algoritmes. Deze algoritmes ondersteunen de uitvoerende functionarissen bij het maken van analyses en bij het nemen van besluiten.

Daarnaast hebben wij vastgesteld dat algoritmes voor ons als onafhankelijk controleur geen *black box* zijn: wij hebben de algoritmes kunnen bekijken en beoordelen. Dat neemt niet weg dat er – anno 2021 – ruimte is voor verbetering, omdat het gebruik van algoritmes de komende jaren alleen maar zal toenemen. Wanneer algoritmes zelflerend en daardoor complexer worden kan aan snelheid, kwaliteit en objectiviteit van besluitvorming worden gewonnen. De uitvoerend functionaris komt daarmee ook op meer afstand te staan van besluiten die door de rijksoverheid over burgers en bedrijven worden genomen.
In dit hoofdstuk presenteren wij de deelconclusies die ons oordeel ondersteunen en voorzien we deze van aanbevelingen.

6.1 Een algoritme hoeft geen black box te zijn

Algoritmes zijn er ter ondersteuning van menselijk handelen. Binnen de inventarisatie toegepaste algoritmes binnen de rijksoverheid zijn er geen algoritmes aangetroffen die volledig autonoom handelen. Er zijn wel algoritmes die eenvoudige beslissingen nemen of routinehandelingen uitvoeren in een niet complexe omgeving. Voorbeelden hiervan zijn automatisch gegenereerde brieven en berichten. Bij het ontwikkelen van een algoritme hoort het maken van keuzes over de verklaarbaarheid en transparantie van dat algoritme. Ook het afleggen van verantwoording over een algoritme is een keuze. Door aan de voorkant van het proces hier actief op te sturen, wordt een algoritme geen *black box* maar een ondersteunend element in het werkproces. Een element waarvan duidelijk is welke data het gebruikt, hoe het model functioneert, welke uitkomsten dat oplevert en welke impact die uitkomsten hebben. Er zijn kansen om het algoritme beter controleerbaar te maken dan bij menselijke analyse mogelijk is. Aandachtspunt hierbij zijn algoritmes die van private partijen worden afgenomen. Die moeten aan dezelfde eisen voldoen als de algoritmes die de overheid zelf ontwikkelt.

6.2 Centraal inzicht ontbreekt, behoefte aan concrete instrumenten

De ontwikkeling van algoritmes vindt vaak bottom-up plaats, vanuit de praktijk van de uitvoering. De ambtelijke leiding van het ministerie en de Chief Information Officer (CIO) van het ministerie hebben weinig inzicht in dit proces. Het gevolg is dat de ministers de risico's en mogelijke (negatieve) impact van algoritmes op dienstverlening door de rijksoverheid niet tijdig kunnen beperken. Dit onderzoek heeft met de inventarisatie een bijdrage geleverd aan de ministers om een beter beeld te krijgen van het gebruik van algoritmes binnen hun ministerie. Daarnaast is er, als het over algoritmes gaat, geen eenduidige terminologie. Dit verklaart onze bevinding dat de functionarissen van ministeries op verschillende manieren invulling geven aan definities van algoritmes, de ontwikkeling van algoritmes, de risico's die daarbij horen en de beperking van die risico's.

De huidige toetsingskaders schieten tekort in het beoordelen van algoritmes. Voor de bevordering van de kwaliteit en betrouwbaarheid en de beperking van risico's die samenhangen met het gebruik van algoritmes, hanteren ministeries algemene normenkaders zoals AVG, BIO, ITIL[24] (*Information Technology Infrastructure Library*) en COBIT. Dat is echter niet bij alle ministeries het geval. Om risico's te beperken, gebruiken ministeries ook de Kamerbrieven over *big data*/algoritmes als handvat.

Medewerkers van slechts 3 ministeries geven expliciet in de inventarisatie aan dat zij ethische aspecten een belangrijk onderdeel vinden van een algoritme. Dit zien wij bevestigd in de praktijkcasussen waar het vaak ontbreekt aan maatregelen om vooroordelen te beperken (zoals de gekozen data of het risico van discriminatie) en ethische aspecten, zoals profilering. De algemene normenkaders zijn niet gespecificeerd voor algoritmes en worden ook niet in samenhang toegepast. Zonder een adequate sturing op en verantwoording van algoritmes, is het niet mogelijk om goede afwegingen te maken over de voor- en nadelen van de inzet van een algoritme. Daarnaast is het niet goed mogelijk de werking van een algoritme uit te leggen. De impact daarvan op burgers (discriminatie, onjuiste profilering, financiële consequenties) kan groot zijn.

Over een normenkader zijn de bij de ministeries betrokken functionarissen eensgezind. Er is behoefte aan bruikbare en heldere definities over algoritmes. Nu is er vaak sprake van interpretatieverschillen. Over de vraag of deze definities specifiek

of generiek moet zijn, lopen de meningen uiteen. Een aantal functionarissen ziet algoritmes als ICT, waarvoor dezelfde overkoepelende normen kunnen gelden. Tegelijkertijd geven functionarissen aan dat de risico's niet altijd generiek zijn, en één generiek normenkader daarom lastig is. De opbrengsten van de denksessie bevestigen onze bevindingen.

6.2.1 Aanbeveling 1: Uniformiteit en eenduidigheid van begrippen en kwaliteitseisen

Wij bevelen het kabinet aan om zorg te dragen voor een eenduidige gemeenschappelijk taal en concrete kwaliteitseisen voor algoritmes. Eenduidigheid en uniformiteit van begrippen en kwaliteitseisen zorgen voor: kennisuitwisseling, het stroomlijnen van processen en het voorkomen van misinterpretaties. Tijdens dit onderzoek hebben deelnemers aan een denksessie deze behoefte aan eenduidigheid en uniformiteit van begrippen binnen de rijksoverheid verder ingevuld en daarmee een basis gelegd voor een 'gemeenschappelijk taal' op het gebied van algoritmes. Het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, het Ministerie van Veiligheid en Justitie en Agentschap Telecom van het Ministerie van Economische Zaken en Klimaat hebben samen met ons deze denksessie georganiseerd om hun voortrekkersrol binnen de rijksoverheid als het gaat om algoritmes verder te kunnen vervullen in de vorm van eenduidige en breed toepasbare richtlijnen en kwaliteitseisen.

6.3 Voorspellende en voorschrijvende algoritmes volop in ontwikkeling met nu nog beperkte impact burger

Op basis van onze inventarisatie concluderen wij dat algoritmes rijksbreed worden ingezet. Het gaat hierbij om zowel eenvoudige als complexe algoritmes. Op hoofdlijnen zijn er 3 verschillende processen/doelen waarvoor algoritmes worden ingezet:

- bij het automatiseren van administratie en eenvoudige wetgeving;
- voor het faciliteren en verbeteren van bedrijfsvoering en/of dienstverlening;
- voor risicogerichte controle en daarmee ook voor gerichte inzet van beschikbare mensen en middelen.

We hebben geen volledig zelflerende algoritmes aangetroffen binnen de rijksoverheid; alleen lerende algoritmes. Automatische besluitvorming vindt alleen plaats bij algoritmes die eenvoudige administratieve handelingen uitvoeren, zonder substantiële impact voor de burger.

6.4 De burger staat onvoldoende centraal

Op dit moment zijn *Data Protection Impact Assessments*[25] (DPIA's), *privacy statements* en verwerkingsregisters onvoldoende toegankelijk en onvoldoende begrijpelijk voor niet-deskundigen en leken. Burgers weten niet waar zij terecht kunnen als ze vragen hebben over algoritmes, waar ze fouten in data kunnen aangeven of hoe ze bezwaar kunnen maken tegen gebruik van data of uitkomsten van algoritmes. Het volstaat naar onze mening niet om alleen formeel te voldoen aan de eisen van de AVG; dat geeft burgers meestal niet voldoende zicht op de algoritmes die hen raken. De rijksoverheid kan voorkomen dat er vooroordelen over algoritmes ontstaan door transparant te communiceren over de inzet van algoritmes, over de effecten die de burger van die algoritmes kan ondervinden en over haar aanspreekbaarheid.

6.4.1 Aanbeveling 2: Geef de burgers inzicht in de toepassing van het algoritme en geef aan waar zij terechtkunnen als ze vragen hebben

Wij bevelen het kabinet aan om burgers op een logische plek inzicht te geven in welke data worden gebruikt in welke algoritmes, hoe die algoritmes op hoofdlijnen functioneren en welke impact de uitkomsten daarvan hebben. Het gaat hierbij om algoritmes die een substantiële impact hebben op overheidshandelen of op beslissingen over een concrete casus, burger of een bedrijf. Er kan bijvoorbeeld gedacht worden aan een dashboard zoals dat ook is ingericht om inzicht te verschaffen in grote ICT-projecten.

6.5 Verbeterpunten voor een verantwoorde inzet en doorontwikkeling van algoritmes

Sturing en verantwoording

Wij signaleren dat de afspraken, rollen, taken en verantwoordelijkheden van partijen die bij de rijksoverheid betrokken zijn bij het gebruik van algoritmes beter belegd moeten worden. Dit is nodig om als ministerie systematisch inzichtelijk te hebben of het algoritme doet wat het moet doen. Zeker in de gevallen waarin meerdere partijen betrokken zijn bij de ontwikkeling, de werking en de instandhouding van het algoritme. We vragen aandacht voor de kwaliteit van het testen van de algoritmes en de continue monitoring door het ministerie.

Wij stelden vast dat er veelal geen *lifecycle management* is ingericht voor algoritmes. Er gaat veel aandacht naar het ontwerp en de ingebruikname van het algoritme, maar niet naar de instandhouding en het onderhoud ervan. Dit leidt ertoe dat het algoritme in de loop van de tijd bijvoorbeeld niet meer voldoet aan de veranderde ethische principes, wet- en regelgeving of gewoon technisch out-of-date is.

6.5.1 Aanbeveling 3: Leg afspraken omtrent de inzet van algoritmes vast en richt de continue monitoring goed in

Wij bevelen het kabinet aan te zorgen voor een goede vastlegging van de uitgangspunten, organisatie, monitoring (bijvoorbeeld *lifecycle*: onderhoud en actualiteit wet- en regelgeving) en evaluatie van het algoritme, zodat inzichtelijk wordt en blijft in hoeverre het algoritme voldoet aan de doelstellingen. Dat maakt het mogelijk om, als dat nodig is, het algoritme bij te sturen. Zeker bij uitbesteding of inkoop bij een andere (externe) partij is het van belang afspraken vast te leggen over aansprakelijkheid. Ons toetsingskader biedt belangrijke aandachtspunten als input voor deze vastlegging.

Model en data

De rijksoverheid zet algoritmes in op verschillende terreinen. Zo zien we eenvoudige beslisbomen, maar ook ingewikkelde algoritmes voor beeldanalyse. Niet alle beoordelingsaspecten uit ons toetsingskader zijn dan ook op alle algoritmes van toepassing. Ook de context speelt een belangrijke rol bij de weging van de bevindingen over een algoritme. In het ene geval kan uitlegbaarheid belangrijk zijn om burgers inzicht te geven, terwijl diezelfde uitlegbaarheid in een ander geval ongewenst is omdat dit beslissers teveel beïnvloedt. Bovendien kan transparantie ontaarden in een handleiding frauderen voor de burger. Ons toetsingskader kan voor elk algoritme worden doorontwikkeld tot een normenkader of een set van minimale kwaliteitseisen.

De aandachtspunten bij model en data gaan zowel over de techniek van het ontwikkelen van modellen voor het algoritme, als over de kwaliteit van de data. Als het gaat om techniek is de deskundigheid van de medewerkers vaak in orde. Wij vragen aandacht voor 2 potentiële risico's bij het datamanagement. Het eerste risico is dat het gebruik van historische data bij een algoritme niet aansluit wanneer maatschappelijke vernieuwingen plaatsvinden. De historische data wordt dan nog steeds op de huidige situatie toegepast. Het tweede risico betreft de vooroordelen in de data. Indien bepaalde bevolkingsgroepen afwijkend zijn behandeld in het verleden neemt het algoritme deze discriminatie over.

Uit de inventarisatie en de 3 casussen concluderen wij dat al bij de ontwikkeling van algoritmes niet alle relevante kennisdisciplines betrokken zijn. Wanneer juridische en ethische kennisdisciplines niet betrokken zijn kan dat ertoe leiden dat een algoritme niet voldoet aan alle wet- en regelgeving, ethische principes of dat het algoritme niet bijdraagt aan het beleidsdoel. Ook ontbreken vaak maatregelen om vooroordelen te beperken (zoals de gekozen data of het risico van discriminatie) en ethische aspecten.

6.5.2 Aanbeveling 4: Draag zorg voor een vertaling van het toetsingskader naar hanteerbare kwaliteitseisen voor algoritmes

Wij bevelen het kabinet aan om via de minister van Binnenlandse Zaken en Koninkrijksrelaties, de *Chief Information Officer* op een ministerie verantwoordelijk te stellen voor de vertaling van het toetsingskader (controle achteraf) naar een hanteerbaar normenkader aan de voorkant of naar kwaliteitseisen voor de ontwikkeling van algoritmes. Dit heeft als doel dat de bruikbaarheid van de kwaliteitseisen omhoog gaat en al aan de voorkant, bij de ontwikkeling van algoritmes, kan worden toegepast.

6.5.3 Aanbeveling 5: Betrek meerdere disciplines al bij ontwikkeling van algoritmes

Wij bevelen het kabinet aan om bij de ontwikkeling van algoritmes alle relevante disciplines of soorten kennis te betrekken. Dit betekent dat naast de technici, ook juristen, ethici en beleidsadviseurs betrokken zouden moeten worden.

Privacy

De burger heeft onvoldoende zicht op de waarborgen voor de privacy bij het gebruik van algoritmes. Dit vertaalt zich concreet in de volgende aandachtspunten:

- Voldoen aan de formele vereisten van de AVG is niet voldoende om de burger inzicht te geven in de werking, de gebruikte data en de impact van een algoritme.
- Het online verwerkingsregister van de rijksoverheid (www.avgregisterrijksoverheid.nl) wekt de verwachting dat daar alle verwerkingsregisters te vinden zijn. Dit is echter niet het geval. Er is ook geen wettelijke plicht om verwerkingsregisters op deze website te publiceren.
- De aanbeveling voor privacy is opgenomen bij § 6.4.1.

***IT General Controls* (ITGC)**

Wanneer het beheer van het algoritme is uitbesteed aan een externe partij, concluderen wij dat op departementaal niveau niet bekend is of de ITGC-normen

voldoende zijn toegepast. Op zichzelf is dat geen probleem, maar zoals het nu is ingericht bij de beoordeelde algoritmes, zien wij risico's.
Ministeries die de ontwikkeling en het beheer van algoritmes hebben uitbesteed, hebben namelijk beperkt zicht op deze algoritmes. Het opdrachtgevend ministerie gaat ervan uit dat deze partij *in control* is, en voldoet aan de (ITGC-)standaarden die wij toetsen. Wij missen hiervoor het bewijs: de verantwoordelijk minister verwijst naar de externe partij, maar heeft zelf geen zicht op de kwaliteit of op documenten over de naleving van deze standaarden.

Bij een ministerie dat het beheer heeft uitbesteed aan een SSO, zien we hetzelfde beeld als bij de externe private partij. De afdeling die het algoritme inzet, verwijst naar de ITGC-richtlijnen op een hoger/ander niveau binnen de organisatie. De verantwoordelijkheid wordt dus verlegd maar de functionarissen, van het ministerie die het algoritme inzet, kunnen niet toelichten op welke manier de organisatiebrede kaders doorwerken naar het specifieke algoritme.

6.5.4. Aanbeveling 6: Draag zorg voor het inzicht hebben en houden in het functioneren van de IT General Controls

Wij bevelen het kabinet aan via de minister van Binnenlandse Zaken en Koninkrijksrelaties, dat de verantwoordelijke bewindslieden ervoor zorgen dat de betrokken functionarissen die werken met het algoritme, inzicht hebben en houden in de kwaliteit van het functioneren van de ITGC van dat specifieke algoritme. Dit kan door bij de beherende partij te vragen om officiële verklaringen, zoals rapportages van IT auditors, waaruit blijkt dat de ITGC voldoende functioneren.

Ethiek

De waarneming die wij tijdens ons onderzoek hebben gedaan is dat de wetgeving soms op gespannen voet staat met de ethische principes. Voor de toetsing op de naleving van de ethische principes fairness en verklaarbaarheid en transparantie moet de onafhankelijk controleur kunnen vaststellen welke data gehanteerd is en deze data kunnen testen. De privacywetgeving zorgt ervoor dat veel data niet lang bewaard wordt waardoor de data voor de auditor achteraf niet controleerbaar is.
De behoefte bij onafhankelijk controleurs aan een wetswijzing op het gebied van privacy voor complexe algoritmes is er al, maar neemt naar verwachting toe wanneer de algoritmes complexer worden. Dit zal de komende jaren blijken op basis van de ontwikkeling van algoritmes.

7.
Reactie en nawoord

De staatssecretaris van Binnenlandse Zaken en Koninkrijksrelaties (BZK) heeft vanuit zijn coördinerende verantwoordelijkheid met betrekking tot ICT binnen de Rijksdienst op 22 december 2020 gereageerd op ons rapport mede namens zijn collega's.

7.1 Reactie staatssecretaris van BZK

In zijn reactie geeft de staatssecretaris van BZK (hierna: de staatssecretaris) aan onze conclusies te herkennen en te waarderen. Onze aanbevelingen ervaart hij als constructief. Hieronder geven we de hoofdlijn van zijn reactie op onze aanbevelingen weer. De volledige reactie is te vinden op www.rekenkamer.nl. We sluiten af met ons nawoord.

Reactie op aanbevelingen

Uw aanbevelingen dragen bij aan verbeterde dienstverlening naar de mensen waar de overheid voor werkt en de daarvoor ingerichte werkprocessen.

1. "Uniformiteit en eenduidigheid van begrippen en kwaliteitseisen."
Aan een eenduidige gemeenschappelijk taal en concrete kwaliteitseisen voor algoritmes wordt gewerkt via onder meer de kennisbundeling en het gestructureerde overleg van de Nederlandse Digitaliseringstrategie (NDS). Een verkenning is uitgevoerd, in samenspraak met de minister voor Rechtsbescherming en de

staatssecretaris van EZK, die onder meer heeft gekeken naar het voorkomen van fragmentatie, de normering van toezicht en het betrekken van publieke en private kennis. De Tweede Kamer is ten tijde van uw onderzoek geïnformeerd[26] en heeft de resultaten van de verkenning besproken met meerdere bewindspersonen.
In deze actie is het zaak een goede balans te zoeken in meerwaarde van rijksbrede eenduidigheid t.o.v. specifieke invullingen per departement en uitvoeringsorganisatie.

2. “Geef de burger inzicht in de toepassing van het algoritme en geef aan waar zij terechtkunnen als ze vragen hebben.”
De opgestelde richtlijnen voor het gebruik van algoritmes door overheden worden verder aangescherpt en geëvalueerd. Daarnaast wordt een model ontwikkeld voor een impact assessment met betrekking tot algoritmes en mensenrechten. Zowel nationale als Europese wetgeving geeft inzicht in de toegepaste voorspellende of voorschrijvende algoritmes.

In uw rapport wordt Syri als voorbeeld genoemd. Op pagina 5 van dit rapport wordt aangegeven dat het systeem Syri binnen de overheid (door het UWV en de Belastingdienst) werd gebruikt om fraude op te sporen met algoritmes. Omdat deze passage de indruk wekt van een breed en generiek gebruik van Syri in het reguliere toezicht, wordt eraan gehecht de context en het gebruik van Syri te verduidelijken. Syri is een systeem voor vergelijking van gegevensbestanden van verschillende overheidsorganisaties (zowel centraal als decentraal) op grond van de Wet Suwi, dat is ingezet in een beperkt aantal specifieke samenwerkingsprojecten op het terrein van voorkoming en terugdringing van belasting- en sociale zekerheidsfraude, overtredingen van arbeidswetgeving en daarmee samenhangende misstanden. Afgelopen 5 februari heeft de rechter zich uitgesproken over gebruik van Syri waaruit bleek dat de privacy van burgers onvoldoende gewaarborgd was. Daarop heeft de overheid het gebruik van Syri meteen stopgezet.

3. “Leg afspraken omtrent de inzet van algoritmes vast en richt de continue monitoring goed in.”
Afspraken omtrent de inzet en monitoring van algoritmes zijn door de intensieve samenwerking van diverse departementen verder vormgegeven. Concreet resultaat hiervan zijn onder meer het Strategisch actieplan voor kunstmatige intelligentie, een beleidsbrief publieke waarden en waarborgen tegen data-analyses door de overheid.

4. “Draag zorg voor een vertaling van het toetsingskader naar hanteerbare kwaliteitseisen voor algoritmes.”

Het kabinet werkt met de Algemene Rekenkamer en de Auditdienst Rijk aan een vertaling van het toetsingskader naar hanteerbare kwaliteitseisen voor algoritmes. Bij artificiële intelligentie (AI) algoritmes moet ook de betrouwbaarheid en kwaliteit van data worden meegenomen, omdat AI-algoritmes gebruik maken van data. Dit borgt het kabinet met een aanpak via een agenda en projectgroep om de input van de verkenning en het rekenkameronderzoek, zowel bij de ontwikkeling als de uitwerking, mee te nemen en een breed gedragen agenda op het terrein van normering en toezicht op algoritmen te realiseren. Een nadere analyse is hierbij nodig naar de mogelijke gevolgen voor administratieve belasting en uitvoerbaarheid bij (decentrale) uitvoerders en de samenhangende benodigde implementatietijd met respect voor bestaande structuren en de autonomie van de departementen en uitvoeringsorganisaties.

5. “Betrek meerdere disciplines al bij ontwikkeling van algoritmes.”

Het betrekken van meerdere disciplines is de norm bij de ontwikkeling van beleid, wetgeving, uitvoerende processen en daaruit volgende algoritmes en het toezicht daarop, conform het integraal afwegingskader. De basis voor de in te zetten instrumenten, algoritmes of anderszins, wordt bij wet bepaald en dient te werken binnen het afgesproken kader. Een ideale mix van disciplines voor dergelijke (ontwikkel)processen blijft echter maatwerk, afhankelijk van beschikbare capaciteit, middelen of tijd.

6. “Draag zorg voor het inzicht hebben en houden in het functioneren van de IT General Controls.”

Voor de werking van reguliere systemen als ook algoritmes is het functioneren van de IT General Controls van belang. Mogelijke additionele rapportagewensen of audit verklaringen dragen bij aan inzicht en controle bij de verantwoordelijke opdrachtgever en eigenaar. Met parallelle acties wordt ook het inzicht vanuit de CIO-kolom versterkt. De dynamiek vraagt echter ook om de balans in ogenschouw te houden tussen controle instrumenten en de organisatorische of administratieve belasting die hierbij gepaard gaat.

Met het door u ontwikkelde toetsingskader heeft u een drietal algoritmes in de praktijk getoetst. De bevindingen zijn veralgemeniseerd. Ik wil daarbij als kant-tekening plaatsen dat daarmee voorbeelden van kwalitatief hoogwaardig gebruik van algoritmes binnen de rijksoverheid, waarbij ook ethische voorwaarden als organiserend principe worden gehanteerd, minder in het rapport tot uiting komen.

Naast het belang van bescherming van de rechten en vrijheden van burgers zal, mede vanuit het oogpunt van toezicht, nader inzicht moeten worden verkregen in de experimenteerruimte voor departementen en uitvoeringsorganisaties op het gebied van de inzet van algoritmen, waarbij zowel de (uitvoerings)praktijk als toezicht-houdende organisaties van elkaar kunnen leren.

Zoals u ook aangeeft in uw rapport is de samenwerking geïntensiveerd, mede dankzij de denksessie van 22 september 2020 waarbij de departementen en externe experts deelnamen met de rekenkamer en de ADR.

Ik dank uw rekenkamer voor het onderzoek en de bijdrage om hiermee het begrip van en over algoritmes te vergroten. Daarmee zal verder gewerkt worden aan verbeterde dienstverlening en beleidsuitvoering vanuit de rijksoverheid."

7.2 Nawoord Algemene Rekenkamer

Wij waarderen de reactie van de staatssecretaris.

De staatssecretaris geeft bij onze aanbeveling over uniformiteit en eenduidigheid van begrippen en kwaliteitseisen (aanbeveling 1) aan dat het belangrijk is een goede balans te zoeken in meerwaarde van rijksbrede eenduidigheid ten opzichte van specifieke invullingen per departement en uitvoeringsorganisatie.
Wij signaleren naar aanleiding van ons onderzoek bij vrijwel alle departementen een significante behoefte bij de met ontwikkeling en uitvoering van algoritmes belaste medewerkers naar meer eenduidigheid van te hanteren begrippen, richtlijnen en standaarden voor kwaliteit. De medewerkers bij het Ministerie van BZK (CIO-Rijk) kunnen hierin een belangrijke rol vervullen. Eenduidigheid bevordert kennisdeling en consistente kwaliteit rijksbreed; zo kan juist meer ruimte voor specifieke invulling ontstaan. Ook de samenwerking tussen departementen op het gebied van afspraken over inzet en monitoring van algoritmes (zie aanbeveling 3) zal hierbij gebaat zijn. Het is van groot belang dat deze afspraken consistent, controleerbaar en niet vrijblijvend zijn.

Wij merken op dat ons toetsingskader risico's bevat voor alle soorten algoritmes ongeacht de context van het ministerie of daaraan verbonden organisaties. We gaan er van uit dat de inzichten uit ons onderzoek worden gebruikt voor een verantwoorde inzet van algoritmes bij alle ministeries.

De door de staatsecretaris genoemde impactassessment en regelgeving (zie aanbeveling 2) vormen ongetwijfeld een goede basis: wij pleiten ervoor bij het geven van meer inzicht aan de burger over het gebruik van algoritmes ook vroegtijdig en voldoende aandacht te besteden aan de praktische uitvoering hiervan en de burger hier actief over de mogelijkheden te informeren.

Wij waarderen de aandacht die de staatssecretaris heeft voor de vertaling van het toetsingskader naar hanteerbare kwaliteitseisen voor algoritmes (zie aanbeveling 4). De betrouwbaarheid en kwaliteit van data luistert bij alle vormen van algoritmes nauw. Inderdaad verdienen AI algoritmes additionele aandacht, omdat zij van zeer veel data gebruik maken maar ook omdat niet altijd navolgbaar is hoe deze verwerkt worden en hoe de uitkomsten tot stand komen.

Het is goed dat interdisciplinaire samenwerking ook in de life cycle van algoritmes door de staatssecretaris als norm wordt gesteld (zie aanbeveling 5); wij hebben echter in ons onderzoek gezien dat dat in de praktijk nog onvoldoende wordt toegepast. Vanzelfsprekend kunnen economische of praktische overwegingen een rol spelen; wij onderstrepen echter het voordelige effect van een interdisciplinaire aanpak op de mitigatie van risico's, zeker op de wat langere termijn.

Aandacht voor de IT general controls (zie aanbeveling 6) zal zeker zijn weerslag vinden in allerlei auditrapportages en verklaringen; waar het ons echter met name om gaat is de mate en de kwaliteit van de toegangs- en data beveiliging, het life cycle management en de continuïteitsmaatregelen, ook voor algoritmes, die door de eigenaar/beheerder moeten worden gewaarborgd.

De staatsecretaris onderschrijft het belang van bescherming van de rechten en vrijheden van burgers in relatie tot de inzet van algoritmes. Hij legt terecht een verband met de wens om nader inzicht te krijgen in de experimenteerruimte voor departementen en uitvoeringsorganisaties op het gebied van de inzet van algoritmes, dat mogelijk voor dilemma's kan zorgen. Wij begrijpen deze overweging. Uitgangspunt moet wel zijn dat het een het ander niet kan en mag bijten en omgekeerd. Bovendien is de mate en de manier van informatieverschaffing, ook over de aard en de reikwijdte van experimentele algoritmes van groot belang. In lijn met het punt van de kwaliteitseisen willen wij erop wijzen dat ons toetsingskader risico's bevat voor alle soorten algoritmes in elke fase van ontwikkeling of toepassing.

De Algemene Rekenkamer hoopt met dit onderzoek een bijdrage te leveren aan het wegnemen van begrijpelijke zorgen van burgers over mogelijk oncontroleerbare en besluitvormende algoritmes bij de rijksoverheid. Uit ons onderzoek blijkt op dit moment dat slechts zeer beperkt sprake is van besluitvormende algoritmes en de door ons onderzochte algoritmes bleken wel degelijk controleerbaar. Dat hoeft niet te betekenen dat er geen terechte zorgen zijn. Er is sprake van een snelle ontwikkeling, we hebben de volledigheid van de aanlevering door de ministeries niet onderzocht. Niettemin geldt voor algoritmes dat zij niet alleen door menselijk handelen gemaakt zijn maar ook controleerbaar zijn door menselijk handelen voor het effect op de burger. Zo moet het ook zijn.

Wij zullen de voortgang op alle aanbevelingen volgen en de komende jaren in onze onderzoeksactiviteiten aandacht besteden aan het onderwerp algoritmes.

Bijlagen

Bijlage 1
Methodologische verantwoording

Inzicht in algoritmes

In dit onderzoek stonden de volgende onderzoeksvragen centraal:

1. Voor welke activiteiten en processen worden algoritmes toegepast bij de rijksoverheid en bij organisaties die aan de rijksoverheid zijn verbonden, welke typen/categorieën zijn er te onderscheiden en wat zijn de effecten en risico's?
2. Hoe is de besturing/*governance* en kwaliteitsbeheersing van algoritmes bij de rijksoverheid vormgegeven, en bij organisaties die aan de rijksoverheid zijn verbonden?

De beantwoording van deze vragen heeft plaatsgevonden door middel van een inventarisatie binnen de rijksoverheid naar soorten algoritmes en bij welke activiteiten deze zijn ingezet. Wij hebben gevraagd om voorschrijvende en voorspellende algoritmes aan te leveren met relevante impact op werkprocessen en/of dienstverlening van de overheid. Daarbij heeft de Algemene Rekenkamer gevraagd om de meest representatieve algoritmes aan te leveren. In de vragenlijst was ruimte voor 10 algoritmes. Het was mogelijk om af te wijken van dit aantal.

De indeling van algoritmes komt voort uit de indeling/typering zoals beschreven in de bijlage bij de Kamerbrief over waarborgen tegen risico's van data-analyses door de overheid.[27] Daarin wordt onder meer onderscheid gemaakt in:

- complexiteit van het algoritme (laag-hoog);
- technische transparantie (laag-hoog).

De bijlage onderscheidt ook hoe algoritmes worden ingezet, waarbij ook een onderscheid gemaakt kan worden met betrekking tot impact. De impact is klein bij beschrijvend en is het grootst bij voorschrijvend:

- beschrijvend;
- diagnostisch;
- voorspellend;
- voorschrijvend.

Gezien de focus van ons onderzoek op substantiële impact hebben wij de keuze gemaakt om voorspellende en voorschrijvende algoritmes te gaan inventariseren. We hebben in dit onderzoek nadrukkelijk niet gestreefd naar een volledige inventarisatie van alle algoritmes bij de rijksoverheid. Wij hebben de ministeries gevraagd zelf te rapporteren over de ingezette algoritmes die zij binnen onze afbakening vonden passen. In een aanvullend gesprek hebben we verdiepingsvragen gesteld. De gesprekken zijn vastgelegd in verslagen en afgestemd met de geïnterviewden.

Toetsingskader en praktijktoets van algoritmes
Wij hebben een toetsingskader ontwikkeld op basis van bestaande normenkaders en richtlijnen inclusief relevante literatuur (zie bijlage 2). Het toetsingskader is aangescherpt met de uitkomsten van de denksessie en de praktijktoets. In die praktijktoets hebben we 3 algoritmes beoordeeld.

Denksessie

Er zijn tal van richtlijnen en kaders die ingaan op deelaspecten van een algoritme. Bijvoorbeeld de AVG (privacy) of de BIO (informatiebeveiliging). Er is echter geen integraal toetsingskader voor algoritmes. Bovendien is er geen gemeenschappelijke taal als het gaat om algoritmes. In onze inventarisatie geven functionarissen aan dat ze behoefte hebben aan meer eenduidigheid in gebruikte definities en terminologie. Wat verstaan we eigenlijk precies onder een algoritme? Wat is uitlegbaarheid? En voor wie moet een algoritme uitlegbaar zijn? En wat bedoelen we met transparantie?

Om aan deze behoeften tegemoet te komen, organiseerden wij op 22 september 2020 een denksessie in samenwerking met het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, het Ministerie van Veiligheid en Justitie en Agentschap Telecom van het Ministerie van Economische Zaken en Klimaat. Deze partijen vervullen een voortrekkersrol binnen de rijksoverheid als het gaat om algoritmes. Het doel was om gezamenlijk tot meer eenduidigheid te komen over de terminologieën die gebruikt worden bij algoritmes. Meer eenduidigheid draagt ook bij aan de

totstandkoming van kaders voor de toepassing van algoritmes in de praktijk, en aan de gewenste manier om daarover verantwoording af te leggen. Door vanuit verschillende perspectieven (onder andere juridisch, technisch, beleidsmatig, wetenschappelijk en toezichthoudend) het gesprek te voeren over de kenmerken en definities van algoritmes, en te leren van elkaars ervaringen, kunnen we het denken over en het begrip van algoritmes een stap verder brengen. Aan de sessie namen 30 experts van binnen en buiten de overheid deel. Hieronder de terugkoppeling van de denksessie.

Verslag denksessie

Doel

Er zijn tal van richtlijnen en kaders die ingaan op deelaspecten van een algoritme. Bijvoorbeeld de AVG (Algemene verordening gegevensbescherming) voor privacy of de BIO (Baseline Informatiebeveiliging Overheid) voor informatiebeveiliging. Er is echter nog geen integraal toetsingskader voor algoritmes. Daarnaast spreekt niet iedereen dezelfde 'taal' als het over algoritmes gaat. Waar hebben we het eigenlijk over, als we over een algoritme spreken? *Wat is uitlegbaarheid, en wat bedoelen we met transparantie? Waar zit het verschil, en voor wie moet een algoritme uitlegbaar zijn? En wat bedoelen we met bias? Was er niet altijd al bias? En wordt dat lastiger nu de bias in het algoritme zit en niet in de mens?*

Om met elkaar te komen tot meer eenduidigheid over de terminologie rondom algoritmes, organiseerde het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, het Ministerie van Veiligheid en Justitie, het Agentschap Telecom van het Ministerie van Economische Zaken en Klimaat en de Algemene Rekenkamer op 22 september 2020 de denksessie algoritmische data-analyses.

Terugblik & opbrengsten

Aan de denksessie namen ruim 30 experts van binnen en buiten de overheid deel. Vanuit hun verschillende rollen, achtergronden en/of expertise (o.a. juridisch, technisch, beleidsmatig, wetenschappelijk en toezichthoudend) spraken zij met elkaar over vijf thema's: datagedreven werken, datakwaliteit, AI en algoritmes, AI bij de overheid en transparantie. Wat viel op?

- Zowel binnen de rijksoverheid als daarbuiten wordt al jarenlang gebruik gemaakt van algoritmes, maar door de negatieve aandacht wordt dit als 'eng' ervaren. Het is wenselijk dat de techniek rondom algoritmes beter wordt begrepen en er meer inzicht wordt geboden in het gebruik van algoritmes zodat de 'hype' er af gaat (demystificatie);
- De behoefte aan een toetsingskader of richtlijn komt voort uit de wens om binnen de Rijksoverheid een gemeenschappelijke basis te creëren met meer concretisering;

- Voor een goed begrip van het algoritme moet zowel naar de context als het doel worden gekeken;
- De bovengenoemde thema's zijn veelal te ruim en abstract. Het is niet realistisch om tot één definitie te komen, maar om deze op te delen in de belangrijkste aspecten en die te concretiseren;
- Algoritme en data vragen om een overheidsbrede sturing. Omdat die overheid ook zelf steeds meer in netwerken van overheidsorganisaties opereert.
- Bij uitbesteding van algoritmes of onderdelen daarvan zitten overheidsorganisaties er niet altijd dicht genoeg op en is er in een aantal gevallen minder grip dan bij algoritmes in eigen beheer;
- Ook al is het algoritme 'op papier' perfect, de mens maakt het algoritme en bepaalt welke data het algoritme gebruikt. Er kan nooit met 100% zekerheid worden gesteld dat modellen (of mensen) geen bias hebben en daarmee niet discrimineren. Het lijkt belangrijk dat ook de politiek zich dat realiseert voor het nemen van effectieve en haalbare beheersmaatregelen.
- We zouden binnen de rijksoverheid (minimale) eisen moeten stellen aan de toepassing van algoritmes zodat hier op een verantwoorde wijze mee wordt omgegaan.

Bijlage 2
Literatuurlijst en bronnen toetsingskader

Deze literatuurbijlage bevat een selectie van de belangrijkste bronnen maar is niet compleet gezien de grote hoeveelheid publicaties op dit vlak.

Parlementaire stukken

- EZK (2019), *Kamerbrief van Minister van EZK over Strategische Actieplan voor Artificiële Intelligentie*, Tweede Kamer, 8 oktober 2019, kamerstuk 26 643, nr. 640
- Kabinet (2019), *Strategisch actieplan Artificiële Intelligentie,* 8 oktober 2019
- J&V (2019), *Kamerbrief van Minister van J&V over Waarborgen tegen risico's van data-analyses door de overheid,* 8 oktober 2019, kamerstuk 26 643, nr. 641
- BZK (2019), *Kamerbrief van Minister van BZK over AI, publieke waarden en mensenrechten,* Tweede Kamer, 8 oktober 2019, kamerstuk 26 643, nr. 642
- Kabinet (2020), *Kabinetsreactie op het onderzoek 'Toezicht op het gebruik van algoritmen door de overheid'* Datum 20 april 2020, bijlage bij kamerstuk 35 212, nr. 3

Nationaal

- Auditdienst Rijk (2018), GITC kader gebaseerd op BIR 2017
- BIR 2017, *Baseline Informatiebeveiliging Rijksdienst,* volledig gebaseerd op de internationale norm ISO/IEC 27002
- BIO, *Baseline Informatiebeveiliging Overheid* per 1 januari 2019 (gepubliceerd in staatscourant 23 mei 2019), vervangt BIR 2017
- Gemeente Amsterdam, *Modelbepalingen voor gemeenten voor verantwoord gebruik van Algoritmische toepassingen,* https://www.amsterdam.nl/wonen-leefomgeving/innovatie/de-digitale-stad/grip-op-algoritmes/
- Hooghiemstra & Partners (2019), *Onderzoek Toezicht op het gebruik van algoritmen door de overheid,* Hooghiemstra & Partners
- Waag (2020), *Algoritme: de mens in de machine,* Waag
- Frans van Bruggen/Joep Beckers (2020), *Nut en noodzaak van toezicht op artificiële intelligentie,* Tijdschrift voor Toezicht
- Montaigne Centrum voor Rechtsstaat en Rechtspleging, Universiteit Utrecht (2020), *Juridische aspecten van algoritmen die besluiten nemen,* Een verkennend onderzoek, Montaigne Centrum
- Dialogic (2020), *Gebruik van en toezicht op AI-toepassingen in telecominfrastructuren, Advies aan de toezichthouder over inrichting van risico gebaseerd AI-toezicht,* Agentschap Telecom

EU & Internationaal

- High-level expert group on artificial intelligence set up by European commission (2019), *Ethics guidelines for trustworthy AI*, Europese Commissie
- Michael Veale (2019), *A Critical Take on the Policy Recommendations of the EU High-Level Expert Group on Artificial Intelligence,* Faculty of Laws, University College London and the Alan Turing Institute
- National Audit Office UK (2016), *Framework to review models*
- Anna Jobin/Marcello Ienca/Effy Vayena (2019), *Artificial Intelligence: the global landscape of ethics guidelines,* Health Ethics & Policy Lab, ETH Zurich
- Thilo Hagendorff (2020), *The Ethics of AI Ethics: An Evaluation of Guidelines, Minds and Machines*
- Europese Commissie (2020), *Whitepaper on Artificial Intelligence – A European approach to excellence and trust,* Europese Commissie
- Daten Ethik Kommission (2019), *Opinion of the Data Ethics Commission, Daten Ethik Kommission*
- OECD (2019), *Recommendation of the Council on Artificial Intelligence,* OECD
- Geron, A. (2017), *Hands-On Machine Learning with Scikit-Learn and TensorFlow*
- Hastie, T., Tibshirani R, Friedman, F. (2009), *The Elements of Statistical Learning*
- Thomas L.C., Oliver R.W., and Hand D.J. (2005), *A survey of the issues in consumer credit modelling research,* Journal of the Operational Research Society, 56, 1006-1015
- ISACA (2018), *Auditing Artificial Intelligence,* ISACA
- ISACA (2012), COBIT 5, *A Business Framework for the Governance and Management of Enterprise IT*, ISACA
- ISACA (2012), COBIT 5, *Enabling Processes,* ISACA

Online bronnen

Kennisbank openbaar bestuur, *Artificiële Intelligentie en publieke waarden*, https://kennisopenbaarbestuur.nl/thema/artifici%C3%ABle-intelligentie-en-publieke-waarden

Bijlage 3
Toetsingskader algoritmes

Het toetsingskader algoritmes, ontwikkeld in het kader van ons onderzoek Aandacht voor Algoritmes, is een praktisch handvat om de belangrijkste risico's voor de rijksoverheid met betrekking tot dit onderwerp te beheersen. Er is gebruik gemaakt van bestaande (normen)kaders, richtlijnen, wet- en regelgeving en kent 5 perspectieven:

1. sturing en verantwoording;
2. model en data;
3. privacy;
4. IT General Controls (ITGC);
5. ethiek.

Voor elk perspectief zijn de belangrijkste risico's geformuleerd. Wij koppelen de te toetsen aspecten en onderzoeksvragen aan die risico's. Met de beantwoording van alle vragen en het geven van een score ontstaat het beeld over de mate waarin risico's beperkt worden bij het gekozen algoritme. Hoe groot de risico's voor een specifiek algoritme zijn, hangt af van de mate waarin geavanceerde technieken worden toegepast en de impact van het algoritme op de burger.
Voordat het kader gebruikt wordt, is de beantwoording van een aantal algemene vragen nodig. De informatie voor de beantwoording van deze vragen geeft een algemeen beeld en context van het algoritme. Deze context en algemene informatie bepalen de selectie van de relevante vragen in het toetsingskader voor het te toetsen algoritme.

Algemene vragen

1. Wat is de naamgeving van het algoritme of het systeem waar het algoritme deel van uit maakt?
2. In welk werkproces of ten behoeve van welk product of dienst speelt dit algoritme een rol?
3. Maakt het algoritme gebruik van persoonsgegevens (AVG)?
4. Is er sprake van een lerend algoritme, dat wil zeggen een algoritme dat ontwikkelt en verbetert in de loop der tijd door gebruik te maken van data en/of ervaringen?
5. Is het algoritme adviserend/ondersteunend ten behoeve van acties/besluit-vorming door mensen of handelt het autonoom/automatisch zonder menselijke tussenkomst?
6. Van welke technologie maakt het algoritme gebruik van en/of welke applicatie/software?
7. Welke data(bronnen) gebruikt het algoritme?

Het toetsingskader

Sturing en verantwoording		
Risico	**Onderzoeksvraag**	**Ethisch principe[28]**
Zonder eenduidigheid over het doel is geen sturing op en verantwoording over het algoritme mogelijk.	Is er een doel van het algoritme en vastgesteld?	4.2
Zonder actueel beeld van risico 's kan er geen goede afweging worden gemaakt of de voordelen van de toepassing van het algoritme opwegen tegen de nadelen.	Is er een vastgelegde (periodieke) afweging (bij start een businesscase) omtrent de risico's (beheersing) van het gebruik van het algoritme?	4.1
Zonder voldoende resources (kwal+kwant) is er een groter risico op fouten.	Beschikt de organisatie over voldoende deskundigheid, kwalitatief en kwantitatief?	
Geen compleet beeld op life-cycle waardoor sturing/beheersing niet mogelijk is.	Is het gehele proces/life-cycle rondom het algoritme gedocumenteerd?	
Onduidelijkheid rondom rollen, taken, verantwoordelijkheden en bevoegdheden creëert risico's.	Zijn de rollen, taken, verantwoordelijkheden en bevoegdheden in het proces beschreven (incl. eigenaarschap) en in de praktijk toegepast?	4.1
Prestatiedoelstellingen en kwaliteitsdoelstellingen zijn niet meetbaar of bespreekbaar als er geen aanpak is.	Is er een overeengekomen en vastgelegde aanpak m.b.t. kwaliteits- en prestatiedoelstellingen voor algoritmes?	4.2
Afhankelijkheid van extern deskundigen die na het ontwikkelen van het algoritme met de betreffende kennis en ervaring weg gaan, waardoor continuïteit en beheersing daarna niet meer geborgd is. Geen controle en beheersing op algoritme.	Zijn bij uitbesteding van onderdelen of activiteiten m.b.t. het algoritme afspraken met betrokken externe partijen gemaakt en vastgelegd?	4.1
Zonder monitoring is er geen beheersing mogelijk, nemen de risico's toe of vergroten.	Vindt periodiek monitoring plaats? (in ieder geval op beschikbaarheid, prestaties/kwaliteit, veiligheid en voldoen aan actuele wet/regelgeving, en op uitbesteding)	

Model & Data		
Risico	**Onderzoeksvraag**	**Ethisch principe**
Risico dat algoritme niet in lijn functioneert met geformuleerde doelstellingen. Zonder gedeeld beeld van de deelstellingen is er een groter risico op fouten, verschillen in interpretatie.	Is er een doel van het algoritme en is dat geoperationaliseerd in bruikbare aspecten i.h.k.v. te gebruiken model en data? Welke taak of onderdeel van de bedrijfsvoering dient het algoritme te ondersteunen?	4.2
Zonder gedeeld beeld van de deelstellingen is er een groter risico op fouten, verschillen in interpretatie.	Is er een gedeelde doel van het algoritme en inzichtelijk/ uitlegbaar voor eigenaar, ontwikkelaar en gebruiker?	4.2
Risico op niet of slecht uitlegbare toepassing van algoritmes.	Is het algoritme uitlegbaar en heeft er een afweging plaats gevonden tussen de uitlegbaarheid van het model en de prestatie van het model?	4.2
Het is niet meer terug te herleiden waarom welke keuzes zijn gemaakt in ontwerp en implementatie (explain).	Zijn de gemaakte overwegingen van het ontwerp en de implementatie vastgelegd?	4.1, 2.1
Geen continuïteit van het proces/ uitvoering werkzaamheden doordat documentatie ontbreekt.	Is er documentatie die de ontwerp en implementatie beschrijft?	4.1
Er heeft een willekeurig selectie van hyperparameters plaatsgevonden en daarbij zijn onjuiste keuzes gemaakt.	Zijn de keuzes m.b.t. de gebruik van hyperparameters beargumenteerd en onderbouwd?	
Ontbreken transparantie voor burgers/ bedrijven/stakeholders, niet voldoen aan wet/regelgeving m.b.t. transparantie.	Is het model (code en werking) gepubliceerd en beschikbaar voor belanghebbenden? Evenals indien mogelijk de gebruikte data of een beschrijving daarvan?	
Gebruik van geautomatiseerde besluitvorming wanneer dat niet is toegestaan of ontbreken van de mogelijkheid van menselijke tussenkomst.	Indien er geautomatiseerde besluitvorming plaatsvindt wordt daarbij voldaan aan de wetgeving/ regelgeving die daarvoor geldt?	1.1, 2.1
Te eenzijdige inbreng vergroot kans op fouten en niet voldoen aan doelen, wet/regelgeving.	Zijn de verschillende stakeholders/ ‘eindgebruikers’ van het algoritme betrokken in het ontwikkelproces?	3.1
Werking niet conform vooraf vastgestelde opzet en werking.	Welke controles zijn geïmplementeerd om de aansluiting te maken tussen de invoer en de uitvoer en daarmee de juistheid en volledigheid van de verwerking te garanderen?	2.1

Model & Data		
Risico	**Onderzoeksvraag**	**Ethisch principe**
Model is ontwikkeld op basis van regelgeving van jaar t-1, en wordt ingezet in jaar t. De regelgeving (grenswaarden, bedragen) kan ondertussen veranderd zijn of bepaalde bepalingen niet meer geldig.	Wordt het model periodiek geactualiseerd in lijn met actuele wet- en regelgeving?	
Onjuiste manier van training/testing kan leiden tot overfitting of underfitting, bias.	Is de kwaliteit geborgd m.b.t. keuzes die zijn gemaakt bij training en testdata?	4.1
Het model creëert onwenselijke systematische afwijking voor specifieke personen, groepen of andere eenheden (bias).	Wordt er geborgd dat er geen bias wordt gecreëerd door keuzes m.b.t. het model?	3.1, 3.2
Er zit onwenselijke systematische afwijking (bias) in de data.	Bevat de data geen onwenselijke bias?	3.1, 3.2
Als er niet wordt gescheiden dan is er sprake van overfitting en het model niet gebruikt kan worden voor nieuwe observaties.	Zijn training-, test- en validatiedata gescheiden verwerkt?	
De data is niet representatief.	Is de gebruikte data representatief voor de toepassing?	2.1, 3.1, 4.1
Afhankelijkheid van derden met betrekking tot gebruikte data.	Heeft de overheid volledige controle en beheersing over de gebruikte data voor het model? (“eigenaarschap”)	
Overtreden van uitgangspunten/ regels m.b.t. dataminimalisatie en proportionaliteit.	Is er sprake van dataminimalisatie, is gekeken naar proportionaliteit en subsidiariteit?	2.1
De performance metrics komen niet overeen met de doelstellingen van het algoritme.	Is de kwaliteit van het model gedocumenteerd?	4.2
De data waarop het model is gebaseerd is niet beschikbaar voordat de uitkomst geobserveerd wordt.	Is er target leakage? Ofwel zit hetgeen wat voorspeld moet worden bij de model features.	
Kwaliteit van de voorspelling is op orde.	Wordt er gebruikgemaakt van prestatie-indicatoren/ performance metrics?	2.1, 4.2
Soms werkt het model in de praktijk niet (meer).	Wordt de output van het model gemonitored?	2.1

Model & Data		
Risico	**Onderzoeksvraag**	**Ethisch principe**
Het is voor mensen niet duidelijk dat zij met een algoritme te maken hebben, welke consequenties dat heeft, welke beperkingen dit kent. Met in dat geval van incidenten/fouten schadeclaims achteraf tot gevolg.	Vindt er externe communicatie over het model/algoritme plaats inclusief de beperkingen (wat kan het wel en niet)?	4.2
Het risico bestaat dat alle focus en effort aan de voorkant wordt gestoken in het ontwikkelen en in productie brengen van het algoritme, zonder overdracht naar degene die dit moeten beheren en de business ook vergeten wordt in het onderhoud.	Vindt er onderhoud en beheer plaats op het algoritme?	

Privacy		
Risico	**Onderzoeksvraag**	**Ethisch principe**
Niet voldoen aan wettelijke verplichting AVG.	Wordt er een register bijgehouden m.b.t. het gebruik van persoons-gegevens?	2.2
Ontwerp en opzet zijn onvoldoende gericht op bescherming van privacy.	Is er sprake van “data protection by design”?	2.2
Niet voldoen aan wettelijke verplichting AVG.	Is er een DPIA uitgevoerd (indien van toepassing)?	2.2
Automatische besluitvorming terwijl dat volgens AVG niet is toegestaan.	Is er sprake van automatische besluitvorming en zo ja is dit toegestaan?	2.2
Niet voldoen aan wettelijke verplichting AVG/menselijke maat.	Hebben de betrokkenen de mogelijkheid niet onderworpen te zijn aan geautomatiseerde besluitvorming (indien van toepassing)?	2.2
Niet proportioneel gebruik/ verzameling van persoonsgegevens.	Is er sprake van data minimalisatie?	2.2
Niet-wettelijk handelen.	Vind de verwerking van gegevens plaats op grond van een wettelijke taak?	2.2
Niet voldoen aan doelbinding/AVG.	Is de verwerking van (bijzondere) persoonsgegevens met het algoritme verenigbaar met het oorspronkelijke doel?	2.2

Privacy		
Risico	**Onderzoeksvraag**	**Ethisch principe**
Niet voldoen aan wettelijke verplichting AVG.	Is de verwerkingsverantwoordelijke en verwerker van de persoons-gegevens met betrekking tot het algoritme en de daarbij gebruikte data vastgesteld?	2.2
Handelen in strijd met art. 1 GW/art.14 EVRM.	Is er geen sprake van discriminatie door gebruikte data en model?	2.2
Profilering in de zin van artikel 4 sub 4 AVG, risico handelen in strijd met AVG.	Is er getoetst in hoeverre er sprake is van profilering en in hoeverre dat is toegestaan?	2.2
Niet voldoen aan wettelijke verplichting AVG.	Is er invulling gegeven aan het proactief of op verzoek informeren van betrokkenen wier gegevens worden verwerkt/gebruikt? (zowel data als algoritme)	2.2
Niet voldoen aan wettelijke verplichting AVG.	Is de logica van het gebruikte algoritme, de gebruikte gegevens, voldoende duidelijk voor betrokkenen?	2.2
Niet voldoen aan wettelijke verplichting AVG.	Zijn de gevolgen van de toepassing van het gebruikte algoritme duidelijk voor betrokkenen?	2.2
Betrokkenen zijn niet op de hoogte van hun rechten, gebruikte algoritmes en data.	Is er een openbaar privacybeleid waarin gebruikte data en algoritmes aan bod komen?	2.2

ITGC		
Risico	**Onderzoeksvraag**	**Ethisch principe**
Zonder loginformatie is niet te achterhalen wanneer er aanpassingen zijn gedaan (audit trail).	Wordt loginformatie omtrent de werking van het algoritme bewaard en toegankelijk gemaakt?	
Toegangsrechten niet meer up-to-date.	Wordt gecontroleerd of toegangsrechten up-to-date zijn met betrekking tot de omgeving waarin het algoritme functioneert?	2.2
Onrechtmatige toegang tot het algoritme.	Worden toegangsrechten aangepast zodra er een uitdiensttreding of functiewijziging van een werknemer plaatsvindt?	2.2

ITGC		
Risico	**Onderzoeksvraag**	**Ethisch principe**
Toegang wordt uitgegeven door persoon die daarvoor niet geautoriseerd is.	Worden toegangsrechten uitgegeven door daarvoor bevoegde personen?	2.2
Kans op manipulatie van het algoritme bij conflicterende toegangsrechten.	Wordt functievermenging voorkómen bij de toegang van gebruikers tot het algoritme?	2.2
Hoe meer toegewezen speciale bevoegdheden, hoe meer kans op manipulatie.	Wordt er gebruik gemaakt van generieke beheeraccounts? Staat het aantal beheeraccounts in logische verhouding met de beheerders?	2.2
Gebruikersgroepen van het algoritme lastig te identificeren.	Wordt er bij het inrichten van toegangsrechten van verschillende gebruikersgroepen/rollen gebruik gemaakt van naamgevingsconventies en systematiek?	2.2
Beheerders en gebruikers van het algoritme lastig te identificeren.	Worden er naamgevingsconventies gebruikt voor gebruikers en beheerders, zodat zij geïdentificeerd kunnen worden?	2.2
Onduidelijkheid in wie wijzigingen/ werkzaamheden aan het algoritme heeft uitgevoerd.	Voeren beheerders werkzaamheden als beheerder en werkzaamheden als gewone gebruiker onder twee verschillende gebruikersnamen uit?	2.2
Indien wel toegang tot onderliggende componenten kan manipulatie van de database plaatsvinden.	Hebben gebruikersaccounts (geen) directe toegang tot onderliggende componenten?	2.2
Indien wel toegang tot onderliggende componenten kan manipulatie van de database plaatsvinden.	Bestaat er een functiescheiding tussen aanvragen, autoriseren en verwerken van wijzigingen in gebruikersaccounts en toegangs-rechten?	2.2
Indien wel toegang tot onderliggende componenten kan manipulatie van de database plaatsvinden.	Is het wachtwoordbeheer interactief en zijn de wachtwoorden van geschikte kwaliteit?	2.2
Ongeautoriseerde toegang, wijziging, beschadiging en/of dataverlies. Niet naleven van wetgeving.	Worden wijzigingen in de code van het algoritme op een gecontroleerde wijze uitgevoerd? (denk aan het testen en accorderen/autoriseren van wijzigingen)	2.2

ITGC		
Risico	**Onderzoeksvraag**	**Ethisch principe**
Ongeautoriseerde toegang en daarmee kans op manipulatie van het algoritme (wijziging, beschadiging, dataverlies).	Is het algoritme beveiligd, zodat er geen risico is op ongeautoriseerde toegang, wijziging, beschadiging en/of dataverlies?	2.2
Back-ups zijn niet in overeenstemming met het back-upbeleid. Er is geen hersteloptie bij uitval van het algoritme en risico van gegevensverlies.	Worden er back-ups van het algoritme gemaakt en kan het algoritme hersteld worden?	
Bij het ontbreken van security by design zijn de risico's veel groter.	Is er sprake van security by design?	2.1

Ethiek[29]		
Ethisch raamwerk	**Ethisch Principe**	**Nummer**
Respect voor menselijke autonomie.	De beslissingen die gemaakt zijn door het algoritme zijn te controleren d.m.v. menselijke tussenkomst.	1.1
Voorkomen van schade.	Het algoritme is veilig en doet ten alle tijden waar het voor gemaakt is.	2.1
	Privacy is gewaarborgd en data is beschermd.	2.2
Fairness (eerlijke algoritmes).	Het algoritme houdt rekening met diversiteit in de populatie en discrimineert niet.	3.1
	Er is bij de ontwikkeling van het algoritme rekening gehouden met impact op maatschappij en milieu.	3.2
Verklaarbaarheid en transparantie.	Er kan verantwoording worden afgelegd over de gevolgde procedures.	4.1
	De werking van het algoritme is te verklaren en uit te leggen.	4.2

Bijlage 4
Eindnoten

1. Aankondiging onderzoek *Zicht op Algoritmes*, Algemene Rekenkamer, februari 2020
2. Zie voor deze uitspraak https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBDHA:2020:865#:~:text=De%20rechtbank%20heeft%20vandaag%20uitspraak,Systeem%20Risico%20Indicatie%20(SyRI).&text=De%20rechtbank%20moest%20toetsen%20of,artikel%208%20lid%202%20EVRM.
3. Een *voorspellend* algoritme wordt ingezet voor een analyse van de vraag 'Wat zal er gebeuren?', een *voorschrijvend* algoritme voor een analyse van de vraag 'Wat moet er gebeuren?' (zie ook § 2.2).
4. Het algoritme ontdekt in de loop van de tijd, op basis van nieuwe data, nieuwe verbanden (correlaties) en genereert op basis daarvan uitkomsten. Het algoritme 'leert'.
5. Rapporten Algemene Rekenkamer (2019 en 2020). (1) I*nformatiebeveiliging Verantwoordingsonderzoek 2019,* (2) *Digitalisering aan de grens* (20 april 2020) en (3) *Digitale dijkverzwaring: cybersecurity en vitale waterwerken* (28 maart 2019).
6. *IT General Controls* (ITGC) zijn de beheersmaatregelen die een organisatie heeft getroffen om ervoor te zorgen dat de IT-systemen betrouwbaar en integer zijn. Het zijn traditionele ICT-maatregelen, zoals het beheer van toegangsrechten (zie hoofdstuk 4.1).
7. Er zijn meerdere definities voor een algoritme. Op hoofdlijnen komen ze allemaal overeen met de formulering die wij hebben gebruikt. Zie bijlage 1 voor de geraadpleegde literatuur in dit onderzoek.
8. *Siri, Siri in my hand, who's the Fairest in the Land?*, 2018, Kaplan & Haenlein.
9. Strategisch Actieplan voor Artificiële Intelligentie, 8 oktober 2019, TK 2019D39726.
10. Bijlage: https://www.rijksoverheid.nl/documenten/rapporten/2019/10/08/tk-bijlage-over-waarborgen-tegen-risico-s-van-data-analyses-door-de-overheid bij de Kamerbrief: https://www.rijksoverheid.nl/documenten/kamerstukken/2019/10/08/tk-waarborgen-tegen-risico-s-van-data-analyses-door-de-overheid.
11. In verband met de corona-maatregelen hebben wij een beperkt aantal experts laten deelnemen.
12. Vanuit de SVB wordt verwezen naar een eerder onderzoek van de Algemene Rekenkamer (2019): *Ouderdomsregelingen ontleed* (13 november 2019).

13. Bijlage bij Kamerbrief ‘Waarborgen tegen risico’s van data-analyses door de overheid’ (8 oktober 2019), TK 26643-641.
14. *Deep learning* is een vorm van *machine learning* waarbij modellen worden gebruikt die overeenkomsten vertonen met neurale netwerken in het brein. *Machine learning* ontwikkelt algoritmes waarmee computers kunnen leren.
15. Zie ons rapport *Datagedreven selectie van aangiften door de Belastingdienst* (11 juni 2019).
16. *Control Objectives for Information and related Technology* (COBIT) is de standaard die tegemoetkomt aan de behoefte aan controle over informatie- en IT-gerelateerde risico’s.
17. Gevoelige gegevens als iemands ras, godsdienst of gezondheid worden bijzondere persoonsgegevens genoemd. Deze zijn door de wetgever extra beschermd (bron: https://autoriteitpersoonsgegevens.nl/nl/over-privacy/persoonsgegevens/wat-zijn-persoonsgegevens).
18. Herijking wil zeggen dat er opnieuw wordt bepaald of het algoritme nog steeds aan de vastgestelde normen voldoet.
19. Met lifecycle management bedoelen we in dit verband het planmatig onderhouden van algoritmes gedurende hun hele looptijd, zodat deze onderdeel zijn en blijven van een duurzaam en toekomstbestendig IT-landschap.
20. Door het vrije verkeer van persoonlijke gegevens goed te regelen –op een manier die het vertrouwen van mensen in de samenleving (en de overheid) vergroot en beschermt– kun je de (dienstverlening in de) digitale samenleving verbeteren. Dat is het uitgangspunt voor het programma Regie op Gegevens (bron: https://www.digitaleoverheid.nl/dossiers/rog-regie-op-gegevens/).
21. MijnOverheid is de overheidswebsite waarop burgers digitale berichten van de overheid kunnen ontvangen en hun persoonlijke gegevens kunnen inzien.
22. *Data Protection Impact Assessments* (DPIA’s) zijn effectbeoordelingen van gegevensbescherming.
23. *Baseline Informatiebeveiliging Overheid* (BIO) gebaseerd op de internationale norm ISO/IEC 27002.
24. I*nformation Technology Infrastructure Library* (ITIL) is een referentiekader voor het inrichten van de beheerprocessen binnen een ICT-organisatie.
25. *Data Protection Impact Assessments* (DPIA’s) zijn effectbeoordelingen van gegevensbescherming.
26. Kamerstuk TK 35212, nr. 5 d.d. 15 oktober 2020.
27. Bijlage: https://www.rijksoverheid.nl/documenten/rapporten/2019/10/08/tk-bijlage-over-waarborgen-tegen-risico-s-van-data-analyses-door-de-overheid bij de brief: https://www.rijksoverheid.nl/documenten/kamerstukken/2019/10/08/tk-waarborgen-tegen-risico-s-van-data-analyses-door-de-overheid.

28. De vragen in het toetsingskader zijn mede opgesteld aan de hand van ethische principes. De nummers refereren aan een ethisch principe, in de tabel onderaan dit document zijn die omschreven.
29. De vragen in het toetsingskader zijn mede opgesteld aan de hand van deze ethische principes. De principes zijn grotendeels afkomstig uit de rapporten *Ethic guidelines for trustworthy AI* (2019) en *Whitepaper on Artificial Intelligence – A European approach to excellence and trust* (2020) van de Europese Commissie (2019).

Algemene Rekenkamer
Voorlichting
Afdeling Communicatie
Postbus 20015
2500 EA Den Haag
telefoon (070) 342 44 00
voorlichting@rekenkamer.nl
www.rekenkamer.nl

Omslag
Ontwerp: Ontwerpwerk
Foto: Getty Images

Den Haag, januari 2021