Document 1

Vragenlijst naar het gebruik van algoritmen binnen de overheid

Achtergrond en toelichting

Aanleiding

In het 30-leden debat van 4 oktober[1] over de Digitale Infrastructuur is gesproken over de mogelijkheden van algoritmen en het gebruik ervan door de overheid. Door de leden Jetten (D66) en Bruins-Slot (CDA) zijn een aantal zorgen geuit over de inzet van algoritmen. Daarom heeft de staatssecretaris van BZK toegezegd om een onderzoek uit te voeren naar de toepassing van algoritmen binnen de overheid. Hij concentreert zich daarbij op twee hoofdvragen:

1. Welke overheidsorganen gebruiken welke algoritmen?
2. Gegeven de complexiteit , en soms het “black-box” karakter van een aantal ervan, kunt u inzicht geven in de werking van deze algoritmen en hoe u er over communiceert?

Op basis daarvan is deze vragenlijst opgezet en uitgezet bij de CIO’s van departementen, provincies, gemeenten en waterschappen. De uitkomsten worden benut voor en onderdeel van de Nationale Agenda Digitalisering Openbaar Bestuur. De uitkomsten worden tevens geanalyseerd en aan de kamer gemeld.

Wat zijn algoritmen?

Een algoritme is een eindige reeks instructies die vanuit een gegeven begintoestand naar een beoogd doel leidt. Algoritmes zijn vergelijkbaar met een recept. Door een aantal stappen uit te voeren ontstaat er een (hopelijk) lekkere maaltijd. Algoritmen staan in beginsel los van computerprogramma's, maar in dit voorstel hebben we het uitsluitend over algoritmes die door computers worden gebruikt.

In de informatica is elke verzameling instructies een algoritme. Dat is niet wat we bedoelen: het berekenen van een gemiddelde is ook een algoritme en dat vinden we in deze context niet relevant. Wat tijdens het debat in het algemeen bedoeld werd met algoritmen zijn “intelligente” algoritmen. Die worden meestal in twee typen onderverdeeld: rule-based en case-based.

1. Rule-based systemen

Rule-based: algoritmen die op basis van een aantal regels tot een bepaalde conclusie komen. Het betreft vaak top-down regels die worden gecreëerd voor een systeem, dat daar vervolgens over kan doorredeneren.

De basis van de meeste rule-based systemen is gelegen in als-dan redeneringen in de trant van: als iemand jonger is dan 18 jaar, dan mag hij niet de supermarkt verlaten met drank. Wet- en regelgeving laat zich relatief eenvoudig vertalen in deze regels. Rule-based systemen hebben als voordeel dat regels, redeneringen en aannames expliciet moeten worden gemaakt. Dit heeft als meestal als gevolg dat de manier waarop een systeem tot zijn conclusie komt expliciet kan worden gemaakt. Wat vaak ook als een voordeel wordt gezien is de modulariteit van de aanpak en uniformiteit van de regels, namelijk alle regels worden uitgeschreven in een zelfde soort format wat de opbouw van een kennisbase vereenvoudigd.

Nadelen zijn er ook: het grootste gevaar is misschien wel dat de kennis base erg groot en complex wordt waardoor nieuwe kennis soms contradictoir is met bestaande kennis. Een ander gevolg van zo’n grote kennisbase is dat niet altijd duidelijk is wat de precieze effecten van het veranderen van één regel op alle mogelijke uitkomsten is, vanwege het grote aantal regels dat afhankelijk is van die ene regel.

[1] 4 oktober 2017, Tweede Kamer, Digitalisering infrastructuur, TK 8, 8-11-2

Case-based: de training van algoritmen

Case-based algoritmen zijn algoritmen die op basis van een aantal casussen geleerd hebben en voorspellingen kunnen doen over nog niet bekende gevallen. Dit veld is nog relatief nieuw, vergeleken met rule-based systemen. Sommige case-based algoritmen worden al veelvuldig ingezet bij de overheid, zoals lineaire regressie en, in mindere mate, beslisbomen. Andere algoritmen, zoals neurale netwerken zijn vermoedelijk nog maar in de onderzoeksfase van een paar bestuursorganen. Zie figuur 1 voor een overzicht.

Voorbeeld: stel dat een algoritme geleerd heeft dat iemand met een uitkering én een Porsche een grote kans heeft een fraudeur te zijn. Dan zal een goed getraind algoritme waarschijnlijk de voorspelling doen dat iemand met een uitkering én een Maserati óók een fraudeur is, ook al heeft het algoritme deze specifieke casus niet langs zien komen.

De crux zit hem in het goed getraind zijn van het algoritme: wat zijn de kenmerken van het goed getraind zijn? Hoe weet ik dat het netwerk goed generaliseert? Wordt er niet, direct of indirect, gediscrimineerd op variabelen zoals geslacht, afkomst of opleiding? Zo lijkt een algoritme dat gebruikt wordt om overtreders te straffen in Wisconsin, zwarten significant zwaarder te bestraffen dan blanken. Goed trainen van een netwerk betekent niet alleen kijken of het algoritme goede voorspellingen kan doen maar ook of het niet teveel 'bias' bevat, onbedoeld op kenmerken selecteert waarop we juist niet willen letten, althans niet in die situatie.



Figuur 1. Globaal overzicht en classificatie van case-based algoritmen. Dit overzicht is niet uitputtend.

Uitlegbaarheid van algoritmen

De reden dat we niet altijd goed weten of een algoritme 'last heeft' van onbedoelde zaken ligt in het feit dat algoritmen niet altijd even makkelijk te doorgronden zijn, zelfs niet altijd door statistici en informatici. Dat betekent dat nagedacht moet worden over andere verantwoordingsmodellen die meer ingaan op de manier waarop de algoritmes zijn getraind, welke gegevens daarvoor zijn gebruikt en de uitlegbaarheid van dat hele proces. Omdat verantwoording essentieel is voor een bestuursorgaan dat besluiten neemt, zijn we in dit onderzoek ook geïnteresseerd hoe een

bestuursorgaan organiseert dat het inzicht kan geven in de werking van een algoritme dat niet altijd even inzichtelijk is.

Worden algoritmes ingezet bij de overheid?

Algoritmes worden zeker ingezet door de overheid. Zo gebruikt de Landelijke Stuurgroep Interventies het Systeem Risico Indicatie (SyRI) om fraudeurs op te sporen, de belastingdienst experimenteert met allerlei algoritmen om foutieve aangiften op te kunnen sporen, de politie is bezig met het voorspellen van waar en wanneer door wie een misdaad zal plaatsvinden en zo verder. Dus ja, algoritmes worden door de overheid ingezet maar we weten nauwelijks door welke overheidsorganisaties, met welk doel en met welke resultaten.

Scope

Om de scope van dit onderzoek te beperken hebben we het in dit onderzoek alleen over algoritmen die de basis vormen voor besluiten over personen of anderszins een behoorlijke impact hebben op een persoon of groepen personen. Ook algoritmen die impact hebben op groepen personen vallen onder de scope van dit onderzoek. Als ergens een dijk wordt volgestopt met sensoren (zoals in Groningen) en een algoritme wordt ingezet om te voorspellen of de dijk op doorbreken staat, valt dat onder de scope omdat het resultaat van het algoritme impact heeft op groepen personen.

Mocht u een externe partij voor uw algoritmen inzetten dan valt dit ook binnen de scope omdat uw organisatie aangesproken kan worden.

Het gaat om algoritmes die worden ingezet in productie en onderzoek.

Wat willen we weten?

Om maar een begin van beleid te kunnen vormen op het gebied van algoritmen is het nodig om een beeld te hebben welke overheidspartijen algoritmen inzetten, met welk doel en met welk succes. Omdat het in veel gevallen een relatief nieuw terrein betreft is het ook nodig om een beeld te krijgen van organisaties die hiernaar een onderzoek doen of ermee aan het experimenteren zijn. Als laatste willen we weten of er behoefte bestaat aan beleid op dit terrein en zo ja, welk. Dat leidt tot de volgende drie hoofdvragen:

1. Zet uw organisatie algoritmen in, welke typen, met welke doelen en met welk succes?
2. Kan de organisatie inzicht geven over de werking van het algoritme?
3. Heeft u behoefte (of juist niet) aan beleid op het gebied van inzet van algoritmen?

De vragenlijst wordt tezamen met dit document meegestuurd. De uitleg vind u in de bijlage.

Bijlage. Toelichting op de Vragenlijst

Deze bijlage bevat de toelichting op de vragen. Afhankelijk van de beantwoording op vraag 1, kan vraag 2 al dan niet beantwoord worden. We vragen u om in alle gevallen vraag 3 te beantwoorden.

I. Past uw organisatie algoritmes toe en zo ja, met welk doel?

- Welke algoritmes gebruikt uw organisatie in de primaire processen? Gebruik daarvoor eventueel de illustratie in figuur 1. Dat overzicht is niet uitputtend en zo uw organisatie algoritmes test of inzet waarvan u vindt dat ze in het overzicht ontbreken dan kunt u deze hier melden.
- Per algoritme, wat is het doel dat ermee beoogd wordt, het resultaat en eventuele ongewenste (neven)effecten?
- Doet uw organisatie onderzoek naar algoritmes? Welke doelen zou u met de onderzochte algoritmes willen bereiken?

De volgende vraag betreffen informatie over het algoritme zelf en uw communicatie erover. We vragen u dit voor maximaal twee algoritmes te beantwoorden. We zijn geïnteresseerd in met name de meer “moeilijk uitlegbare” algoritmes.

We verzoeken u uit de algoritmes die u toepast in onderzoek of productie, twee door u gebruikte algoritmes te kiezen die wat moeilijker uit te leggen zijn: één rule-based algoritme en één case-based algoritme. Mocht u geen case-based algoritmen gebruiken dan graag twee rule-based kiezen en omgekeerd: twee case-based indien geen rule-based wordt gebruikt.

Wat de Kamer graag wil weten of het algoritme op een goede manier wordt ingezet. De vragen zijn niet bedoeld als controle of u uw werk goed doet. We zijn juist geïnteresseerd in de maatregelen die u genomen heeft om u zich van de correcte werking van het algoritme te overtuigen.

De reden dat deze vragen gesteld worden ligt in het feit dat algoritmen niet altijd even makkelijk te doorgronden zijn, zelfs niet altijd voor statistici en informatici.

Dat betekent dat nagedacht moet worden over de uitleg over algoritmes. Denk aan de manier waarop de algoritmes zijn getraind, welke gegevens daarvoor zijn gebruikt, de code en de uitlegbaarheid van dat hele proces. Omdat uitleg en inzicht essentieel is voor een bestuursorgaan dat besluiten neemt, zijn we in dit onderzoek geïnteresseerd hoe een bestuursorgaan communiceert over een algoritme dat niet altijd even inzichtelijk is.

II. Inzicht in het algoritme

- Hoe legt u het algoritme uit aan een geïnteresseerde leek?
- Kunt u beschrijven welke training- en validatiegegevens zijn gebruikt? Zijn deze openbaar of bent u bereid deze openbaar te maken? Waarom niet?
- Hoe heeft u bij de ontwikkeling concepten als privacy by design, ethics by design of responsible innovation toegepast?
- Hoe toetst u of er indirect op verborgen kenmerken wordt geselecteerd?
- Op welke gronden heeft uw organisatie besloten om juist dit algoritme te gebruiken?
- Hoe wordt het algoritme op de juiste werking geaudit?
    - Is het wetenschappelijk onderbouwd?
    - Is er een gedetailleerde omschrijving van het algoritme?
    - Zijn de implementaties van de algoritmen open-source?
    - Zijn training en validatie sets open data?
- Op welke manier worden het gebruik van algoritmen, het algoritme zelf, de methodologie, het proces en de uitkomst transparant gemaakt voor eindgebruikers?
- Kan de werking van het algoritme worden geverifieerd op basis van het algoritme, de implementatie ervan en de gebruikte data? Door zowel rechters, toezichthouders als burgers?
- Welke positieve en (wellicht onverwacht) negatieve effecten heeft het gebruik van het algoritme?

III. heeft u behoefte aan beleid op het gebied van algoritmen?

- Heeft u behoefte aan beleid of regelgeving op het gebied van algoritmen en zo ja, welk?
- Of heeft u juist liever geen beleid op bepaalde terreinen, zo ja, welke?

Vragenlijst onderzoek algoritmen

We vragen u onderstaande vragen lijst in te vullen. Afhankelijk van de beantwoording op vraag 1, kan vraag 2 al dan niet beantwoord worden. We vragen u om in alle gevallen vraag 3 te beantwoorden.

U kunt de ingevulde vragenlijst, uiterlijk 25 april, mailen aan: [redacted]@minbzk.nl

Mocht u vragen hebben, kunt u die stellen op dat email adres, of bellen op [redacted].

I. Past uw organisatie algoritmes toe en zo ja, met welk doel?

I.a Welke algoritmes gebruikt uw organisatie in de primaire processen? Gebruik daarvoor eventueel de illustratie in figuur 1 van de toelichting. Dat overzicht is niet uitputtend en zo uw organisatie algoritmes test of inzet waarvan u vindt dat ze in het overzicht ontbreken dan kunt u deze hier melden.

I.b Per algoritme, wat is het doel dat ermee beoogd wordt, het resultaat en eventuele ongewenste (neven)effecten?

I.c Doet uw organisatie onderzoek naar algoritmes? Welke doelen zou u met de onderzochte algoritmes willen bereiken?

II. Inzicht in het algoritme

II.a Hoe legt u het algoritme uit aan een geïnteresseerde leek?

II.b Kunt u beschrijven welke training- en validatiegegevens zijn gebruikt? Zijn deze openbaar of bent u bereid deze openbaar te maken? Waarom niet?

II.c Hoe heeft u bij de ontwikkeling concepten als privacy by design, ethics by design of responsible innovation toegepast?

II.d Hoe toetst u of er indirect op verborgen kenmerken wordt geselecteerd?

II.e Op welke gronden heeft uw organisatie besloten om juist dit algoritme te gebruiken?

II.f. Hoe wordt het algoritme op de juiste werking geaudit?

II.f.1 Is het wetenschappelijk onderbouwd?

II.f.2 Is er een gedetailleerde omschrijving van het algoritme?

II.f.3 Zijn de implementaties van de algoritmen open-source?

II.f.4 Zijn training en validatie sets open data?

II.g Op welke manier worden het gebruik van algoritmen, het algoritme zelf, de methodologie, het proces en de uitkomst transparant gemaakt voor eindgebruikers?

II.h Kan de werking van het algoritme worden geverifieerd op basis van het algoritme, de implementatie ervan en de gebruikte data? Door zowel rechters, toezichthouders als burgers?

II.i Welke positieve en (wellicht onverwacht) negatieve effecten heeft het gebruik van het algoritme?

III. heeft u behoefte aan beleid op het gebied van algoritmen?

III.a Heeft u behoefte aan beleid of regelgeving op het gebied van algoritmen en zo ja, welk?

III.b Of heeft u juist liever geen beleid op bepaalde terreinen, zo ja, welke?

Document 3

Ministerie van Binnenlandse Zaken en Koninkrijksrelaties

Gedeputeerde Staten van de Provincie Drenthe
Postbus 122
9400 AC Assen

www.rijksoverheid.nl
www.facebook.com/minbzk
www.twitter.com/minbzk

Kenmerk
2018-0000204233

Uw kenmerk

Bijlage(n)
2

Datum 16 APR 2018
Betreft Verzoek om medewerking aan onderzoek naar algoritmen

Geachte heer, mevrouw,

Bij deze vraag ik uw medewerking aan een onderzoek dat ik laat uitvoeren naar de toepassing van algoritmen binnen de overheid.

Naar aanleiding van een overleg met de Tweede Kamer over de toepassing van algoritmen door de overheid, wordt door het ministerie van Binnenlandse Zaken en Koninkrijksrelaties een onderzoek uitgevoerd. Doel van het onderzoek is inzicht te verwerven in welke bestuursorganen welke algoritmen toepassen, of bestuursorganen onderzoek doen naar algoritmen en hoe deze zich kunnen verantwoorden over de toepassing van deze algoritmen. Het onderzoek wordt uitgezet bij departementen, hun ZBO's, provincies, gemeenten en waterschappen. De uitkomsten zullen later dit jaar aan de Tweede Kamer worden verzonden.

Graag vraag ik uw medewerking bij de uitvoering van het onderzoek door het invullen van de vragenlijst. Ik stel het zeer op prijs als de ingevulde vragenlijst uiterlijk 10 mei wordt geretourneerd aan het ministerie (zie bijlage voor contactgegevens en aanspreekpunt voor eventuele vragen).

Ik dank u bij voorbaat voor uw medewerking.

De staatssecretaris van Binnenlandse zaken en Koninkrijksrelaties,
namens deze,

Simone Roos
Directeur-generaal Overheidsorganisatie

Document 4

Van: Postbus CIOberaad
Aan: [redacted]
Cc: [redacted]
Onderwerp: Onderzoek algoritmen
Datum: maandag 26 maart 2018 08:57:54
Bijlagen: Vragenlijst Algoritmen.docx
Vragenlijst Algoritmen Toelichting.docx
kst-26643-507.pdf
Informatie_over_onderzoek_naar_algoritmen.pdf

Beste leden van het CIO-beraad,
In het CIO-beraad van 7 februari hebben we een onderzoek naar het gebruik van algoritmen binnen de overheid aangekondigd. Staatssecretaris Knops heeft dit onderzoek aan de Kamer toegezegd, volgend op een dertigledendebat met EZK in 2017 (zie bijgaande Kamerbrieven voor meer informatie). Onze zusterdirectie DIO voert het onderzoek uit en heeft mij gevraagd om dit voor de sector Rijk naar jullie te verspreiden. Parallel wordt het onderzoek ook via de koepels van gemeenten, provincies en waterschappen uitgezet. Het verzoek van DIO is om bijgaande vragenlijst in te vullen voor het kerndepartement, eventuele agentschappen en ZBO's. Zij mogen de vragenlijst ook zelfstandig invullen. Bij vraag II van de vragenlijst worden veel subvragen gesteld. Daarom is deze vraag gelimiteerd tot maximaal twee algoritmen om de hoeveelheid invulwerk te beperken (meer mag natuurlijk altijd). De ingevulde vragenlijsten kunnen jullie sturen naar [redacted]@minbzk.nl). Met eventuele vragen kunnen jullie hem direct benaderen. De deadline van het onderzoek is 18 april 2018.
Met vriendelijke groet,
Hans Wanders
CIO Rijk

Document 5

Ministerie van Binnenlandse Zaken en Koninkrijksrelaties

Unie van Waterschappen
t.a.v. dhr. [redacted]
Postbus 93218
2509 AE Den Haag

www.rijksoverheid.nl
www.facebook.com/minbzk
www.twitter.com/minbzk

Kenmerk
2018-0000190688

Uw kenmerk

Datum 28 MRT 2018

Betreft Verzoek om medewerking aan onderzoek naar algoritmen

Geachte heer [redacted]

Bij deze vraag ik uw medewerking aan een onderzoek dat ik laat uitvoeren naar de toepassing van algoritmes binnen de overheid.

Naar aanleiding van een overleg met de Tweede Kamer over de toepassing van algoritmen door de overheid, wordt door het ministerie van Binnenlandse Zaken en Koninkrijksrelaties een onderzoek uitgevoerd. Doel van het onderzoek is inzicht te verwerven in welke bestuursorganen welke algoritmen toepassen, of bestuursorganen onderzoek doen naar algoritmen en hoe deze zich kunnen verantwoorden over de toepassing van deze algoritmen. Het onderzoek wordt uitgezet bij departementen, hun ZBO's, provincies, gemeenten en waterschappen. De uitkomsten zullen later dit jaar aan de Tweede Kamer worden verzonden.

Graag vraag ik uw medewerking bij de uitvoering van het onderzoek door het uitzetten van de vragenlijst bij uw leden. Ik stel het zeer op prijs als de ingevulde vragenlijst uiterlijk 25 april wordt geretourneerd aan het ministerie (zie bijlage voor contactgegevens en aanspreekpunt voor eventuele vragen).

De uitkomsten van dit onderzoek zijn van belang voor zowel Rijk als gemeenten, provincies en waterschappen. Daarom nodig ik u bij deze uit om uw organisatie deel te laten nemen aan een begeleidingsgroep die mede op basis van de uitkomsten van het onderzoek een visie en mogelijk beleid rondom de toepassing van algoritmen voorbereidt. Gedacht wordt aan iemand die goed is ingevoerd in het IT-beleid van uw organisatie en in staat is de consequenties van intelligente algoritmes te doordenken. Ik dank u bij voorbaat voor uw medewerking.

De staatssecretaris van Binnenlandse zaken en Koninkrijksrelaties,
namens deze,

[redacted]

Simone Roos
Directeur-generaal Overheidsorganisatie

Vragenlijst naar het gebruik van algoritmen binnen de overheid

Achtergrond en toelichting

Aanleiding

In het 30-leden debat van 4 oktober[1] over de Digitale Infrastructuur is gesproken over de mogelijkheden van algoritmen en het gebruik ervan door de overheid. Door de leden Jetten (D66) en Bruins-Slot (CDA) is een aantal zorgen geuit over de inzet van algoritmen. Daarom heeft de staatssecretaris van BZK toegezegd om een onderzoek uit te voeren naar de toepassing van algoritmen binnen de overheid. Hij concentreert zich daarbij op twee hoofdvragen:

1. Welke overheidsorganen gebruiken welke algoritmen?
2. Gegeven de complexiteit , en soms het "black-box" karakter van een aantal ervan, kunt u inzicht geven in de werking van deze algoritmen en hoe u er over communiceert?

Op basis daarvan is deze vragenlijst opgezet en uitgezet bij de CIO's van departementen, provincies, gemeenten en waterschappen. De uitkomsten worden benut voor de vorming van een beleidsvisie op algoritmen en onderdeel van de Nationale Agenda Digitalisering Openbaar Bestuur. De uitkomsten worden tevens geanalyseerd en aan de Tweede Kamer gemeld.

Wat zijn algoritmen?

Een algoritme is een eindige reeks instructies die vanuit een gegeven begintoestand naar een beoogd doel leidt. Algoritmes zijn vergelijkbaar met een recept. Door een aantal stappen uit te voeren ontstaat er een (hopelijk) lekkere maaltijd. Algoritmen staan in beginsel los van computerprogramma's, maar in dit voorstel hebben we het uitsluitend over algoritmes die door computers worden gebruikt.

In de informatica is elke verzameling instructies een algoritme. Dat is niet wat we bedoelen: het berekenen van een gemiddelde is ook een algoritme en dat vinden we in deze context niet relevant. Wat tijdens het debat in het algemeen bedoeld werd met algoritmen zijn "intelligente" algoritmen. Die worden meestal in twee typen onderverdeeld: rule-based en case-based.

1. Rule-based systemen

Rule-based: algoritmen die op basis van een aantal regels tot een bepaalde conclusie komen. Het betreft vaak top-down regels die worden gecreëerd voor een systeem, dat daar vervolgens over kan doorredeneren.

De basis van de meeste rule-based systemen is gelegen in als-dan redeneringen in de trant van: als iemand jonger is dan 18 jaar, dan mag hij niet de supermarkt verlaten met drank. Wet- en regelgeving laat zich relatief eenvoudig vertalen in deze regels. Rule-based systemen hebben als voordeel dat regels, redeneringen en aannames expliciet moeten worden gemaakt. Dit heeft meestal als gevolg dat de manier waarop een systeem tot zijn conclusie komt expliciet kan worden gemaakt. Wat vaak ook als een voordeel wordt gezien is de modulariteit van de aanpak en uniformiteit van de regels, namelijk alle regels worden uitgeschreven in een zelfde soort format wat de opbouw van een kennisbase vereenvoudigt.

Nadelen zijn er ook: het grootste gevaar is misschien wel dat de kennisbase erg groot en complex wordt waardoor nieuwe kennis soms contradictoir is met bestaande kennis. Een ander gevolg van zo'n grote kennisbase is dat niet altijd duidelijk is wat de precieze effecten van het veranderen van één regel op alle mogelijke uitkomsten zijn, vanwege het grote aantal regels dat afhankelijk is van die ene regel.

---

[1] 4 oktober 2017, Tweede Kamer, Digitalisering infrastructuur, TK 8, 8-11-2

Case-based: de training van algoritmen

Case-based algoritmen zijn algoritmen die op basis van een aantal casussen geleerd hebben en voorspellingen kunnen doen over nog niet bekende gevallen. Dit veld is nog relatief nieuw, vergeleken met rule-based systemen. Sommige case-based algoritmen worden al veelvuldig ingezet bij de overheid, zoals lineaire regressie en, in mindere mate, beslisbomen. Andere algoritmen, zoals neurale netwerken zijn vermoedelijk nog maar in de onderzoeksfase van een paar bestuursorganen. Zie figuur 1 voor een overzicht.

Voorbeeld: stel dat een algoritme geleerd heeft dat iemand met een uitkering én een Porsche een grote kans heeft een fraudeur te zijn. Dan zal een goed getraind algoritme waarschijnlijk de voorspelling doen dat iemand met een uitkering én een Maserati óók een fraudeur is, ook al heeft het algoritme deze specifieke casus niet langs zien komen.

De crux zit hem in het goed getraind zijn van het algoritme: wat zijn de kenmerken van het goed getraind zijn? Hoe weet ik dat het netwerk goed generaliseert? Wordt er niet, direct of indirect, gediscrimineerd op variabelen zoals geslacht, afkomst of opleiding? Zo lijkt een algoritme dat gebruikt wordt om overtreders te straffen in Wisconsin, zwarten significant zwaarder te bestraffen dan blanken. Goed trainen van een netwerk betekent niet alleen kijken of het algoritme goede voorspellingen kan doen maar ook of het niet teveel 'bias' bevat, onbedoeld op kenmerken selecteert waarop we juist niet willen letten, althans niet in die situatie.



Figuur 1. Globaal overzicht en classificatie van case-based algoritmen. Dit overzicht is niet uitputtend.

Uitlegbaarheid van algoritmen

De reden dat we niet altijd goed weten of een algoritme 'last heeft' van onbedoelde zaken ligt in het feit dat algoritmen niet altijd even makkelijk te doorgronden zijn, zelfs niet altijd door statistici en informatici. Dat betekent dat nagedacht moet worden over andere verantwoordingsmodellen die meer ingaan op de manier waarop de algoritmes zijn getraind, welke gegevens daarvoor zijn gebruikt en de uitlegbaarheid van dat hele proces. Omdat verantwoording essentieel is voor een bestuursorgaan dat besluiten neemt, zijn we in dit onderzoek ook geïnteresseerd hoe een

bestuursorgaan organiseert dat het inzicht kan geven in de werking van een algoritme dat niet altijd even inzichtelijk is.

Worden algoritmes ingezet bij de overheid?

Algoritmes worden zeker ingezet door de overheid. Zo gebruikt de Landelijke Stuurgroep Interventies het Systeem Risico Indicatie (SyRI) om fraudeurs op te sporen, de belastingdienst experimenteert met allerlei algoritmen om foutieve aangiften op te kunnen sporen, de politie is bezig met het voorspellen van waar en wanneer door wie een misdaad zal plaatsvinden en zo verder. Dus ja, algoritmes worden door de overheid ingezet maar we weten nauwelijks door welke overheidsorganisaties, met welk doel en met welke resultaten.

Scope

Om de scope van dit onderzoek te beperken hebben we het in dit onderzoek alleen over algoritmen die de basis vormen voor besluiten over personen of anderszins een behoorlijke impact hebben op een persoon of groepen personen. Ook algoritmen die impact hebben op groepen personen vallen onder de scope van dit onderzoek. Als ergens een dijk wordt volgestopt met sensoren (zoals in Groningen) en een algoritme wordt ingezet om te voorspellen of de dijk op doorbreken staat, valt dat onder de scope omdat het resultaat van het algoritme impact heeft op groepen personen.

Mocht u een externe partij voor uw algoritmen inzetten dan valt dit ook binnen de scope omdat uw organisatie aangesproken kan worden.

Het gaat om algoritmes die worden ingezet in productie en onderzoek.

Wat willen we weten?

Om maar een begin van beleid te kunnen vormen op het gebied van algoritmen is het nodig om een beeld te hebben welke overheidspartijen algoritmen inzetten, met welk doel en met welk succes. Omdat het in veel gevallen een relatief nieuw terrein betreft is het ook nodig om een beeld te krijgen van organisaties die hiernaar een onderzoek doen of ermee aan het experimenteren zijn. Als laatste willen we weten of er behoefte bestaat aan beleid op dit terrein en zo ja, welk. Dat leidt tot de volgende drie hoofdvragen:

1. Zet uw organisatie algoritmen in, welke typen, met welke doelen en met welk succes?
2. Kan de organisatie inzicht geven over de werking van het algoritme?
3. Heeft u behoefte (of juist niet) aan beleid op het gebied van inzet van algoritmen?

De vragenlijst wordt tezamen met dit document meegestuurd. De uitleg vind u in de bijlage.

Bijlage. Toelichting op de Vragenlijst

Deze bijlage bevat de toelichting op de vragen. Afhankelijk van de beantwoording op vraag 1, kan vraag 2 al dan niet beantwoord worden. We vragen u om in alle gevallen vraag 3 te beantwoorden.

I. Past uw organisatie algoritmes toe en zo ja, met welk doel?

- Welke algoritmes gebruikt uw organisatie in de primaire processen? Gebruik daarvoor eventueel de illustratie in figuur 1. Dat overzicht is niet uitputtend en zo uw organisatie algoritmes test of inzet waarvan u vindt dat ze in het overzicht ontbreken dan kunt u deze hier melden.
- Per algoritme, wat is het doel dat ermee beoogd wordt, het resultaat en eventuele ongewenste (neven)effecten?
- Doet uw organisatie onderzoek naar algoritmes? Welke doelen zou u met de onderzochte algoritmes willen bereiken?

De volgende vraag betreft informatie over het algoritme zelf en uw communicatie erover. We vragen u dit voor maximaal twee algoritmes te beantwoorden. We zijn geïnteresseerd in met name de meer "moeilijk uitlegbare" algoritmes.

We verzoeken u uit de algoritmes die u toepast in onderzoek of productie, twee door u gebruikte algoritmes te kiezen die wat moeilijker uit te leggen zijn: één rule-based algoritme en één case-based algoritme. Mocht u geen case-based algoritmen gebruiken dan graag twee rule-based kiezen en omgekeerd: twee case-based indien geen rule-based wordt gebruikt.

Wat de Kamer graag wil weten of het algoritme op een goede manier wordt ingezet. De vragen zijn niet bedoeld als controle of u uw werk goed doet. We zijn juist geïnteresseerd in de maatregelen die u genomen heeft om u zich van de correcte werking van het algoritme te overtuigen.

De reden dat deze vragen gesteld worden ligt in het feit dat algoritmen niet altijd even makkelijk te doorgronden zijn, zelfs niet altijd voor statistici en informatici.

Dat betekent dat nagedacht moet worden over de uitleg over algoritmes. Denk aan de manier waarop de algoritmes zijn getraind, welke gegevens daarvoor zijn gebruikt, de code en de uitlegbaarheid van dat hele proces. Omdat uitleg en inzicht essentieel is voor een bestuursorgaan dat besluiten neemt, zijn we in dit onderzoek geïnteresseerd hoe een bestuursorgaan communiceert over een algoritme dat niet altijd even inzichtelijk is.

II. Inzicht in het algoritme

- Hoe legt u het algoritme uit aan een geïnteresseerde leek?
- Kunt u beschrijven welke training- en validatiegegevens zijn gebruikt? Zijn deze openbaar of bent u bereid deze openbaar te maken? Waarom niet?
- Hoe heeft u bij de ontwikkeling concepten als privacy by design, ethics by design of responsible innovation toegepast?
- Hoe toetst u of er indirect op verborgen kenmerken wordt geselecteerd?
- Op welke gronden heeft uw organisatie besloten om juist dit algoritme te gebruiken?
- Hoe wordt het algoritme op de juiste werking geaudit?
    - Is het wetenschappelijk onderbouwd?
    - Is er een gedetailleerde omschrijving van het algoritme?
    - Zijn de implementaties van de algoritmen open-source?
    - Zijn training en validatie sets open data?
- Op welke manier worden het gebruik van algoritmen, het algoritme zelf, de methodologie, het proces en de uitkomst transparant gemaakt voor eindgebruikers?
- Kan de werking van het algoritme worden geverifieerd op basis van het algoritme, de implementatie ervan en de gebruikte data? Door zowel rechters, toezichthouders als burgers?
- Welke positieve en (wellicht onverwacht) negatieve effecten heeft het gebruik van het algoritme?

III. heeft u behoefte aan beleid op het gebied van algoritmen?

- Heeft u behoefte aan beleid of regelgeving op het gebied van algoritmen en zo ja, welk?
- Of heeft u juist liever geen beleid op bepaalde terreinen, zo ja, welke?

Vragenlijst onderzoek algoritmen

We vragen u onderstaande vragenlijst in te vullen. Afhankelijk van de beantwoording op vraag 1, kan vraag 2 al dan niet beantwoord worden. We vragen u om in alle gevallen vraag 3 te beantwoorden.

U kunt de ingevulde vragenlijst, uiterlijk 25 april, mailen aan: [redacted]@minbzk.nl

Mocht u vragen hebben, kunt u die stellen op dat email adres, of bellen op [redacted].

I. Past uw organisatie algoritmes toe en zo ja, met welk doel?

I.a Welke algoritmes gebruikt uw organisatie in de primaire processen? Gebruik daarvoor eventueel de illustratie in figuur 1 van de toelichting. Dat overzicht is niet uitputtend en zo uw organisatie algoritmes test of inzet waarvan u vindt dat ze in het overzicht ontbreken dan kunt u deze hier melden.

I.b Per algoritme, wat is het doel dat ermee beoogd wordt, het resultaat en eventuele ongewenste (neven)effecten?

I.c Doet uw organisatie onderzoek naar algoritmes? Welke doelen zou u met de onderzochte algoritmes willen bereiken?

II. Inzicht in het algoritme

II.a Hoe legt u het algoritme uit aan een geïnteresseerde leek?

II.b Kunt u beschrijven welke training- en validatiegegevens zijn gebruikt? Zijn deze openbaar of bent u bereid deze openbaar te maken? Waarom niet?

II.c Hoe heeft u bij de ontwikkeling concepten als privacy by design, ethics by design of responsible innovation toegepast?

II.d Hoe toetst u of er indirect op verborgen kenmerken wordt geselecteerd?

II.e Op welke gronden heeft uw organisatie besloten om juist dit algoritme te gebruiken?

II.f. Hoe wordt het algoritme op de juiste werking geaudit?

II.f.1 Is het wetenschappelijk onderbouwd?

II.f.2 Is er een gedetailleerde omschrijving van het algoritme?

II.f.3 Zijn de implementaties van de algoritmen open-source?

II.f.4 Zijn training en validatie sets open data?

II.g Op welke manier worden het gebruik van algoritmen, het algoritme zelf, de methodologie, het proces en de uitkomst transparant gemaakt voor eindgebruikers?

II.h Kan de werking van het algoritme worden geverifieerd op basis van het algoritme, de implementatie ervan en de gebruikte data? Door zowel rechters, toezichthouders als burgers?

II.i Welke positieve en (wellicht onverwacht) negatieve effecten heeft het gebruik van het algoritme?

III. heeft u behoefte aan beleid op het gebied van algoritmen?

III.a Heeft u behoefte aan beleid of regelgeving op het gebied van algoritmen en zo ja, welk?

III.b Of heeft u juist liever geen beleid op bepaalde terreinen, zo ja, welke?

Document 6

Ministerie van Binnenlandse Zaken en
Koninkrijksrelaties

Vereniging van Nederlandse Gemeenten
t.a.v. dhr. [redacted]
Postbus 30435
2500 GK Den Haag

www.rijksoverheid.nl
www.facebook.com/minbzk
www.twitter.com/minbzk

Kenmerk
2018-0000190690

Uw kenmerk

Datum 28 MRT 2018

Betreft Verzoek om medewerking aan onderzoek naar algoritmen

Geachte heer [redacted]

Bij deze vraag ik uw medewerking aan een onderzoek dat ik laat uitvoeren naar de toepassing van algoritmes binnen de overheid.

Naar aanleiding van een overleg met de Tweede Kamer over de toepassing van algoritmen door de overheid, wordt door het ministerie van Binnenlandse Zaken en Koninkrijksrelaties een onderzoek uitgevoerd. Doel van het onderzoek is inzicht te verwerven in welke bestuursorganen welke algoritmen toepassen, of bestuursorganen onderzoek doen naar algoritmen en hoe deze zich kunnen verantwoorden over de toepassing van deze algoritmen. Het onderzoek wordt uitgezet bij departementen, hun ZBO's, provincies, gemeenten en waterschappen. De uitkomsten zullen later dit jaar aan de Tweede Kamer worden verzonden.

Graag vraag ik uw medewerking bij de uitvoering van het onderzoek door het uitzetten van de vragenlijst bij uw leden. Ik stel het zeer op prijs als de ingevulde vragenlijst uiterlijk 25 april wordt geretourneerd aan het ministerie (zie bijlage voor contactgegevens en aanspreekpunt voor eventuele vragen).

De uitkomsten van dit onderzoek zijn van belang voor zowel Rijk als gemeenten, provincies en waterschappen. Daarom nodig ik u bij deze uit om uw organisatie deel te laten nemen aan een begeleidingsgroep die mede op basis van de uitkomsten van het onderzoek een visie en mogelijk beleid rondom de toepassing van algoritmen voorbereidt. Gedacht wordt aan iemand die goed is ingevoerd in het IT-beleid van uw organisatie en in staat is de consequenties van intelligente algoritmes te doordenken. Ik dank u bij voorbaat voor uw medewerking.

De staatssecretaris van Binnenlandse zaken en Koninkrijksrelaties,
namens deze,

[redacted]

Simone Roos
Directeur-generaal Overheidsorganisatie

Vragenlijst naar het gebruik van algoritmen binnen de overheid

Achtergrond en toelichting

Aanleiding

In het 30-leden debat van 4 oktober[1] over de Digitale Infrastructuur is gesproken over de mogelijkheden van algoritmen en het gebruik ervan door de overheid. Door de leden Jetten (D66) en Bruins-Slot (CDA) is een aantal zorgen geuit over de inzet van algoritmen. Daarom heeft de staatssecretaris van BZK toegezegd om een onderzoek uit te voeren naar de toepassing van algoritmen binnen de overheid. Hij concentreert zich daarbij op twee hoofdvragen:

1. Welke overheidsorganen gebruiken welke algoritmen?
2. Gegeven de complexiteit , en soms het "black-box" karakter van een aantal ervan, kunt u inzicht geven in de werking van deze algoritmen en hoe u er over communiceert?

Op basis daarvan is deze vragenlijst opgezet en uitgezet bij de CIO's van departementen, provincies, gemeenten en waterschappen. De uitkomsten worden benut voor de vorming van een beleidsvisie op algoritmen en onderdeel van de Nationale Agenda Digitalisering Openbaar Bestuur. De uitkomsten worden tevens geanalyseerd en aan de Tweede Kamer gemeld.

Wat zijn algoritmen?

Een algoritme is een eindige reeks instructies die vanuit een gegeven begintoestand naar een beoogd doel leidt. Algoritmes zijn vergelijkbaar met een recept. Door een aantal stappen uit te voeren ontstaat er een (hopelijk) lekkere maaltijd. Algoritmen staan in beginsel los van computerprogramma's, maar in dit voorstel hebben we het uitsluitend over algoritmes die door computers worden gebruikt.

In de informatica is elke verzameling instructies een algoritme. Dat is niet wat we bedoelen: het berekenen van een gemiddelde is ook een algoritme en dat vinden we in deze context niet relevant. Wat tijdens het debat in het algemeen bedoeld werd met algoritmen zijn "intelligente" algoritmen. Die worden meestal in twee typen onderverdeeld: rule-based en case-based.

1. Rule-based systemen

Rule-based: algoritmen die op basis van een aantal regels tot een bepaalde conclusie komen. Het betreft vaak top-down regels die worden gecreëerd voor een systeem, dat daar vervolgens over kan doorredeneren.

De basis van de meeste rule-based systemen is gelegen in als-dan redeneringen in de trant van: als iemand jonger is dan 18 jaar, dan mag hij niet de supermarkt verlaten met drank. Wet- en regelgeving laat zich relatief eenvoudig vertalen in deze regels. Rule-based systemen hebben als voordeel dat regels, redeneringen en aannames expliciet moeten worden gemaakt. Dit heeft meestal als gevolg dat de manier waarop een systeem tot zijn conclusie komt expliciet kan worden gemaakt. Wat vaak ook als een voordeel wordt gezien is de modulariteit van de aanpak en uniformiteit van de regels, namelijk alle regels worden uitgeschreven in een zelfde soort format wat de opbouw van een kennisbase vereenvoudigt.

Nadelen zijn er ook: het grootste gevaar is misschien wel dat de kennisbase erg groot en complex wordt waardoor nieuwe kennis soms contradictoir is met bestaande kennis. Een ander gevolg van zo'n grote kennisbase is dat niet altijd duidelijk is wat de precieze effecten van het veranderen van één regel op alle mogelijke uitkomsten zijn, vanwege het grote aantal regels dat afhankelijk is van die ene regel.

---

[1] 4 oktober 2017, Tweede Kamer, Digitalisering infrastructuur, TK 8, 8-11-2

Case-based: de training van algoritmen

Case-based algoritmen zijn algoritmen die op basis van een aantal casussen geleerd hebben en voorspellingen kunnen doen over nog niet bekende gevallen. Dit veld is nog relatief nieuw, vergeleken met rule-based systemen. Sommige case-based algoritmen worden al veelvuldig ingezet bij de overheid, zoals lineaire regressie en, in mindere mate, beslisbomen. Andere algoritmen, zoals neurale netwerken zijn vermoedelijk nog maar in de onderzoeksfase van een paar bestuursorganen. Zie figuur 1 voor een overzicht.

Voorbeeld: stel dat een algoritme geleerd heeft dat iemand met een uitkering én een Porsche een grote kans heeft een fraudeur te zijn. Dan zal een goed getraind algoritme waarschijnlijk de voorspelling doen dat iemand met een uitkering én een Maserati óók een fraudeur is, ook al heeft het algoritme deze specifieke casus niet langs zien komen.

De crux zit hem in het goed getraind zijn van het algoritme: wat zijn de kenmerken van het goed getraind zijn? Hoe weet ik dat het netwerk goed generaliseert? Wordt er niet, direct of indirect, gediscrimineerd op variabelen zoals geslacht, afkomst of opleiding? Zo lijkt een algoritme dat gebruikt wordt om overtreders te straffen in Wisconsin, zwarten significant zwaarder te bestraffen dan blanken. Goed trainen van een netwerk betekent niet alleen kijken of het algoritme goede voorspellingen kan doen maar ook of het niet teveel 'bias' bevat, onbedoeld op kenmerken selecteert waarop we juist niet willen letten, althans niet in die situatie.



Figuur 1. Globaal overzicht en classificatie van case-based algoritmen. Dit overzicht is niet uitputtend.

Uitlegbaarheid van algoritmen

De reden dat we niet altijd goed weten of een algoritme 'last heeft' van onbedoelde zaken ligt in het feit dat algoritmen niet altijd even makkelijk te doorgronden zijn, zelfs niet altijd door statistici en informatici. Dat betekent dat nagedacht moet worden over andere verantwoordingsmodellen die meer ingaan op de manier waarop de algoritmes zijn getraind, welke gegevens daarvoor zijn gebruikt en de uitlegbaarheid van dat hele proces. Omdat verantwoording essentieel is voor een bestuursorgaan dat besluiten neemt, zijn we in dit onderzoek ook geïnteresseerd hoe een

bestuursorgaan organiseert dat het inzicht kan geven in de werking van een algoritme dat niet altijd even inzichtelijk is.

Worden algoritmes ingezet bij de overheid?

Algoritmes worden zeker ingezet door de overheid. Zo gebruikt de Landelijke Stuurgroep Interventies het Systeem Risico Indicatie (SyRI) om fraudeurs op te sporen, de belastingdienst experimenteert met allerlei algoritmen om foutieve aangiften op te kunnen sporen, de politie is bezig met het voorspellen van waar en wanneer door wie een misdaad zal plaatsvinden en zo verder. Dus ja, algoritmes worden door de overheid ingezet maar we weten nauwelijks door welke overheidsorganisaties, met welk doel en met welke resultaten.

Scope

Om de scope van dit onderzoek te beperken hebben we het in dit onderzoek alleen over algoritmen die de basis vormen voor besluiten over personen of anderszins een behoorlijke impact hebben op een persoon of groepen personen. Ook algoritmen die impact hebben op groepen personen vallen onder de scope van dit onderzoek. Als ergens een dijk wordt volgestopt met sensoren (zoals in Groningen) en een algoritme wordt ingezet om te voorspellen of de dijk op doorbreken staat, valt dat onder de scope omdat het resultaat van het algoritme impact heeft op groepen personen.

Mocht u een externe partij voor uw algoritmen inzetten dan valt dit ook binnen de scope omdat uw organisatie aangesproken kan worden.

Het gaat om algoritmes die worden ingezet in productie en onderzoek.

Wat willen we weten?

Om maar een begin van beleid te kunnen vormen op het gebied van algoritmen is het nodig om een beeld te hebben welke overheidspartijen algoritmen inzetten, met welk doel en met welk succes. Omdat het in veel gevallen een relatief nieuw terrein betreft is het ook nodig om een beeld te krijgen van organisaties die hiernaar een onderzoek doen of ermee aan het experimenteren zijn. Als laatste willen we weten of er behoefte bestaat aan beleid op dit terrein en zo ja, welk. Dat leidt tot de volgende drie hoofdvragen:

1. Zet uw organisatie algoritmen in, welke typen, met welke doelen en met welk succes?
2. Kan de organisatie inzicht geven over de werking van het algoritme?
3. Heeft u behoefte (of juist niet) aan beleid op het gebied van inzet van algoritmen?

De vragenlijst wordt tezamen met dit document meegestuurd. De uitleg vind u in de bijlage.

Bijlage. Toelichting op de Vragenlijst

Deze bijlage bevat de toelichting op de vragen. Afhankelijk van de beantwoording op vraag 1, kan vraag 2 al dan niet beantwoord worden. We vragen u om in alle gevallen vraag 3 te beantwoorden.

I. Past uw organisatie algoritmes toe en zo ja, met welk doel?

- Welke algoritmes gebruikt uw organisatie in de primaire processen? Gebruik daarvoor eventueel de illustratie in figuur 1. Dat overzicht is niet uitputtend en zo uw organisatie algoritmes test of inzet waarvan u vindt dat ze in het overzicht ontbreken dan kunt u deze hier melden.
- Per algoritme, wat is het doel dat ermee beoogd wordt, het resultaat en eventuele ongewenste (neven)effecten?
- Doet uw organisatie onderzoek naar algoritmes? Welke doelen zou u met de onderzochte algoritmes willen bereiken?

De volgende vraag betreft informatie over het algoritme zelf en uw communicatie erover. We vragen u dit voor maximaal twee algoritmes te beantwoorden. We zijn geïnteresseerd in met name de meer "moeilijk uitlegbare" algoritmes.

We verzoeken u uit de algoritmes die u toepast in onderzoek of productie, twee door u gebruikte algoritmes te kiezen die wat moeilijker uit te leggen zijn: één rule-based algoritme en één case-based algoritme. Mocht u geen case-based algoritmen gebruiken dan graag twee rule-based kiezen en omgekeerd: twee case-based indien geen rule-based wordt gebruikt.

Wat de Kamer graag wil weten of het algoritme op een goede manier wordt ingezet. De vragen zijn niet bedoeld als controle of u uw werk goed doet. We zijn juist geïnteresseerd in de maatregelen die u genomen heeft om u zich van de correcte werking van het algoritme te overtuigen.

De reden dat deze vragen gesteld worden ligt in het feit dat algoritmen niet altijd even makkelijk te doorgronden zijn, zelfs niet altijd voor statistici en informatici.

Dat betekent dat nagedacht moet worden over de uitleg over algoritmes. Denk aan de manier waarop de algoritmes zijn getraind, welke gegevens daarvoor zijn gebruikt, de code en de uitlegbaarheid van dat hele proces. Omdat uitleg en inzicht essentieel is voor een bestuursorgaan dat besluiten neemt, zijn we in dit onderzoek geïnteresseerd hoe een bestuursorgaan communiceert over een algoritme dat niet altijd even inzichtelijk is.

II. Inzicht in het algoritme

- Hoe legt u het algoritme uit aan een geïnteresseerde leek?
- Kunt u beschrijven welke training- en validatiegegevens zijn gebruikt? Zijn deze openbaar of bent u bereid deze openbaar te maken? Waarom niet?
- Hoe heeft u bij de ontwikkeling concepten als privacy by design, ethics by design of responsible innovation toegepast?
- Hoe toetst u of er indirect op verborgen kenmerken wordt geselecteerd?
- Op welke gronden heeft uw organisatie besloten om juist dit algoritme te gebruiken?
- Hoe wordt het algoritme op de juiste werking geaudit?
    - Is het wetenschappelijk onderbouwd?
    - Is er een gedetailleerde omschrijving van het algoritme?
    - Zijn de implementaties van de algoritmen open-source?
    - Zijn training en validatie sets open data?
- Op welke manier worden het gebruik van algoritmen, het algoritme zelf, de methodologie, het proces en de uitkomst transparant gemaakt voor eindgebruikers?
- Kan de werking van het algoritme worden geverifieerd op basis van het algoritme, de implementatie ervan en de gebruikte data? Door zowel rechters, toezichthouders als burgers?
- Welke positieve en (wellicht onverwacht) negatieve effecten heeft het gebruik van het algoritme?

III. heeft u behoefte aan beleid op het gebied van algoritmen?

- Heeft u behoefte aan beleid of regelgeving op het gebied van algoritmen en zo ja, welk?
- Of heeft u juist liever geen beleid op bepaalde terreinen, zo ja, welke?

Vragenlijst onderzoek algoritmen

We vragen u onderstaande vragenlijst in te vullen. Afhankelijk van de beantwoording op vraag 1, kan vraag 2 al dan niet beantwoord worden. We vragen u om in alle gevallen vraag 3 te beantwoorden.

U kunt de ingevulde vragenlijst, uiterlijk 25 april, mailen aan: [redacted]@minbzk.nl

Mocht u vragen hebben, kunt u die stellen op dat email adres, of bellen op [redacted].

I. Past uw organisatie algoritmes toe en zo ja, met welk doel?

I.a Welke algoritmes gebruikt uw organisatie in de primaire processen? Gebruik daarvoor eventueel de illustratie in figuur 1 van de toelichting. Dat overzicht is niet uitputtend en zo uw organisatie algoritmes test of inzet waarvan u vindt dat ze in het overzicht ontbreken dan kunt u deze hier melden.

I.b Per algoritme, wat is het doel dat ermee beoogd wordt, het resultaat en eventuele ongewenste (neven)effecten?

I.c Doet uw organisatie onderzoek naar algoritmes? Welke doelen zou u met de onderzochte algoritmes willen bereiken?

II. Inzicht in het algoritme

II.a Hoe legt u het algoritme uit aan een geïnteresseerde leek?

II.b Kunt u beschrijven welke training- en validatiegegevens zijn gebruikt? Zijn deze openbaar of bent u bereid deze openbaar te maken? Waarom niet?

II.c Hoe heeft u bij de ontwikkeling concepten als privacy by design, ethics by design of responsible innovation toegepast?

II.d Hoe toetst u of er indirect op verborgen kenmerken wordt geselecteerd?

II.e Op welke gronden heeft uw organisatie besloten om juist dit algoritme te gebruiken?

II.f. Hoe wordt het algoritme op de juiste werking geaudit?

II.f.1 Is het wetenschappelijk onderbouwd?

II.f.2 Is er een gedetailleerde omschrijving van het algoritme?

II.f.3 Zijn de implementaties van de algoritmen open-source?

II.f.4 Zijn training en validatie sets open data?

II.g Op welke manier worden het gebruik van algoritmen, het algoritme zelf, de methodologie, het proces en de uitkomst transparant gemaakt voor eindgebruikers?

II.h Kan de werking van het algoritme worden geverifieerd op basis van het algoritme, de implementatie ervan en de gebruikte data? Door zowel rechters, toezichthouders als burgers?

II.i Welke positieve en (wellicht onverwacht) negatieve effecten heeft het gebruik van het algoritme?

III. heeft u behoefte aan beleid op het gebied van algoritmen?

III.a Heeft u behoefte aan beleid of regelgeving op het gebied van algoritmen en zo ja, welk?

III.b Of heeft u juist liever geen beleid op bepaalde terreinen, zo ja, welke?

Centraal Bureau
voor de Statistiek

Algemene voorwaarden

Algemene voorwaarden van Centraal Bureau voor de Statistiek

1 Begrippen

1.1 Opdrachtnemer: Het Centraal Bureau voor de Statistiek, afkorting: “CBS”.

1.2 Opdrachtgever: de contractspartij van de Opdrachtnemer.

1.3 Partijen: Opdrachtnemer en Opdrachtgever.

1.4 Offerte: het voorstel voor de uitvoering van betaald werk in opdracht inclusief bijbehorende bijlagen.

1.5 Overeenkomst: de Offerte waarover Partijen overeenstemming hebben bereikt, waarop deze Algemene voorwaarden van toepassing zijn.

1.6 Microdata: data ten behoeve van statistisch of wetenschappelijk onderzoek door CBS vervaardigd met betrekking waartoe passende maatregelen zijn genomen om herkenning van afzonderlijke personen, huishoudens, ondernemingen of instellingen te voorkomen, zoals bedoeld in artikel 41 van de Wet op het Centraal Bureau voor de Statistiek, nader te noemen; “CBS-wet”.

2 Toepassing

2.1 Deze Algemene voorwaarden zijn van toepassing op iedere Overeenkomst tussen Opdrachtgever en Opdrachtnemer tenzij Partijen schriftelijk anders zijn overeengekomen.

2.2 De toepasselijkheid van algemene voorwaarden die de Opdrachtgever gebruikt, wordt uitdrukkelijk van de hand gewezen.

3 Geldigheid aanbod en totstandkoming Overeenkomst

3.1 De Overeenkomst tussen Partijen komt tot stand op het moment dat Partijen overeenstemming hebben bereikt over de Offerte.

3.2 Partijen kunnen in de Overeenkomst van deze Algemene voorwaarden afwijken. In geval van strijd tussen de Overeenkomst en de Algemene voorwaarden prevaleert de Overeenkomst.

Centraal Bureau
voor de Statistiek

4 Offertetraject

4.1 Indien een Offerte niet leidt tot een Overeenkomst, worden geen kosten in rekening gebracht voor het uitbrengen van de Offerte.

4.2 In de Offerte staat vermeld voor welke termijn de Offerte geldt. Indien Opdrachtgever de Offerte binnen de gestelde termijn niet heeft ondertekend en geretourneerd aan Opdrachtnemer, heeft de Opdrachtnemer de mogelijkheid om de Offerte in te trekken.

5 Uitvoering van de Overeenkomst

5.1 Opdrachtnemer zal de Overeenkomst zorgvuldig, vakbekwaam en onafhankelijk uitvoeren met inachtneming van de CBS-wet.

5.2 Het verklaren van verschillen tussen de onderzoeksresultaten en andere (CBS) publicaties valt buiten het bereik van de Overeenkomst, voor zover dit niet is overeengekomen in de Offerte.

5.3 Opdrachtnemer bepaalt de methode die wordt gebruikt voor het onderzoek alsmede de uitkomsten van het onderzoek.

5.4 Opdrachtnemer doet geen beleidsmatige uitspraken over de uitkomsten van het onderzoek dat hij verricht binnen de Overeenkomst.

5.5 Opdrachtnemer verzamelt gegevens om statistieken te maken en Opdrachtnemer maakt gebruik van deze statistieken om onderzoek uit te voeren. Opdrachtnemer beschikt daarom over informatie van burgers, bedrijven en organisaties. De juridische kaders waarbinnen Opdrachtnemer met privacygevoelige informatie om mag gaan zijn strikt geregeld in de CBS-wet en de Wet bescherming personengegevens (Wbp). Aanvullend heeft Opdrachtnemer een eigen gedragscode. Opdrachtnemer heeft een functionaris voor de gegevensbescherming. Deze persoon waakt over de naleving door Opdrachtnemer van de wettelijke bepalingen inzake de privacybescherming en houdt daarnaast een meldingenregister bij van de verwerkingen van persoonsgegevens.

5.6 Wanneer Opdrachtnemer voor de uitvoering van een onderzoek gebruikt maakt van een bestand dat door Opdrachtgever is aangeleverd, dan is Opdrachtgever verantwoordelijk voor de kwaliteit van de geleverde gegevens. Opdrachtnemer voert geen kwaliteitscontroles uit, tenzij tussen Partijen schriftelijk anders is overeengekomen.

Centraal Bureau
voor de Statistiek

6 Kwaliteit en aansprakelijkheid

6.1 In het geval van eventuele tekortkomingen bij de uitvoering van de Opdracht, voor zover deze het gevolg zijn van het niet in acht nemen van de zorgvuldigheid en deskundigheid waarop daarbij vertrouwd mag worden, zal Opdrachtnemer de onderzoeksresultaten kosteloos corrigeren.

6.2 Iedere aansprakelijkheid van Opdrachtnemer voor eventuele door Opdrachtgever of derden geleden schade is uitgesloten, tenzij er sprake is van opzet of grove schuld aan de zijde van Opdrachtnemer.

6.3 Indien aan de zijde van Opdrachtgever of van derden niet tijdig dan wel onjuiste of gebrekkige gegevens worden verstrekt, zal Opdrachtnemer voor de gevolgen niet aansprakelijk zijn.

6.4 Opdrachtgever is ervoor verantwoordelijk dat de gegevens die hij verstrekt aan Opdrachtnemer op een rechtmatige wijze zijn verkregen en worden overgedragen conform geldende wet- en regelgeving inzake privacy en intellectueel eigendom, waaronder in ieder geval de Wet bescherming persoonsgegevens en de Auteurswet. Opdrachtgever vrijwaart Opdrachtnemer voor iedere aansprakelijkheid voortvloeiend uit het niet of niet volledig naleven van bovengenoemde wet- en regelgeving.

7 Levertermijn en levering

7.1 In de Overeenkomst is de opleveringstermijn opgenomen van de onderzoeksresultaten en andere afgesproken producten, indien de termijn niet wordt gehaald stelt Opdrachtnemer direct Opdrachtgever op de hoogte.

7.2 Van de opleveringstermijn kan door Opdrachtnemer worden afgeweken als er bijvoorbeeld sprake is van een tussentijdse wijzigingen in de opdracht, meerwerk, niet tijdige, onjuiste en/of gebrekkige toelevering van gegevens en/of informatie door Opdrachtgever. Bovenstaande opsomming van voorbeelden is niet limitatief.

7.3 Een verzoek van Opdrachtgever om een overeengekomen planning te wijzingen, kan slechts gehonoreerd worden indien daartoe voldoende mogelijkheden zijn binnen de totale planning van CBS.

7.4 Tot 30 dagen gerekend vanaf de dag na oplevering van de onderzoeksresultaten kan Opdrachtgever zonder aanvullende kosten om nadere toelichting op de resultaten vragen, voor zover voor de beantwoording van deze vragen geen nadere analyse nodig is en het verzoek binnen de grenzen van redelijkheid en billijkheid vallen. Voor het beantwoorden van vragen waarvoor nadere analyse nodig is of die na de termijn gesteld worden, worden kosten in rekening gebracht. Hiervoor zal Opdrachtnemer van te voren een prijsopgave geven.

Centraal Bureau
voor de Statistiek

8 Meer en minderwerk

8.1 Indien door aanvullende wensen of gewijzigde inzichten van Opdrachtgever de prestaties die Opdrachtnemer op grond van de Overeenkomst moet verrichten, aantoonbaar worden verzwaard dan wel uitgebreid, is sprake van meerwerk, dat voor vergoeding in aanmerking komt. Tot meerwerk worden niet gerekend aanvullende werkzaamheden of gewijzigde inzichten die Opdrachtnemer bij het sluiten van de Overeenkomst had behoren te voorzien.

8.2 Indien door gewijzigde inzichten van Opdrachtgever de prestaties die Opdrachtnemer op grond van de Overeenkomst moet verrichten, aantoonbaar worden verlicht dan wel verminderd, is sprake van minderwerk, dat voor verrekening in aanmerking komt.

8.3 Indien een Partij van oordeel is dat van meerwerk of minderwerk sprake is, doet zij daarvan zo spoedig mogelijk mededeling aan de andere Partij.

8.4 Opdrachtnemer vangt niet aan met meerwerk alvorens hij daartoe schriftelijke opdracht van Opdrachtgever heeft gekregen en tussen Partijen overeenstemming is over de prijs voor het meerwerk, de planning van de werkzaamheden en de oplevertermijn.

9 Vergoeding en betaling

9.1 De vergoeding die als tegenprestatie voor de uitvoering van de Overeenkomst is verschuldigd is beschreven in de Overeenkomst.

9.2 Het door Opdrachtnemer gehanteerde uurtarief is vastgesteld voor het jaar waarin de Overeenkomst wordt gesloten. Indien de werkzaamheden plaatsvinden in het daarop volgende jaar, gelden de tarieven van dat volgende jaar. Indien deze tarieven afwijken van de reeds gemaakte afspraken, stelt Opdrachtnemer Opdrachtgever hiervan direct op de hoogte. Opdrachtgever verklaart zich akkoord met een afwijking van maximaal 5% van de oorspronkelijk afgesproken tarieven. Indien de uurtarieven meer dan 5% afwijken, heeft Opdrachtgever de mogelijkheid om binnen vier weken de Overeenkomst kosteloos te ontbinden.

9.3 De in de Overeenkomst genoemde bedragen zijn steeds in euro's en hierover wordt geen omzetbelasting (BTW) en/of andere heffingen van overheidswege geheven. Betaling aan Opdrachtnemer dient eveneens te geschieden in euro's.

9.4 Indien de Overeenkomst door Opdrachtgever voortijdig wordt beëindigd, is Opdrachtgever aan Opdrachtnemer een vergoeding verschuldigd voor de tot op dat moment verrichte werkzaamheden en gemaakte kosten.

9.5 Binnen 30 dagen na eindoplevering door Opdrachtnemer dan wel na voortijdige beëindiging van de Overeenkomst dient Opdrachtnemer een factuur met eindafrekening in bij Opdrachtgever. Na de vervaldag van de betalingstermijn van de factuur wordt de wettelijke rente in rekening gebracht.

Centraal Bureau
voor de Statistiek

10 Publicatie en intellectueel eigendom

10.1 Onderzoeksresultaten worden gratis beschikbaar gesteld via de website van Opdrachtnemer. De onderzoeksresultaten van Opdrachtnemer zijn in principe voor iedereen gelijktijdig beschikbaar en toegankelijk. Uitzonderingen hierop zijn gepubliceerd op de website van Opdrachtnemer. De onderzoeksresultaten worden vergezeld van voldoende uitleg om te kunnen begrijpen waar de cijfers betrekking op hebben.

10.2 Opdrachtnemer is eigenaar van de onderzoeksmethode en uitkomsten en behoudt het intellectuele eigendom hierover. Elk onderzoek wordt voorzien van een verantwoording van de onderzoeksmethode. Opdrachtnemer is niet verplicht de uitgewerkte onderzoeksmethode of programmatuur te delen met Opdrachtgever en/of derden.

10.3 Programmatuur die door Opdrachtnemer geschreven is voor de uitvoering van de Overeenkomst wordt tot vijf jaar na publicatie van het onderzoek bewaard door Opdrachtnemer.

10.4 Uit de verstrekte onderzoeksresultaten mag geen informatie kunnen worden herleid over individuele personen, bedrijven of organisaties. Dit betekent dat de resultaten kunnen worden afgerond en/of geheimgehouden.

10.5 Door Opdrachtnemer in opdracht van Opdrachtgever samengestelde of verrijkte Microdata worden aan Opdrachtgever ter beschikking gesteld uitsluitend in de beveiligde (remote access of on site) netwerkomgeving van Opdrachtnemer.

10.6 Voor het werken op microdata via remote access of on site dient een apart contract te worden afgesloten met de Microdataservices van Opdrachtnemer conform de vigerende voorwaarden.

10.7 Opdrachtnemer en Opdrachtgever zullen over eventuele persberichten of andere publicitaire uitingen nadere schriftelijke afspraken maken waarbij de wederzijdse communicatie-afdelingen betrokken worden. Dit laat onverlet het recht van Opdrachtnemer om desgewenst op enig moment zelfstandig te publiceren of communiceren over de onderzoeksresultaten. In voorkomende gevallen zal Opdrachtgever daarover vooraf worden geïnformeerd.

10.8 Wanneer Opdrachtgever de door Opdrachtnemer samengestelde onderzoeksresultaten verwerkt in een eigen publicatie dient Opdrachtnemer als bron te worden vermeld. Wanneer Opdrachtgever eigen bewerking op de resultaten heeft uitgevoerd, dient te worden vermeld: Eigen bewerking op basis van CBS.

Centraal Bureau
voor de Statistiek

11 Wijze van kennisgeving

Mededelingen, waaronder begrepen toezeggingen of (nadere) afspraken, van de ene aan de andere Partij van belang voor de uitvoering van de Overeenkomst, binden Partijen alleen indien ze schriftelijk door een daartoe bevoegde persoon zijn gedaan of bevestigd.

12 Schriftelijkheidsvereiste

Waar in deze Algemene voorwaarden wordt gesproken over 'schriftelijk' wordt tevens 'langs elektronische weg' verstaan.

13 Geheimhouding

13.1 Opdrachtnemer maakt hetgeen hem bij de uitvoering van de Overeenkomst ter kennis komt en waarvan hij het vertrouwelijke karakter kent of redelijkerwijs kan vermoeden op geen enkele wijze verder bekend, behalve voor zover enig wettelijke voorschrift of een uitspraak van de rechter hem tot bekendmaking daarvan verplicht.

13.2 Opdrachtnemer verplicht zijn personeel deze geheimhoudingsverplichting na te leven en staat ervoor in dat deze personen die verplichting nakomen.

14 Integriteitsverklaring

Opdrachtnemer verklaart dat hij ter verkrijging van de Opdracht personeel van Opdrachtgever generlei voordeel heeft geboden, gegeven, doen aanbieden of doen geven. Hij zal dat ook niet alsnog doen teneinde personen in dienst van Opdrachtgever te bewegen enige handeling te verrichten of na te laten.

15 Geschillen en toepasselijke recht

15.1 Ieder geschil tussen Partijen ter zake van de Overeenkomst en/of deze Algemene voorwaarden en/of andere schriftelijke afspraken wordt bij uitsluiting voorgelegd aan de daartoe bevoegde rechter in het arrondissement Den Haag, tenzij Partijen alsnog een andere vorm van geschillenbeslechting zullen overeenkomen.

15.2 Op de Overeenkomst is Nederlands recht van toepassing.

Document 8

Centraal Bureau
voor de Statistiek

Offerte

aan [redacted]
cc [redacted]
van [redacted]

onderwerp onderzoek algoritmen
16-04-2018
referentie 180288

Inleiding

In het 30-leden debat van 4 oktober 2017 over de Digitale Infrastructuur is gesproken over de mogelijkheden van algoritmen en het gebruik ervan door de overheid. Een algoritme is een eindige reeks instructies die vanuit een gegeven begintoestand naar een beoogd doel leidt. Door de leden Jetten (D66) en Bruins-Slot (CDA) zijn een aantal zorgen geuit over de inzet van algoritmen. Daarom heeft de staatssecretaris van BZK toegezegd om een onderzoek uit te voeren naar de toepassing van algoritmen binnen de overheid. Hij concentreert zich daarbij op twee hoofdvragen:

1. Welke overheidsorganen gebruiken welke algoritmen?
2. Gegeven de complexiteit , en soms het "black-box" karakter van een aantal ervan, kunt u inzicht geven in de werking van deze algoritmen en hoe u er over communiceert?

Om deze vragen te kunnen beantwoorden heeft het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK) een vragenlijst (zie bijlage 1) opgesteld. Deze vragenlijst is door BZK uitgezet bij de CIO's van departementen, zelfstandige bestuursorganen (zbo's), provincies, gemeenten en waterschappen. BZK heeft aan het CBS gevraagd om de ingevulde vragenlijsten te analyseren. De uitkomsten van het onderzoek zullen benut worden voor een onderdeel van de Nationale Agenda Digitalisering Openbaar Bestuur. De uitkomsten worden tevens geanalyseerd en aan de kamer gemeld.

Projectopdracht

Het CBS kan in het verzoek van BZK voorzien door een rapport op te stellen met de volgende onderdelen: resultaten van het onderzoek, methodebeschrijving en technische aanbevelingen ten opzichte van het onderzoek. Het verzamelen van de data heeft BZK gedaan. Door de gekozen methode van BZK is er een niet representatieve steekproef ontstaan. Dat is de reden waarom de resultaten van de analyse **alleen** betrekking zullen hebben op de overheidsorganisaties die de vragenlijst hebben ingevuld. Daarmee zijn de resultaten niet generaliseerbaar naar de gehele populatie.

Afhankelijk van de kwantiteit en kwaliteit van de respons, zal er getracht worden om de volgende vragen te beantwoorden in het rapport:

1. Hoeveel organisaties van de gerespondeerde organisaties zetten algoritmen in?
2. Welke type organisaties van de gerespondeerde organisaties zetten algoritmen in? Hierbij kan gedacht worden aan de volgende categorieën:
    a. Ministeries
    b. ZBO's
    c. Provincies
    d. Gemeenten
    e. Waterschappen
3. Welke type algoritmen worden ingezet bij de gerespondeerde organisaties? Hierbij kan gedacht worden aan de volgende categorieën:
    a. Rule-based
    b. Case-based
4. Voor welke doeleinden worden algoritmen ingezet bij de gerespondeerde organisaties?
5. Welke successen worden er geboekt met de inzet van algoritmen bij de gerespondeerde organisaties?
6. Kunnen de gerespondeerde organisaties inzicht geven over de werking van algoritmen?
7. Is er behoefte bij gerespondeerde organisaties aan beleid op het gebied van inzet van algoritmen?

Het onderzoek bestaat uit de volgende onderdelen. Het eerste onderdeel is om de ingevulde vragenlijsten te verwerken in een database. Als volgt zullen de antwoorden op de open vragen worden geclassificeerd door middel van tekstanalyse. Hierbij zullen de antwoorden op open vragen zoveel mogelijk gegroepeerd worden om verdere analyse van deze vragen te vergemakkelijken. Uiteindelijk zal er een rapport worden geschreven met een beschrijving van de resultaten, methodebeschrijving en technische aanbevelingen ten opzichte van het onderzoek.

Projectresultaat en publicatie

Het projectresultaat bestaat uit een rapport in pdf-formaat (indicatie: 10 à 15 pagina's), met de volgende onderdelen:

- Onderdeel 1: beschrijving van de belangrijkste resultaten;
- Onderdeel 2:methodebeschrijving van het onderzoek;
- Onderdeel 3: technische aanbevelingen t.o.v. het onderzoek.

De methodebeschrijving van het onderzoek (onderdeel 2) en de technische aanbevelingen t.o.v. het onderzoek (onderdeel 3) worden gepubliceerd op de website van het CBS. Beschrijving van de belangrijkste resultaten (onderdeel 1) zal **niet** worden gepubliceerd op de website van het CBS omdat het onderzoek geen gegevens van het CBS bevat. Het rapport met alle onderdelen wordt als digitale versie aangeboden aan de opdrachtgever.

Risico's en voorwaarden

Op dit projectvoorstel zijn de Algemene Voorwaarden van CBS van toepassing (zie bijlage 2). Daarnaast gelden de volgende risico's en voorwaarden.

- Indien de opdrachtgever, na aanvaarding, tussentijds wijzigingen wil aanbrengen in de projectopdracht, wordt in overleg een nieuwe capaciteits- en urenplanning vastgesteld.
- Het CBS en de opdrachtgever zullen over eventuele persberichten of andere publicitaire uitingen nadere afspraken maken waarbij de wederzijdse communicatie-afdelingen betrokken worden.
- Om onthulling van individuele organisaties te voorkomen zullen de aantallen in het rapport worden afgerond op 5-tallen.
- Het rapport wordt onder embargo geleverd. Dit betekent dat het rapport niet verder mag worden verspreid tot na de publicatiedatum. Gebeurt dit wel dan wordt het rapport direct gepubliceerd op de website van het CBS.
- Het CBS staat niet in voor de kwaliteit van de vragenlijst. De vragenlijst is opgesteld door BZK en daarmee is BZK verantwoordelijk voor de kwaliteit van de vragenlijst.
- Het CBS is niet verantwoordelijk voor het afnemen van de steekproef. BZK is verantwoordelijk voor het afnemen van de steekproef.
- De kwaliteit van de resultaten is afhankelijk van de kwantiteit en kwaliteit van de respons.
- Het onderzoek bevat een niet representatieve steekproef. Dat is de reden waarom de resultaten van de analyse **alleen** betrekking zullen hebben op de overheidsorganisaties die de vragenlijst hebben ingevuld. Daarmee zijn de resultaten niet generaliseerbaar naar de gehele populatie.

Projectplanning

Belangrijke data

Om onderstaande deadline te halen, moet dit projectvoorstel uiterlijk op **20 april 2018** zijn geaccepteerd door de opdrachtgever. Het onderzoek zal dan van start gaan op 1 mei 2018. BZK dient vóór 1 mei 2018 alle ingevulde vragenlijsten in Word aan te leveren aan het CBS die zij tot dan toe hebben ontvangen. BZK kan tot uiterlijk **21 mei 2018** een klein aantal nagekomen ingevulde vragenlijsten aanleveren bij het CBS ten behoeve van de analyse. Het CBS levert het concept rapport onder embargo dan uiterlijk **14 juni 2018** aan BZK. In de **week van 25 juni** zal er een overleg plaatsvinden met BZK. Tijdens dit overleg zal het conceptrapport besproken worden. Het definitieve rapport (met alle drie de onderdelen) wordt uiterlijk op **vrijdag 6 juli** als digitale versie aangeboden aan de opdrachtgever. Tevens zullen onderdelen 2 en 3 gepubliceerd worden op de website van het CBS.

Prijs en specificatie

Het is mogelijk om dit onderzoek uit te voeren voor **22.000 euro**.

De kosten zijn in onderstaande tabel gespecificeerd:

Activiteit	Bedrag (in euro's)
Voorbereiding en afstemming – uitwerken plan van aanpak – projectleiding	
Verwerken ingevulde vragenlijsten – maken van een database – variabelenselectie en hercoderingen – plausibiliteitscontrole	
Classificeren antwoorden open vragen – labels toekennen aan de antwoorden op open vragen	
Analyse – Uitvoeren van analyses – plausibiliteitscontrole	
Rapportage – schrijven van rapport met resultaten, methodebeschrijving en technische aanbevelingen t.o.v. het onderzoek	
Totaal	

Document 9

Ministerie van Binnenlandse Zaken en Koninkrijksrelaties

Aan DIO/DA
Van [redacted]

DGOO
DIO/DA

Contactpersoon
[redacted]

Datum
17 mei 2018

Kenmerk
2018-0000293152

nota

Opdrachtverstrekking aan CBS voor uitvoeren gegevensanalyse onderzoek algoritmen

Aanleiding
DIO wil het CBS een analyse laten uitvoeren ten behoeve van het onderzoek naar de toepassing van algoritmen door de overheid.

Advies/actie
U wordt geadviseerd de bijgevoegde opdrachtovereenkomst, ter waarde van [redacted] te tekenen. Omdat het een contract is met een overheidspartij wordt hierover geen BTW berekend.

Betrokken BZK onderdelen
DIO

Toel chting
Het doel van de opdracht is het analyseren en rapporteren van gegevens die zijn ingezameld voor het onderzoek naar de toepassing van algoritmen door de overheid.

Pol tieke context
Het onderzoek naar de toepassing van algoritmen door de overheid is de invulling van een toezegging van de staatssecretaris van BZK aan de Tweede Kamer.

Communicatie
N.v.t.

Document 11

Van: [redacted]
Aan: [redacted]
Onderwerp: Besprekingspunten BZK-CBS
Datum: maandag 23 april 2018 14:35:25

Hoi Fanny,
Vragen:

- Hoe het geregeld is dat informatie niet naar de individuele bestuursorganen terug te herleiden is?
- Kan de informatie via het CBS worden opgevraagd?
- Kan die informatie via de aanleverancier van gegevens worden opgevraagd?
- Kan die informatie bij de bron worden opgevraagd.

Vriendelijke groet,

[redacted]
Senior beleidsmedewerker
.....
Digitale Agenda, Directie Informatiesamenleving en Overheid
Directoraat-generaal Overheidsorganisaties
Ministerie van Binnenlandse Zaken en Koninkrijksrelaties
Turfmarkt 147 | 2511 DP | Den Haag | Zuid, 11e verdieping
Postbus 20011 | 2500 EA | Den Haag
.....
[redacted]
www.rijksoverheid.nl
www.facebook.com/minbzk | www.twitter.com/minbzk
.....
DGOO werkt aan een moderne overheid
.....

Document 12

[redacted]

Van: [redacted]
Verzonden: donderdag 3 mei 2018 13:07
Aan: [redacted]
Onderwerp: RE: Kontaktgegevens

Hoi [redacted]

Dank je wel! Ik heb ondertussen alle documenten die ik nu heb geupload naar CBS.

Groet,

[redacted]

Van: [redacted]
Verzonden: donderdag 3 mei 2018 11:42
Aan: [redacted]
CC: Postbus DGOO financieel beheer
Onderwerp: RE: Kontaktgegevens

Hallo [redacted]

Hierbij:

Centraal Bureau voor de Statistiek
t.a.v. [redacted]
Postbus 24500
2490 HA Den Haag

Bankgegevens:

Rabobank
Postbus 17100
3500 HG Utrecht
Rekeningnummer: [redacted]
Swiftnummer.....: [redacted]
IBANnummer.....: [redacted]
KvKnummer.....: 51197073

Met vriendelijke groet,

[redacted]
Centrum voor Beleidsstatistiek

CBS |CBS-weg 11| Postbus 4481 | 6401 CZ Heerlen
[redacted]

Kijk voor informatie over onderzoek op maat op https://www.cbs.nl/nl-nl/onze-diensten/maatwerk-en-microdata/onderzoek-op-maat

Van:
Verzonden: donderdag 3 mei 2018 10:17
Aan:
CC: Postbus DGOO financieel beheer
Onderwerp: Kontaktgegevens

Hoi

De opdracht zal binnenkort gegund worden. Daartoe hebben wij nog wat gegevens nodig:
- De naam van de opdrachtnemer aan wie de gunningsbrief wordt gestuurd
- Postadres van CBS waar de brief heen moet
- Het rekeningnr. waar de bijdrage op gestort dient te worden

Zou je die gegevens willen mailen aan Postbus DGOO financieel beheer die in de CC staat vermeld?

Bij voorbaat hartelijk dank!

Met vriendelijke groet,

Senior beleidsmedewerker
.....
Digitale Agenda, Directie Informatiesamenleving en Overheid

Directoraat-generaal Overheidsorganisaties
Ministerie van Binnenlandse Zaken en Koninkrijksrelaties
Turfmarkt 147 | 2511 DP | Den Haag | Zuid, 11e verdieping
Postbus 20011 | 2500 EA | Den Haag
.....

www.rijksoverheid.nl
www.facebook.com/minbzk | www.twitter.com/minbzk
.....
DGOO werkt aan een moderne overheid
.....

Dit bericht kan informatie bevatten die niet voor u is bestemd. Indien u niet de geadresseerde bent of dit bericht abusievelijk aan u is toegezonden, wordt u verzocht dat aan de afzender te melden en het bericht te verw jderen. De Staat aanvaardt geen aansprakel jkheid voor schade, van we ke aard ook, die verband houdt met risico's verbonden aan het elektronisch verzenden van berichten.
This message may contain information that is not intended for you. If you are not the addressee or if this message was sent to you by mistake, you are requested to inform the sender and delete the message. The State accepts no liability for damage of any kind resulting from the risks inherent in the electronic transmission of messages.

Document 13

[redacted]

Van: [redacted]@cbs.nl>
Verzonden: woensdag 13 juni 2018 15:55
Aan: [redacted]
Onderwerp: RE: Bespreken concept rapport algortimen onderzoek

Hoi [redacted]

Helder, bedankt!
4 oktober tussen 13 en 14 uur zal ik een uitnodiging voor sturen.

Groeten,
[redacted]

Van: [redacted]
Verzonden: woensdag 13 juni 2018 15:02
Aan: [redacted]
CC: [redacted]
Onderwerp: RE: Bespreken concept rapport algortimen onderzoek

Hoi [redacted]

Ik kan hier goed mee leven. 4 oktober is ok tussen 13 en 14. Nodig [redacted] hier ook voor uit (zie CC). Of een aanvullende opdracht nodig is moeten we dan maar bezien. Ik hoop van niet.

DG's zijn in principe een onderdeel van het ministerie waar ze onder vallen, uitzonderingen zijn Belastingdienst en RIVM. ZBO's en agentschappen zou ik onder de uitvoeringsorganisaties laten vallen.

Voor de statistieken: het aantal benaderde gemeenten is 230. Voor de anderen geldt dat we hebben aangenomen dat ze niets aan algoritmen doen, dat hoeft niet perse juist te zijn maar omdat ze op geen enkele wijze laten blijken dat ze iets doen aan data-analyse of statistiek is de aanname waarschijnlijk wel juist.

Kun je hier wat mee?

Vriendelijke groet,

[redacted]

Van: [redacted]@cbs.nl>
Verzonden: woensdag 13 juni 2018 14:26
Aan: [redacted]@minbzk.nl>
CC: [redacted]@cbs.nl>
Onderwerp: RE: Bespreken concept rapport algortimen onderzoek

Hoi [redacted]

Ik heb ons telefoontje van vanmorgen net even met mijn collega besproken en we stellen de volgende aangepaste planning voor.

- Wij maken komende twee weken het concept rapport af. De vragenlijsten die we deze week (t/m 15 juni) nog krijgen nemen we daarin mee, de vragenlijsten die daarna komen kunnen we niet meer mee nemen.
- Wij zullen jullie **20 september** het conceptrapport toesturen.

- Dit kunnen we dan **4 oktober** bespreken. Dan ben ik ook weer aanwezig. Hoe laat zou voor jou uitkomen? Wij kunnen tussen 10 en 16 uur.
- We publiceren het rapport dan uiterlijk **19 oktober** op onze website.

Het overleg op 4 oktober kunnen we gebruiken om het concept rapport te bespreken en feedback van jullie kant op het rapport te ontvangen.
Doordat we het project nu tijdelijk stilleggen kunnen wij dit project wel minder efficiënt uitvoeren. Ik wil daarom benadrukken dat we ons qua het verwerken van feedback zullen beperken tot het aanpassen van onjuistheden, veranderingen in formuleringen of zaken die te gevoelig liggen uit het rapport halen. Er is geen ruimte meer om dan nog extra analyses uit te voeren of het rapport volledig om te gooien. Mocht er wel behoefte zijn aan substantiële aanvullende werkzaamheden, zullen wij een nieuwe offerte opstellen.

Tot slot nog een inhoudelijke vraag zodat we verder kunnen met het conceptrapport:

We proberen de resultaten waar mogelijk in te delen naar:
- Ministeries
- ZBO's/ Uitvoeringsorganisatie/ Agentschappen
- Provincies
- Gemeenten
- Waterschappen

We lopen er hierbij tegenaan dat we soms niet weten hoe we om moeten gaan met de antwoorden van verschillende DG's. Bijvoorbeeld:

Ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK) heeft vragenlijsten gestuurd voor →
- Directoraat-generaal Algemene Inlichtingen- en Veiligheidsdienst (AIVD)
- Directoraat-generaal Bestuur, Ruimte en Wonen
- Directoraat-generaal Overheidsorganisatie
- Directoraat-generaal Vastgoed en Bedrijfsvoering Rijk
- Programma directoraat-generaal Omgevingswet

Ministerie van Financiën heeft vragenlijsten gestuurd voor →
- Directoraat-generaal Belastingdienst
- Directoraar-generaal voor Fiscale Zaken
- Directoraat-generaal Rijksbegroting

Ministerie van Infrastructuur en Waterstaat (IenW) heeft vragenlijsten gestuurd voor →
- Directoraat-generaal Luchtvaart en Maritieme Zaken
- Directoraat-generaal Mobiliteit
- Directoraat-generaal Milieu en Internationaal
- Directoraat-generaal Water en Bodem

Moeten we deze DG's behandelen als los onderdeel van de overheid of kunnen we deze antwoorden scharen onder die van de ministeries waar ze onder vallen?
Bijvoorbeeld in het geval van de belastingdienst of de AIVD? Tellen die antwoorden dan voor het ministerie waar ze onder vallen of wil je dat we deze antwoorden indelen bij ZBO's/ Uitvoeringsorganisatie/ Agentschappen.

Kun je laten weten of je akkoord met ons planningsvoorstel en zou je onze vraag kunnen beantwoorden?

Alvast heel erg bedankt en succes met alle deadlines de komende tijd!

Groeten,

Van: [redacted]@minbzk.nl>
Verzonden: woensdag 13 juni 2018 07:58
Aan: [redacted]@cbs.nl>
Onderwerp: RE: Bespreken concept rapport algortimen onderzoek

Hoi [redacted]

Dan wordt het begin september, want dan ben ik van vakantie terug. Wat dacht je van 13 september?

Groet,

[redacted]

Van: [redacted]@cbs.nl>
Verzonden: dinsdag 12 juni 2018 16:35
Aan: [redacted]@minbzk.nl>
Onderwerp: Bespreken concept rapport algortimen onderzoek

Beste [redacted]

Zullen we vast een overleg inplannen om begin jullie het conceptrapport van het algoritmen onderzoek te bespreken?
Wat ons betreft is 3, 4, 5 of 6 juli nog een mogelijkheid. Hoor graag van je wanneer het jou uitkomt!

Met vriendelijke groet,

[redacted]
Statistisch onderzoeker

CBS | Henri Faasdreef 312 | Postbus 24500 | 2490 HA Den Haag
[redacted]
Volg statistiekcbs op twitter | facebook | instagram

Voor wat er **feitelijk** gebeurt

Dit bericht kan informatie bevatten die niet voor u is bestemd. Indien u niet de geadresseerde bent of dit bericht abusievelijk aan u is toegezonden, wordt u verzocht dat aan de afzender te melden en het bericht te verwijderen. De Staat aanvaardt geen aansprakelijkheid voor schade, van welke aard ook, die verband houdt met risico's verbonden aan het elektronisch verzenden van berichten.
This message may contain information that is not intended for you. If you are not the addressee or if this message was sent to you by mistake, you are requested to inform the sender and delete the message. The State accepts no liability for damage of any kind resulting from the risks inherent in the electronic transmission of messages.

Document 14

[redacted]

Van: [redacted]
Verzonden: vrijdag 15 juni 2018 12:39
Aan: [redacted]
Onderwerp: Laatste tranche

Hoi [redacted]

Kreeg vandaag nog de vragenlijst van het UWV. Dat is het laatste wat mij betreft. Hopelijk kunnen jullie het meenemen.

Groet, [redacted]

[redacted]

Document 15

[redacted]

Van: [redacted]
Verzonden: woensdag 6 juni 2018 09:12
Aan: [redacted]
Onderwerp: RE: Planning onderzoek algoritmen

Hoi [redacted]

Vorige week is de brief verzonden. Die was al eerder verzonden maar er stond een verkeerd adres op en kwam weer terug.

Wat je nu hebt, op [redacted] na is het wel ongeveer.

Vriendelijke groet,

[redacted]

Van: [redacted]
Verzonden: dinsdag 5 juni 2018 15:46
Aan: [redacted]
Onderwerp: RE: Planning onderzoek algoritmen

Beste [redacted]

Dank voor de nieuwe lading vragenlijsten, we hebben ze inmiddels ingevoerd.
Kun jij aangeven hoeveel vragenlijsten je nu nog verwacht? [redacted]
Dan hebben wij een beetje een idee van of we al in de buurt van het totaal beeld zit.

Daarnaast vroeg ik mij af of er al een contract richting het CBS is gegaan?

Vriendelijke groet,
[redacted]

Van: [redacted]@minbzk.nl>
Verzonden: donderdag 31 mei 2018 11:11
Aan: [redacted]@cbs.nl>
Onderwerp: RE: Planning onderzoek algoritmen

Beste [redacted]

Ik heb vandaag de derde tranche geupload. Voor wat betreft het uitzetten het volgende: Ik heb aan de koepels gevraagd om het bij hun leden uit te zetten. Aan de leden is gevraagd om dit bij hun onderscheiden beleidsonderdelen uit te zetten, dus:

- Rijk: alle departementen, elke departement bij hun agentschappen en zbo's
- Alle provincies
- Alle waterschappen
- Voor gemeenten die gemeenten die deel uitmaakten van de vereniging big data. Het precieze aantal leden moet ik nog opvragen
- Ik krijg nog vragenlijsten van [redacted] dat wordt echter pas juni

Met vriendelijke groet,

[redacted]

Van: [redacted]@cbs.nl>
Verzonden: vrijdag 18 mei 2018 15:05
Aan: [redacted]@minbzk.nl>
Onderwerp: RE: Planning onderzoek algoritmen

Beste [redacted]

Ik heb nog één andere vraag.

Zou jij ons een overzicht kunnen sturen met welke organisaties (of in elk geval hoeveel) de vragenlijst is toegestuurd?
Dan weten wij straks wat de response is geweest en wat we wellicht nog kunnen verwachten.

Bedankt alvast!

Groeten,
[redacted]

Van: [redacted]
Verzonden: vrijdag 18 mei 2018 13:21
Aan: [redacted]@minbzk.nl>
CC: [redacted]@cbs.nl>; [redacted]@cbs.nl>
Onderwerp: Planning onderzoek algoritmen

Beste [redacted]

Nog even een mail n.a.v. ons telefoongesprek afgelopen woensdag.

Onderstaande data zijn genoemd in het huidige onderzoeksvoorstel:

- Het CBS levert het concept rapport onder embargo dan uiterlijk **14 juni 2018** aan BZK.
- In de **week van 25 juni** zal er een overleg plaatsvinden met BZK. Tijdens dit overleg zal het conceptrapport besproken worden.
- Het definitieve rapport (met alle drie de onderdelen) wordt uiterlijk op **vrijdag 6 juli** als digitale versie aangeboden aan de opdrachtgever.

Wij zijn een week later van start kunnen gaan dan waar de planning in het projectvoorstel vanuit ging.
Je gaf aan dat 14 juni een belangrijk moment is om tussentijdse resultaten te ontvangen. Zullen we daarom afspreken dat wij uiterlijk **14 juni een notitie met een samenvatting van de tussentijdse resultaten onder embargo** opleveren. Deze resultaten mogen dan alleen onder betrokkenen binnen BZK worden gedeeld, maar nog niet naar buiten worden gebracht.

De overige data zouden wij graag een week opschuiven. Ons voorstel is:

- Het CBS levert het concept rapport onder embargo dan uiterlijk **21 juni 2018** aan BZK.
- In de **week van 2 juli** zal er een overleg plaatsvinden met BZK. Tijdens dit overleg zal het conceptrapport besproken worden.
- Het definitieve rapport (met alle drie de onderdelen) wordt uiterlijk op **vrijdag 13 juli** als digitale versie aangeboden aan de opdrachtgever.

Zullen we meteen iets inplannen voor het overleg in de week van 2 juli? Dan staat dat maar vast. Mijn agenda biedt nog veel mogelijkheden die week.

Tot slot vroeg ik mij nog af of er inmiddels al een contract/bevestigingsbrief naar het CBS is gegaan zodat wij een officieel akkoord hebben voor onze offerte?

Met vriendelijke groet,

Statistisch onderzoeker

CBS | Henri Faasdreef 312 | Postbus 24500 | 2490 HA Den Haag
@cbs.nl
Volg statistiekcbs op twitter | facebook | instagram

Voor wat er **feitelijk** gebeurt

Dit bericht kan informatie bevatten die niet voor u is bestemd. Indien u niet de geadresseerde bent of dit bericht abusievelijk aan u is toegezonden, wordt u verzocht dat aan de afzender te melden en het bericht te verw jderen. De Staat aanvaardt geen aansprakel jkheid voor schade, van we ke aard ook, die verband houdt met risico's verbonden aan het elektronisch verzenden van berichten.
This message may contain information that is not intended for you. If you are not the addressee or if this message was sent to you by mistake, you are requested to inform the sender and delete the message. The State accepts no liability for damage of any kind resulting from the risks inherent in the electronic transmission of messages.

Document 16

[redacted]

Van: [redacted]@cbs.nl>
Verzonden: maandag 7 mei 2018 15:22
Aan: [redacted]
Onderwerp: Vragenlijst Algoritmen

Beste [redacted]

Ik ben projectleider voor het project dat CBS-CVB zal uitvoeren m.b.t. de vragenlijsten over algoritmen.

Fijn dat het gelukt is om al een deel van de vragenlijsten te uploaden! Ik heb inmiddels van zo'n 30 overheidsorganisaties de reactie ontvangen n.a.v. wat jij geüpload hebt. In verband met onze planning vroeg ik mij af of je nog meer vragenlijsten terug verwacht en zo ja hoeveel en op welke termijn kunnen wij deze verwachten?

Voor ons zou het fijn zijn dat steeds wanneer er weer vragenlijsten binnen komen deze zo snel mogelijk bij ons terecht komen. Zou je mij een seintje kunnen geven mochten er nieuwe vragenlijsten geüpload zijn, dan kan ik ze hier binnenhalen.

Alvast heel erg bedankt!

Met vriendelijke groet,

[redacted]
Statistisch onderzoeker

CBS | Henri Faasdreef 312 | Postbus 24500 | 2490 HA Den Haag
[redacted]
Volg statistiekcbs op twitter | facebook | instagram

Voor wat er **feitelijk** gebeurt