



Gegevensbeschermings- effectbeoordeling (PIA)

Doelgericht ingrijpen op een aangevraagde toeslag (M1354): Het
Risicoclassificatiemodel

Ministerie van Financiën, Belastingdienst/Toeslagen

Utrecht, november 2020

Maatregelen
nemen
Privacybewustwording
Doelbinding
PIA
Noodzaak
Effecten in kaart
Bescherming van
persoonsgegevens
Risico's
minimaliseren
Richtinggevend
Rechtsgrond
Met open vizier

Vaststelling verwerkersverantwoordelijke: Selecteer/typ datum

Naam: Typ naam/functie

Advies Selecteer/typ datum

Naam: Typ naam/functie

Advies Selecteer/typ datum

Naam: Typ naam/functie

Gegevensbeschermings- effectbeoordeling (PIA)

Ministerie van Financiën, Belastingdienst/Toeslagen
Toeslagen

Contact:

Ministerie van Financiën, directie Belastingdienst/Toeslagen

Contact Persoon:

Versie: 1.9

Inhoudsopgave

Inleiding	5
A. Beschrijving kenmerken gegevensverwerkingen	6
1. Voorstel	6
2. Persoonsgegevens	96
3. Gegevensverwerkingen	118
4. Verwerkingsdoeleinden	1412
5. Betrokken partijen	1614
6. Belangen bij de gegevensverwerking	1715
7. Verwerkingslocaties	1845
8. Techniek en methode van gegevensverwerking	1816
9. Juridisch en beleidsmatig kader	3836
10. Bewaartermijnen	3937
B. Beoordeling rechtmatigheid gegevensverwerkingen.....	4038
11. Rechtsgrond	4038
12. Bijzondere persoonsgegevens	4139
13. Doelbinding	4240
14. Noodzaak en evenredigheid	4340
15. Rechten van de betrokkene	4543
C. Beschrijving en beoordeling risico's voor de betrokkenen.....	4645
16. Risico's	4745
D. Beschrijving voorgenomen maatregelen	4849
17. Maatregelen	4849
Bijlage 1: Bronnen	
Bijlage 2: Halfproducten	
Bijlage 3: HT indicatoren	
Bijlage 4: KOT indicatoren	
Bijlage 5: Control Matrix	
Bijlage 6: Toeslagen Toezicht Werkplan 2018	

Inleiding

Per 25 mei 2018 is de General data protection regulation (GDPR) in werking getreden. Beter bekend als de Algemene verordening gegevensbescherming (AVG). In deze wet ligt vastgelegd dat het verplicht is om een PIA uit te voeren als er sprake is van geautomatiseerde grootschalig data verwerking, waaronder profilering en waarop besluiten worden gebaseerd waaraan voor de werking voor de natuurlijke persoon op vergelijkbare wijze wezenlijk treffen. De term PIA staat voor Privacy Impact Assessment. Ook wel Gegevensbeschermingseffectbeoordeling (GEB) genoemd. In dit rapport noemen we het een PIA.

Het doel van een PIA is om privacyrisico's in kaart te brengen van nieuwe projecten en initiatieven of bestaande verwerkingen van persoonsgegevens. Het draagt bij aan het vermijden of verminderen van privacyrisico's. Door middel van een PIA wordt op systematische wijze inzichtelijk gemaakt hoe groot de kans is dat de privacy van betrokkenen geschaad wordt en welke maatregelen er nodig zijn. Na het uitvoeren van de PIA heeft de organisatie antwoorden op deze vragen en kan er op deze basis een proces beschreven worden over de te nemen maatregelen als er zich calamiteiten of incidenten voordoen.

Deze PIA is uitgevoerd bij Belastingdienst/Toeslagen (B/T). Voor deze PIA wordt het proces "doelgericht ingrijpen op een aangevraagde toeslag" (binnen Toeslagen benoemd het risicoclassificatiemodel) beschreven zoals dat bij Toeslagen gebruikt wordt. Het Risicoclassificatiemodel (RCM) wordt ingezet om een inschatting te maken van het risicogehalte van toelag aanvragen, oftewel de kans dat een aanvraag na controle aangepast of afgekeurd moet worden. Deze risicoclassificatie helpt Toeslagen bij het maken van behandelkeuzes zoals bijvoorbeeld meteen uitbetalen of eerst bewijsstukken opvragen. Deze verwerking valt onder de hoofdverwerking Toezicht.

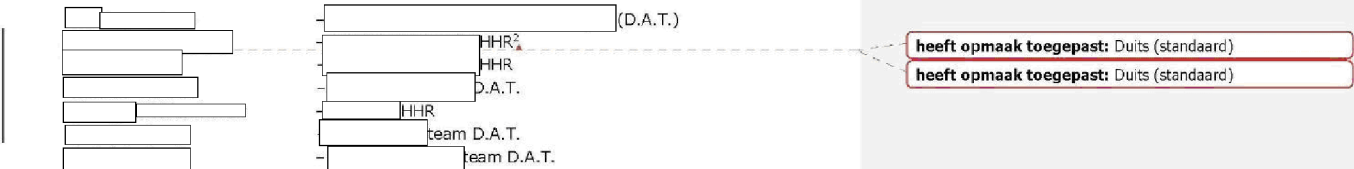
Vanwege het vertrouwelijk karakter van het Risicoclassificatiemodel worden burgers niet bij de workshop voor de PIA betrokken. We zijn wel voornemens (het wordt in Q1 2021 besproken) het concept van RCM aan een gebruikerspanel voor te leggen en daarbij te vragen wat ze van deze aanpak vinden en welke risico's (bedenkingen) zij kunnen identificeren. Hier kunnen we dan rekening mee houden.

Deze PIA wordt minimaal eens in de 3 jaar, of bij grote wijzigingen herijkt. Het gaat dan om wijzigingen aan het proces of aan systemen waarbij risico's voor betrokkenen kunnen ontstaan. Het resultaat van herijken wordt opgenomen als bijlage aan de PIA. In sommige gevallen kan er besloten om de PIA opnieuw uit te voeren. Het proces van herijken werkt als volgt. In een multidisciplinaire workshop wordt getoetst of de PIA nog de werkelijke verwerking c.q. situatie beschrijft. Daarnaast wordt gekeken naar de verwerkte gegevens en of de gegevens nog proportioneel en subsidair zijn. Verder wordt de risicoanalyse voor betrokkenen opnieuw gemaakt en toetsen we de maatregelen in de actualiteit. Jaarlijks wordt getoetst of de maatregelen ingevoerd zijn.

Aan deze PIA hebben de volgende personen bijgedragen:

	-	B/T
	-	HHR
	-	B/T
	-	team D.A.T. ¹

¹ Domain Analytics Toeslagen



A. Beschrijving kenmerken gegevensverwerkingen

Beschrijf op gestructureerde wijze de voorgenomen gegevensverwerkingen, de verwerkingsdoeleinden en de belangen bij de gegevensverwerkingen.

Onder A wordt de eerste stap beschreven van de PIA: een overzicht van de relevante feiten van de voorgenomen gegevensverwerkingen. Als de feiten onduidelijk zijn, werkt dit door in de beoordeling.

1. Voorstel

Beschrijf het voorstel waar de gegevensbeschermingseffectbeoordeling op ziet en context waarbinnen deze plaatsvindt op hoofdlijnen.

De missie van Belastingdienst/Toeslagen is het juist, tijdig en rechtmatig uitbetalen van toeslagen. Om dit zo efficiënt mogelijk te doen is handhaving onmisbaar. Wij willen daarbij maatwerk leveren door de juiste combinatie van dienstverlening, voorkomen van fouten, toezicht en fraudebestrijding. Toekenningseertheid in een zo vroeg mogelijk stadium is een belangrijk onderdeel van deze strategie.

Maandelijkse verwerkt B/T² honderdduizenden mutaties in toeslagaanvragen. Periodiek (ongeveer eens per maand) worden alle mutaties in een keer beschikt: de wijziging in de aanvraag wordt effectief gemaakt door een formele beschikking aan te maken. Na een formele beschikking wordt de toeslag uitgekeerd.

Om grip te houden op mogelijke foutieve aanvragen en mutaties, wordt vlak voor het periodieke beschikkingsproces voor alle aanvragen (waarvoor een conceptbeschikking is afgegeven) een risicoscore berekend door middel van een lerend risicoclassificatie-model (RCM). Deze score drukt de kans uit dat een aanvraag volgens het model onjuist is. De aanvragen met de hoogste risicoscore worden aangeboden aan behandelaars van de Toezicht team voor handmatige controle op juistheid. Het proces waarmee deze risicoscore berekend en opgevolgd wordt heet risicoclassificatie.

Het Risicoclassificatie model wordt toegepast voor Huurtoeslag en Kinderopvangtoeslag, afhankelijk van de indicatoren die worden gebruikt in het model.

² Handhavingregie
³ Belastingdienst/Toeslagen

i

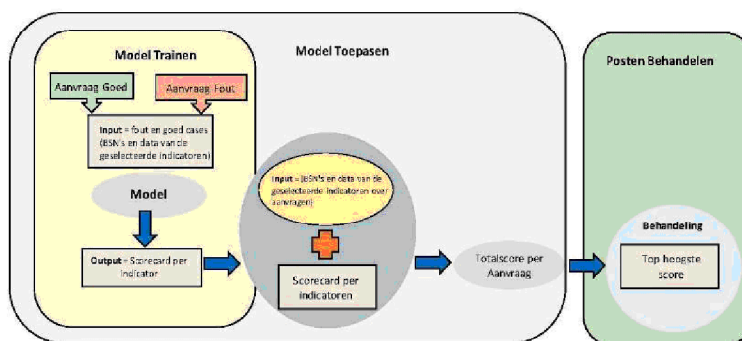
Met opmerkingen (1): Het gaat dus om nieuwe aanvragen en wijzigingen? Verderop lijkt het alleen om nieuwe aanvragen te gaan.

Met opmerkingen (2): Hoezo is dat afhankelijk van de indicatoren? Of is bedoeld dat afhankelijk van het gebruik de indicatoren worden bepaald? (net andersom dus)

Het Risicoclassificatiemodel wordt getraind volgens de supervised learning⁴ methode. Dat betekent dat het model leert van bekende voorbeelden van rechtmatig/onrechtmatig gebruik en de geleerde kennis toepast op onbekende cases.

Het Risicoclassificatie proces bestaat uit 3 onderdelen:

- Het trainen van het model;
- Het toepassen van het model;
- Het behandelen van de geselecteerde posten.



Onderdelen A en B worden verder in deze PIA geanalyseerd. Onderdeel C maakt al deel uit van PIA M8777 Integraal Subjectgericht Toezicht en Toezicht en zal hier dus niet behandeld worden.

Hoe het risicoclassificatie proces begint

HHR geeft opdracht aan DAT om een bepaald aantal toeslagaanvragen te selecteren met de hoogste kans op fouten. Het aantal posten wordt bepaald aan de hand van de beschikbare behandelcapaciteit en er wordt bepaald waar aanvragen aan moeten voldoen om in aanmerking te komen voor selectie door het model (selectie van de indicatoren). De indicatoren zijn variabelen zoals aantal kinderen, leeftijd, huurprijs, etc.

Hoe het model wordt getraind

Het risicoclassificatiemodel is een zelflerend model dat wordt getraind door het toevoegen van goede (rechtmatige) en foute (onrechtmatige) aanvragen. De selectie van goede en foute cases wordt gebaseerd op eerdere behandelingen of ervaringen. De trainingscases bevatten alleen BSN nummers, maar om het model te kunnen trainen zijn meer gegevens nodig. Die extra gegevens zijn de indicatoren die uit verschillende bronnen komen (halfproducten en beschikkingen). De indicatoren worden geselecteerd tijdens brainstormsessies tussen HHR en de behandelaars van het Toezicht team. Die selectie is gebaseerd op huidige en eerdere ervaringen uit het de toezichtpraktijk. Het model wordt gevoerd door het bestand met de datapreparatie in SAS te verwerken; op die manier wordt het model getraind. Tijdens dit proces wordt er bepaald welke indicatoren voorspellen of een aanvraag goed- of fout-afgekeurd wordt en hoe zwaar deze indicatoren wegen. Het resultaat is een scorecard met te beoordelen indicatoren en hun gewicht. Met de beschikbare scorecard worden later alle aanvragen getoetst door het model. Randvoorwaarden voor een betrouwbare scorecard zijn een frequente toevoer van nieuwe testcases zodat het model blijft leren van data die representatief is zijn.

⁴ https://en.wikipedia.org/wiki/Supervised_learning

- Met opmerkingen [3]:** Wat is hiermee bedoeld? Er zullen toch ook toeslaggegevens in zitten?? Of is bedoeld dat bv. geen NAW is opgenomen?
- Met opmerkingen [4]:** Wat zijn halfproducten? Waarom geven die voldoende betrouwbare indicatoren als ze 'half' zijn?
- Met opmerkingen [5]:** Hoe wordt gezorgd voor een objectieve keuze?
- Met opmerkingen [6]:** Wat is dat?
- Met opmerkingen [7]:** Dit gaat dus geautomatiseerd, dat doet SAS?
- Met opmerkingen [8]:** De scorecard is dan toch in feite het model? Want de scorecard bevat de indicatoren en het gewicht. Of bepaalt de scorecard de indicatoren die het model gaat toepassen? De relatie tussen scorecard en model zou mi. wat scherper beschreven kunnen worden.

Hoe het model wordt toegepast

Op de scorecard krijgt elke indicator een bepaalde score toegewezen. Om het model te draaien zijn gegevens van alle aanvragen nodig. De input voor het model zijn de gegevens die gerelateerd zijn aan de indicatoren van de scorecard. Het model wordt daarna in SAS gedraaid.

Het Risicoclassificatiemodel toetst alle aanvragen in een maand aan de hand van de scorecard.

Elke indicator op de scorecard heeft een waarde⁵ en deze waarde resulteert voor iedere aanvraag in een uiteindelijke totaalscore die wordt gebruikt als risicoscore.

De aanvragen met de hoogste risicoscore (ongeveer 30 per maand, afhankelijk van beschikbare capaciteit) worden aangeboden aan behandelaars van het Toezicht team voor handmatige controle op juistheid. De rest van de aanvragen krijgen een formele beschikking conform de aanvraag.

Geïdentificeerd risico's:

Tijdens de analyse van het risicoclassificatieproces zijn er in totaal 10 risico's geïdentificeerd waarvoor ook maatregelen zijn voorgesteld (zie hoofdstuk 16 "Risico's" en 17 "Maatregelen"). Hieronder een overzicht van de eerste 4 risico's met de hoogste kans en impact:

Risico's	Maatregelen
Het risico kan ontstaan dat het model niet meer representatief is omdat er niet voldoende wordt getraind met actuele cases.	Het model regelmatig trainen en testen en oude trainingcases verwijderen. Dmv steekproeven de nauwkeurigheid van het model meten.
Het risico kan ontstaan dat een aantal van de indicatoren onrechtmatig en/of onnodig worden gebruikt.	Gegevens periodiek reviewen en toetsen op dataminimalisatie en grondslag. Check op aanwezigheid van proxies. Waar nodig gegevens niet meer als bron gebruiken en het model hertrainen. Een audit proces vastleggen en goed documenteren.
Het risico kan ontstaan dat er geen goed evenwicht is tussen rechtmatige en onrechtmatige cases (selectiebias).	Om dit risico te voorkomen zijn steekproeven zeer belangrijk. Door middel van een gestratificeerde steekproef wordt de populatie in representatieve groepen verdeeld en kan beter gezorgd worden voor een evenwicht tussen rechtmatige en onrechtmatige cases. Het is ook belangrijk dat er genoeg capaciteit wordt ingezet voor het checken van ten minste 20% van de rechtmatige cases.
Het risico kan ontstaan dat het transparantiebeginsel van de AVG niet wordt nageleefd omdat de informatie in de privacyverklaring beperkt is.	Er komt een addendum voor Toeslagen op de privacyverklaring van de B/T.

Met opmerkingen (9): Is hier niet bedoeld: input voor de analyse?

Met opmerkingen (10): Hier lijkt het alleen om (nieuwe) aanvragen te gaan, niet om wijzigingen. Misschien meer in overeenstemming brengen met de inleiding.

Met opmerkingen (11): Worden er op de 'honderdduizenden' aanvragen/wijzigingen (zoals het in de inleiding staat), maar 360 per maand gecontroleerd?

Met opmerkingen (12): Dit is toch geen risico, want je gaat een model toch niet in productie nemen als je proces voor ontwikkelen en toepassen niet op orde is?? Dat is wmb. net zoals als het risico dat ik onder een auto kom als ik niet links en recht kijk voor ik oversteek. Dat kijken doe ik juist om geen risico te lopen.

Met opmerkingen (13): Moet dat dan niet ook van de onrechtmatige cases (als je evenwicht wil)?

Met opmerkingen (14): Zie opmerking bij eerste risico, dit is van dezelfde aard.

⁵ Bijvoorbeeld: bij de indicator 'aantal kinderen' bepaalt het model de risicoscore aan de hand van het aantal kinderen; het hebben van meer dan 2 kinderen kan dan bijvoorbeeld een hogere score opleveren dan het hebben van meer dan 3 kinderen. Dit is afhankelijk van wat het model heeft geleerd op basis van de trainingcases.

2. Persoonsgegevens

Som alle categorieën van persoonsgegevens op die worden verwerkt. Geef per categorie van persoonsgegevens tevens aan op wie die betrekking hebben. Deel deze persoonsgegevens in onder de typen: gewoon, bijzonder, strafrechtelijk en wettelijk identificerend.

De gegevens die worden verwerkt binnen-in het risicoclassificatiemodel zijn afkomstig uit verschillende bronnen. Met die bronnen worden halfproducten gemaakt die worden gekoppeld aan beschikkingen. Die datapreparatie wordt gebruikt om het model te trainen en toe te passen. Het aantal persoonsgegevens dat wordt gebruikt binnen-voor de datapreparatie is afhankelijk van de indicatoren die geselecteerd worden.

Gewone persoonsgegevens

ID	Veldnaam	Bron1 ⁶	Bron 2 ⁷	Betrekking	Relevant voor onderdeel
1	NAW gegevens: - Postcode en Huisnummer - Woonland - Postadres, Adres woning	BRP ⁸ , CIV ⁹		Av ¹⁰ , Tp ¹¹	A, B, C
2	Persoonlijke gegevens: - Geboortedatum - Leeftijd - Aantal kinderen - Overlijdensdatum	BRP, FLG ¹² (CIV)		Av, (Tp, Mb ¹³ en kind voor geboortedatum)	A, B, C
3	Inkomensgegevens: - Bijstand (WWB uitkering)* - Inkomen - Dagtaak - Draagkracht	FLG (CIV)	Av, Tp, Mb	Av, Tp	A, B, C
4	Huurotoeslag gegevens: - Huurprijs woning - Soort woning - Dienst en Recreatieruimten kosten - Energiekosten - Rekenhuur - Schoonmaakkosten - Standplaatshuur - Woonobject identificatie - Maximale huurgrens	VHA ¹⁴	Av	Av	A, B, C
5	Recht op Kinderbijslag op basis van de AKW ¹⁵	SVB ¹⁶		Av, Tp	A, B, C
6	Gegevens KOO ¹⁷ : Naam	SZW (LRK) ¹⁸ , CIV, DUO	Av	Av, KOO	A, B, C

⁶ Externe bronnen behalve TVS, trainingcases en intensief toezicht die interne bronnen zijn.

⁷ Gegevens worden door de aanvrager zelf verstrekt.

⁸ Basisregistratie Persoonsgegevens

⁹ Centraal Informatie Voorzieningen

¹⁰ Aanvrager

¹¹ Toeslag Partner

¹² Fiscale Loon Gegevens

¹³ Medebewoner

¹⁴ Verhuurders Administratie

¹⁵ Algemene Kinderbijslagwet

¹⁶ Sociale Verzekeringsbank

¹⁷ Kinderopvang Organisatie

¹⁸ Ministerie van Sociale Zaken en Werkgelegenheid

i

Met opmerkingen (15): zie eerdere opmerking

Met opmerkingen (16): Niet duidelijk is waar dit op terugslaat, datapreparatie komt net voor in de vorige zinnen.

Met opmerkingen (17): Rechtstreeks of via BvR?

Met opmerkingen (18): Wat is dit? Een applicatie, en waarmee gevuld? Op basis van BvR?

Met opmerkingen ((19): Waarom komen deze persoonsgegevens uit FLG en niet uit de BRP/BvR?

Met opmerkingen (20): Participatiewet toch?

	- Adres en plaats opvanglocatie en gastouderbureau - Aantal uren opvang - Ingangsdatum kinderopvangtoeslag - periode waarop het kind kinderopvang heeft - uurtarief - soort opvang (dag-/of buitenschools, kindercentrum of gastouder) - LRK-id van opvanglocatie				
7	Aanvraag gegevens: - Type regeling - Ingangsdatum aanvraag - Wijzigingen van de aanvraag op datum en tegemoetkomingsjaar - Product- en procesresultaten van de behandeling van de aanvraag zoals de hoogte van de toeslag (beschikking).	CIV	Av	Av	A, B, C
8	Huishouden gegevens	Belastingdienst (TVS)		Av	A, B, C
9	Persoonsrelaties: Ouder-kind relatie	CIV	Av	Av, Kind	A, B, C
10	Nederlandschap*	BRP		Av	A, B
11	Status voorderingen en beschikkingen: - Toeslagjaar - Beschikking-nummer - Datum-dagtekening - Jaarbedrag	CIV			
12	Wanbetaler zorg status*	CVZ ¹⁹ (CIV)		Av	A, B
13	WOZ onroerend goed bezitting	CIV		Av	A, B, C

Met opmerkingen (21): Waarom niet BRP? En dubbeling met gegevens onder 2?

Met opmerkingen (22): Op grond van Kadaster?

* Deze gegevens (WWB-uitkering, Nederlandschap, Wanbetaler zorg status) worden niet meer als indicator gebruikt maar ze maken nog steeds deel uit van de datapreparatie. Ze zullen in Q1 2021 structureel uit de datapreparatie verwijderd worden (zie Risico's en maatregelen).

Wettelijke identificerend persoonsnummersgegevens

Id	Veldnaam	Bron 1	Bron 2	Betrekking op	Relevant voor onderdeel
14	BSN ²⁰	BVR, DUO ²¹ , COA ²² , CIV, TVS, trainingcases, intensiefToezichtlijst ²³	Av, Tp, Mb, Kind	A, B, C	A, B, C

¹⁹ College voor Zorgverzekeringen

²⁰ Burgerservicenummer

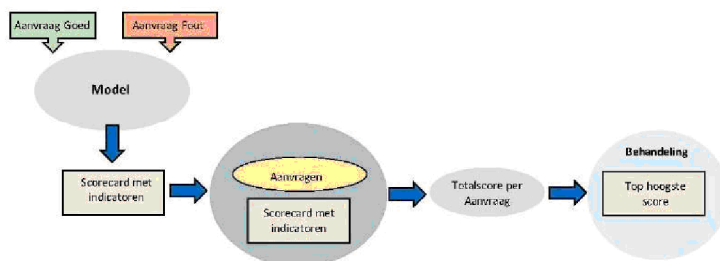
²¹ Dienst Uitvoering Onderwijs

²² Centrale Ontvangers Administratie

²³ Lijst met BSN's die bevat de bewezen fraudegevallen behandeld door het IntensiefToezichtteam

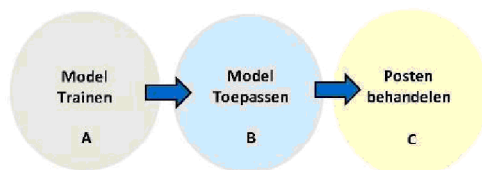
3. Gegevensverwerkingen

Geef alle voorgenomen gegevensverwerkingen weer.



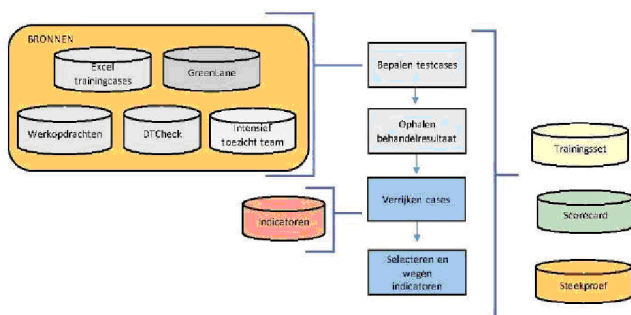
Het risicoclassificatiemodel proces bevat 3 onderdelen:

- A. Het trainen van het model
- B. Het toepassen van het model
- C. Het behandelen van de geselecteerde posten



De PIA is ingedeeld volgens deze driedeling, waarbij dus alleen onderdeel A en B beschreven worden in deze PIA.

A. Trainen van het model



Het RCM is 'lerend': aan de hand van aanvragen uit het verleden waarvan bekend is of deze fout (onrechtmatig) of goed (rechtmatig) waren. Dit is bekend als zijnde de 'trainingscases'. Hierbij wordt uitgerekend welke combinaties van indicatoren risicovol zijn.

De trainingscases vormen het ruggengraat van de risicoclassificatiemodel: zonder een gebalanceerde populatie van kwalitatief goede trainingscases is het onmogelijk een goed model te trainen, en hierop in beleid en uitvoering op te steunen.

Doel van de trainingscases proces is om de toevoer van goede trainingscases te continueren.

Hierdoor wordt de trainingsset geleidelijk ververs, en groeit het model automatisch mee met de patronen van niet compliant cases die recent geconstateerd zijn.

Het model wordt gevoed met trainingscases waaraan historische contra-informatie toegevoegd wordt. Deze reconstructie van trainingscases wordt 'tijdreizen' genoemd.

Alleen wanneer nieuwe trainingscases geïdentificeerd worden zal het tijdreizen aangepast en opnieuw gerund moeten worden. Wanneer geen nieuwe trainingscases beschikbaar zijn zal de trainingsset, en daarmee ook het model, gelijk blijven.

Er worden vijf bronnen voor de trainingscases gebruikt:

1. Trainingscases in Excel
2. GreenLane
3. Werkopdrachten (TVS)
4. Intensief toezicht teams
5. DTCheck

Er wordt per bron bepaald of het behandelresultaat goed of fout is geweest. De bronnen bevatten BSN nummers van aanvragen uit het verleden waarvan bekend is dat deze fout (onrechtmatig) of goed (rechtmatig) zijn geweest. Voor elke bron zijn businessrules opgesteld die bepalen wanneer een case 'fout' of 'goed' genoemd wordt. (zie hoofdstuk 8. Techniek en methode van gegevensverwerking - 2. Bepalen behandelresultaat (goed / fout))

Aan de trainingscases voegt DAT gegevens over de aanvraag en aanvrager toe die pas in de vervolgstap worden geanalyseerd op voorspellende kracht. Dit zijn de indicatoren die gebruikt worden om het scoremodel te bepalen.

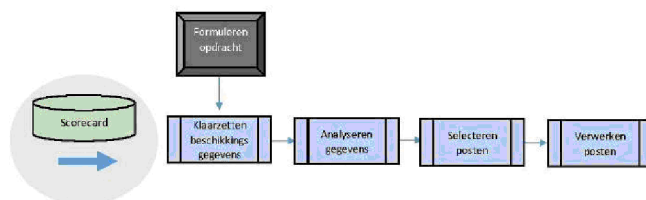
De indicatoren zijn variabelen (bijvoorbeeld: aantal kinderen of aantal verhuizingen) die een indicatie kunnen zijn of onrechtmatigheid. De indicatoren zijn geselecteerd door HHR en PR op basis van eerdere onrechtmatige cases. De informatie die gerelateerd is aan de indicatoren komt uit verschillende bronnen (halfproducten en beschikkingen, zie hoofdstuk 8. Techniek en methode van gegevensverwerking - 3. Klaarzetten beschikking gegevens) en wordt aan het BSN gekoppeld. Dit bestand met BSN's en gekoppelde informatie van de geselecteerde indicatoren wordt gebruikt om het model te kunnen trainen. De verschillende versies van dit bestand worden in SAS opgeslagen. Alleen geautoriseerde DAT medewerkers hebben toegang tot het bestand.

Maar er zijn een aantal oude SAS bestanden die ook op de Q-schijf worden bewaard. Dit stamt nog uit de tijd dat we niet de beschikking hadden over de SAS Grid en alles nog op een aparte pc op moesten slaan. Die bestanden moeten nog bewaard worden om het model te kunnen trainen, maar het is niet meer nodig om die bestanden op de Q-schijf te bewaren - het kan ook alleen op de SAS Grid.

Dit zal in Q1 2021 worden aangepast (zie Risico 9 in Hoofdstuk 16 en 17 Risico's en Maatregelen)

Door middel van een logistisch regressiemodel wordt er tijdens het trainen van het model een scorecard gemaakt. Deze scorecard bevat de indicatoren met hogere scores die daarna worden gebruikt om de risicoscore van de aanvragen te bepalen.

B. Toepassen van het model



De opdrachtgever HHR²⁴ wil data inzetten om risicogebaseerd keuzes te maken over de aan te bieden posten voor toezichtbehandeling in de fase van voorlopige toekenning: de posten waar het risico het grootst is dat de aanvraag/wijziging niet juist en/of rechtmatig is. De opdrachtnemer D.A.T. maakt hiervoor gebruik van het risicoclassificatiemodel waarin door middel van meerdere indicatoren het risico wordt bepaald of een aanvraag (deels) onrechtmatig is.

²⁴ Handhavingsregie

Vergelijkbaar met het trainen van het model selecteert HHR een aantal indicatoren die worden gekoppeld aan de BSN's van alle lopende en nieuwe aanvragen. De verschillende versies van dit bestand worden in SAS opgeslagen. Zoals net bij A benoemd zijn er een aantal oude SAS bestanden op de Q-schijf opgeslagen en dit zal in Q1 2021 aangepast worden. Alleen geautoriseerde DAT medewerkers hebben toegang tot het bestand.

De toepassing van het risicoclassificatiemodel resulteert in een rangschikking van de aanvragen op basis van de risicoscore. Samen met de selectie op basis van harde criteria (zie hoofdstuk 8. Techniek en methode van gegevensverwerking - 5.1 Filteren beschikkingen) wordt de top van deze rangschikking voor handmatige behandeling aangeboden aan toezichtmedewerkers (team Toezicht). In deze behandeling wordt de aanvraag gecontroleerd op het rechtmatig en juist toekennen van de toeslag voordat er een voorlopige toekenning wordt afgegeven.

De teams van Toeslagen die betrokken zijn bij het Risicoclassificatiemodel zijn: HHR in de rol van opdrachtgever, D.A.T. in de rol van opdrachtnemer en Toezicht in de rol van behandelaar van de te controleren posten (uitvoering).

C. Het behandelen van de geselecteerde posten.

De Toezicht team zal als behandelaar de geselecteerde posten controleren. Dit proces valt binnen de scope van PIA M8777 Integraal Subjectgericht Toezicht en Toezicht. Hiermee valt het buiten scope van de PIA en wordt het niet hier beschreven.

4. Verwerkingsdoeleinden

Beschrijf de doeleinden van de voorgenomen gegevensverwerkingen.

Id	Veldnaam	Verwerkingsdoel
1	Postcode en Huisnummer, Woonland, Postadres, Adres woning	Verwerking van persoonsgegevens in het kader van de uitvoering van hun taak en om uitwisseling te kunnen maken met (authentieke) bronnen. Het adres is nodig om de geografische locatie te bepalen en de afstand tot het opvangadres te berekenen (indicator afst_adr_opvang).
2	Geboortedatum, Aantal kinderen, Overlijdensdatum	Geboortedatum wordt gebruikt om met voldoende zekerheid te kunnen stellen of het echt om de gegevens van persoon X gaat (vergewisplicht). Voor kinderopvangtoeslag zijn de leeftijden 4 en 12 jaar relevant als indicatie om te bepalen of er recht is op kinderopvangtoeslag op basis van dagopvang of op basis van naschoolse opvang ²⁵ . Dit is afhankelijk van of een kind naar school gaat. Voor de huurtoeslag dient de aanvrager meerderjarig te zijn. ²⁶

²⁵ In artikel 1.1, lid 1 Wko en artikel 1, sub b Besluit kinderopvangtoeslag.

²⁶ Artikel 1, lid 1, onder c van de wet op de zorgtoeslag en artikel 8 wet op de huurtoeslag.

		Voor de huurtoeslag is de leeftijd van belang bij het vaststellen van de hoogte van het recht op huurtoeslag. Ouderenhuishouden, medebewoner en leeftijd van het kind i.v.m. vrijstellingsdeel van het eventuele inkomen.
3	Huurprijs woning, soort woning, Huurprijs woning, soort woning, Dienst en Recreatieruimten kosten, Energiekosten, Rekenhuur, Schoonmaakkosten, Standplaatsuur, Woonobject identificatie, Maximale huurgrens	Om te kunnen bepalen of er recht is op huurtoeslag, en de hoogte van de huurtoeslag te kunnen bepalen.
4	Statusgegevens: Tp of Av heeft recht op kindgebonden budget op basis van de AKW voor een kind	Deze gegevens zijn nodig omdat het kindgebonden budget afhankelijk is van het recht op kindgebonden budget. ²⁷
5	Toeslagpartner heeft werkt, draagkracht, inkomen	Om te kunnen bepalen bij welke bron de contra-informatie over gewerkte uren kan worden geverifieerd. Ter bepaling van de draagkracht ²⁸ wordt het toetsingsinkomen ²⁹ , van de toeslagaanvrager, toeslagpartner, kind (>18 jaar) en evt. medebewoner in aanmerking genomen.
6	Gegevens KOO: Naam, Adres en plaats opvanglocatie en gastouderbureau, Aantal uren opvang, Ingangsdatum kinderopvangtoeslag, Periode waarop het kind kinderopvang heeft genoten (begin- en einddatum in deze kalendermaand), uurtarief, Soort opvang (dag-/of buitenschools, kindercentrum of gastouder), LRK-id van opvanglocatie	Deze gegevens zijn noodzakelijk om vast te stellen of er sprake is van recht op kinderopvangtoeslag en dit te communiceren met de burger en de KOO in het aanvraag- en mutatieproces. De hoogte van de kinderopvangtoeslag is afhankelijk van het aantal uren kinderopvang per kind in het berekeningsjaar ³⁰ . De hoogte van de kinderopvangtoeslag is afhankelijk van de voor die kinderopvang te betalen prijs ³¹ .
7	Aanvraag gegevens: Type regeling, Ingangsdatum aanvraag, Wijzigingen van de aanvraag op datum en tegemoetkomingsjaar, Product- en procesresultaten van de behandeling van de aanvraag, zoals de hoogte van de toeslag (beschikking).	Om de periode van recht op toeslag vast te kunnen stellen
8	Huishouden gegevens: aantal personen	Voor het vaststellen van het recht op en de hoogte van een tegemoetkoming zijn het aantal personen op een adres van belang.
9	Persoonsrelaties: Ouder kind relatie	Verwerking van persoonsrelaties ten behoeve van het vaststellen van toeslagpartnerschap en de samenstelling van het huishouden teneinde de toeslagen correct te kunnen vaststellen.
10	Status voorderingen en beschikkingen	Om de financiële situatie van de aanvrager te kunnen bepalen.

²⁷ Artikel 2, lid 1 wet op het kindgebonden budget.

²⁸ In artikel 7 lid 1 Awir.

²⁹ In artikel 8 lid 1 Awir.

³⁰ In artikel 1.7 lid 1 sub b ten eerste Wko.

³¹ In artikel 1.7 lid 1 sub b ten eerste Wko.

11	WOZ onroerend goed bezitting	Om te kunnen bepalen of er recht is op huurtoeslag, en de hoogte van de huurtoeslag te kunnen bepalen.
12	BSN	Verwerking van persoonsgegevens ³² in het kader van de uitvoering van hun taak en om uitwisseling te kunnen maken met (authentieke) bronnen. Om het model te trainen worden trainingscases gebruikt. Die bevatten een lijst van BSN's met het resultaat fout of goed. Daarnaast wordt het model getraind met een lijst met posten die vanwege de opzet niet compliant zijn. Die lijst komt van het Intensief toezicht team en bevat alleen BSN nummers. Het BSN wordt gebruikt als unieke identifier om de cases te kunnen identificeren en te kunnen koppelen aan de informatie uit de indicatoren (meer informatie in hoofdstuk 12. Bijzondere persoonsgegevens).

De volgende gegevens worden sinds een aantal jaar geleden niet meer als indicator gebruikt in het model, maar de gegevens worden nog wel verwerkt als bron (ze zijn onderdeel van de datapreparatie). Deze zullen in Q1 2021 structureel uit het model verwijderd worden. (zie Risico's en Maatregelen).

Id	Veldnaam	Verwerkingsdoel
13	Toeslagpartner heeft bijstand (WWB-uitkering)	WWB-uitkering (hoogte) werd als risico indicator in het model opgenomen. WWB-uitkering was een fraudesignaal in het verleden
14	Nederlandschap ja/nee	Het bepalen of sprake is van wel of geen Nederlandschap is van belang om te kunnen vaststellen of een belanghebbende recht heeft op een toeslag. De manier waarop dit werkt is dat er een script naar de bron in BvR kijkt en bepaalt of de waarde van Nederlandschap ja of nee is. Nationaliteit als gegeven wordt niet op een andere manier verwerkt in het proces. Deze gegevens worden ook niet opgeslagen, behalve de output van het script (Nederlandschap ja/nee of 0/1)
15	Wanbetaler zorg status	Het werd als risico indicator in het model opgenomen: aanvrager heeft geen zorgpremie afgedragen (hoger risico).

5. Betrokken partijen

Benoem welke organisaties betrokken zijn bij welke gegevensverwerkingen. Deel deze organisaties per gegevensverwerking in onder de rollen: verwerkingsverantwoordelijke, verwerker, verstrekker en ontvanger. Benoem tevens welke functionarissen binnen deze organisaties toegang krijgen tot welke persoonsgegevens.

De Belastingdienst heeft convenanten gesloten met een aantal organisaties. Dit zijn meestal overheidsorganisaties maar dat hoeft niet altijd zo te zijn. Hierbij gaat het om afspraken over:

³² In artikel 10 Wet algemene bepalingen Burgerservicenummer (Wvab).

- de gezamenlijke handhaving van overheidsregelingen
- het onderling uitwisselen van informatie

De overheid wil dat haar organisaties meer samenwerken waardoor de overheid als geheel efficiënter gaat werken. In dat verband wordt onderling ook informatie uitgewisseld. Uiteraard worden daarbij de wettelijke bepalingen strikt in acht genomen zoals bijvoorbeeld de privacyregelingen. De Belastingdienst hecht aan transparantie en publiceert daarom de samenwerkingsconvenanten met andere organisaties.

Rol	Naam
Verwerkingsverantwoordelijke	Minister van Financiën, voor deze de Belastingdienst, voor deze de Toeslagen
Externe verstrekkers	Holland2stay ³³ , Gemeente, UWV ³⁴ , FLG, VHA, SVB, SZW, CIV, IBS-COA
Interne verstrekkers	CAP ³⁵ , Intensief toezicht team
Interne verwerker	HHR, DAT, PR ³⁶ , CAP-UPG ³⁷ , Toezicht
Interne ontvanger	Werkvoorbereiding Queryloket Handhavingsregie (postbus)

6. Belangen bij de gegevensverwerking

Beschrijf alle belangen die de verwerkingsverantwoordelijke en anderen hebben bij de voorgenomen gegevensverwerkingen.

Het belang van burgers is dat zij tijdig het juiste toeslagbedrag krijgen. Mocht het feit zich voordoen dat zij geen recht hebben, dan is het in het belang van de aanvrager dat deze ook geen voorschot krijgt, zodat er niet onterecht toeslag uitgekeerd wordt en zodat geen terugvordering plaats hoeft te vinden. Daarnaast is de burger erbij gebaat dat hij zo snel mogelijk het juiste toeslagbedrag krijgt en daarbij niet onterecht in lange toezichtprocessen verzeilt raakt.

Het belang van Toeslagen is het juist, tijdig en rechtmatig uitbetalen van toeslagen. Het RCM is daarbij behulpzaam om zo efficiënt en effectief mogelijk de handmatige behandelcapaciteit in te zetten en burgers daardoor ook zoveel mogelijk meteen het juiste toeslagbedrag uit te keren ipv dat achteraf het toeslagbedrag toch anders blijkt te zijn.

Daarnaast speelt het maatschappelijk belang dat er verwacht mag worden dat de overheid zorgvuldig omgaat met en gebruik maakt van gegevens van burgers. De gegevens van burgers op individueel niveau moeten ook inzichtelijk (gemaakt kunnen worden) zijn.

³³ Verhuurder van tijdelijk woonruimte

³⁴ Uitvoeringsinstituut Werknemersverzekeringen

³⁵ Centrale Administratieve Processen

³⁶ Productieregie

³⁷ CAP-UPG (Unit Productiekwiteit en Gebruiksbeheer)

7. Verwerkingslocaties

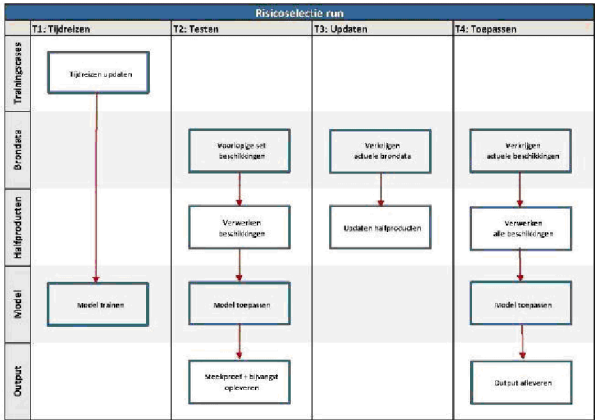
Benoem in welke landen de voorgenomen gegevensverwerkingen plaatsvinden.

Uitsluitend in Nederland.

8. Techniek en methode van gegevensverwerking

Beschrijf op welke wijze en met gebruikmaking van welke (technische) middelen en methoden de persoonsgegevens worden verwerkt. Benoem of sprake is van (semi-)geautomatiseerde besluitvorming, profilering of big data-verwerkingen en, zo ja, beschrijf waaruit een en ander bestaat.

Het risicoclassificatiemodel is lerend en kan op basis van historische aanvragen waarvan de rechtmatigheid/onrechtmatigheid bekend is (trainingscases) de kans voorspellen dat een aanvraag onrechtmatig is. Het model wordt in SAS Enterprise Miner getraind. Tijdens een run van het model wordt in de meeste gevallen eerst een nieuw model berekend. Dit nieuwe model wordt vervolgens toegepast op de meest actuele set beschikkingen. In de praktijk wordt dit niet zo heel vaak gedaan; de huidige scorecard stamt nog uit 2019.



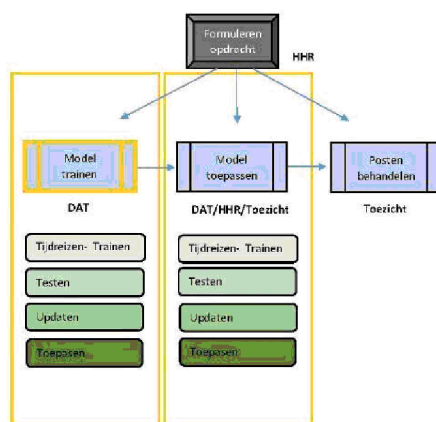
Deze toepassing bestaat vier fasen:

- 1. Trainen:** de set met trainingscases wordt, indien beschikbaar, aangevuld met nieuwe cases. Het model wordt op basis van de nieuwe trainingsets getraind. Tijdreizen is de samenstelling van de trainingset op basis van historische contra-informatie.
- 2. Testen:** het opnieuw getrainde model wordt getest op een voorlopige set van conceptbeschikkingen. Voor de meeste halfproducten wordt hierbij gebruik gemaakt van brondata van de vorige run. Dit leidt tot een steekproef die handmatig beoordeeld wordt door een team medewerkers die fungeren als 'steekproevers'.

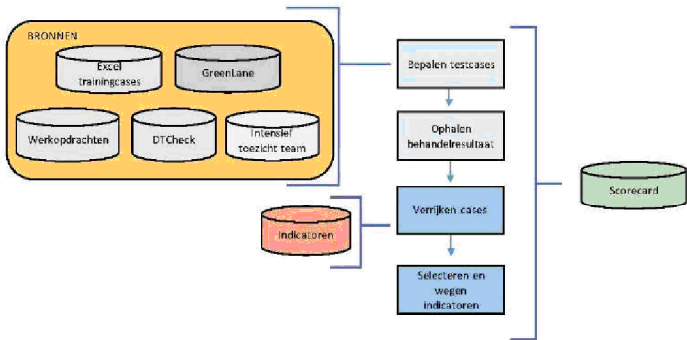
3. **Updaten:** de meeste halfproducten worden ververs met actuele brondata. Als de opvolging van steekproef hier aanleiding toe geeft, wordt het model inhoudelijk aangepast. Het model wordt dan nogmaals toegepast, en uit de resultaten wordt wederom een steekproef getrokken.
4. **Toepassen:** alle actuele conceptbeschikkingen worden verkregen en wordt het model voor de laatste keer toegepast. Dit leidt tot definitieve output van het model, op basis waarvan aanvragen stopgezet, onderzocht en gerapporteerd kunnen worden.

Het RCM proces bevat 3 onderdelen:

- A. Het trainen van het model
- B. Het toepassen van het model
- C. Het behandelen van de geselecteerde posten



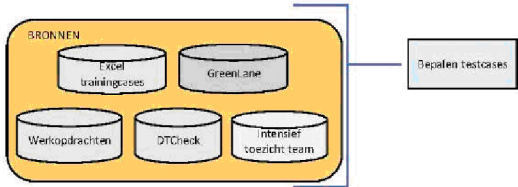
A. Trainen van het model



Proces	Omschrijving	Organisatie
1. Bepalen testcases	Geautoriseerde ²³ medewerkers van DAT en HHR bepalen welk type handmatig behandelde aanvragen worden gebruikt als testcase	DAT
2. Bepalen behandelresultaat (goed / fout)	Geautoriseerde medewerkers van DAT geven het label goed of fout aan de testcases	DAT
3. Verrijken cases	Geautoriseerde medewerkers van DAT koppelen indicatoren aan de testcases	DAT, HHR, Toezicht
4. Selecteren en wegen indicatoren	Geautoriseerde medewerkers van DAT bepalen of hoeveel weging de indicatoren hebben in het voorspellen van goed of fout	DAT

Met opmerkingen **23**: Wat zijn hierbij de criteria? Hoe voorkom je 'bias' in dit proces? Zijn er ook criteria voor wie geautoriseerd wordt (kennis, ervaring ed.)?

1. Bepalen testcases



Voor het bepalen van de testcases zijn vijf bronnen beschikbaar die geautoriseerde medewerkers van DAT gebruiken om een trainingsset te vormen waarmee een scorecard gemaakt wordt:

²³ De autorisaties voor het systeem SAS worden door de Teamleader DAT geregeld.

Met opmerkingen **24**: trainingcases?

- Trainingscases (Excel bestand):** in het verleden (vanaf mei 2013) zijn handmatig trainingscases verzameld, zowel voor KOT als HT. De oorsprong hiervan is velerlei, en niet altijd vastgelegd. De kwaliteit hiervan is daarom in het algemeen (vooral voor de periode mei – december 2013) wisselend. Ondanks de wisselende kwaliteit vormt deze verzameling posten een substantieel deel van de trainingsset. In december 2015 is besloten om cases geïdentificeerd in 2013 te bevroren/bevriezen, en geen nieuwe cases voor toeslagjaar 2013 meer toe te voegen. Daarom worden alle cases met toeslagjaar 2013 of eerder hier al uitgesloten van verdere verwerking. Met ingang van april 2019 zijn ook de jaren 2014 en 2015 bevroren. Sinds 2019 zijn geen cases meer toegevoegd. De cases zijn in de SAS Enterprise tool verwerkt. De output is een Excel-bestand die dat een lijst van BSN's met het resultaat fout of goed bevat. De output is in SAS bewaard. Een backup is ook in een map op de Q-schijf bewaard. Het bestand is toegankelijk voor alle D.A.T. medewerkers.
- GreenLane:** als onderdeel van de GreenLane³⁹ pilots zijn in meerdere iteraties een groot aantal posten behandeld door Toezicht. Deze behandelaars hebben gegeven een groot aantal stukken contra-informatie beoordeeld of deze posten 'de GreenLane op mochten', oftewel: vanwege een laag risico zonder verdere behandeling definitief beschikt mochten worden. De cases die door de rechters als GreenLane waardig beoordeeld zijn, worden in deze flow als complante case meegenomen in de trainingset. Het GreenLane project is gestopt en er worden geen nieuwe trainingscases meer toegevoegd vanuit dit project. De gegevens (in een onversleuteld Excel bestand) waren vanuit productieregie op de Q-schijf geplaatst. Die waren daarna ingelezen in SAS. Het output bestand bevat een lijst van BSN's met het resultaat fout of goed. De output is in SAS bewaard. Een backup is ook in een map op de Q-schijf bewaard. Het bestand is toegankelijk voor alle D.A.T. medewerkers.
- Werkopdrachten:** in deze stap worden de posten toegevoegd waarvan we uit de werkopdrachten weten dat ze in toezicht hebben gezeten. Vervolgens worden daar de beschikkingen aan gekoppeld van vóór en na het toezicht. Afhankelijk van het soort toeslag en verschil tussen het bedrag voor en na de controle wordt bepaald of een post als goed of fout wordt geclassificeerd. Deze verwerking gebeurt in SAS. Gegevens uit het data fundamenteel (TVS en beschikkingen) worden gebruikt. Het output bestand bevat een lijst van BSN's met het resultaat goed of fout. De output is in SAS bewaard. Het bestand is toegankelijk voor alle D.A.T. medewerkers (zie Risico's en Maatregelen).
- Intensief toezicht teams:** de Intensief toezicht teams houden een centrale lijst bij van posten die vanwege opzet niet compliant zijn. Een selectie hiervan kan als fout case toegevoegd worden aan het model. Ook deze stroom is lopend, hoewel aantallen relatief beperkt zijn. De gegevens (in een onversleuteld Excel bestand) worden door de Intensief toezicht team op de Q-schijf geplaatst. Die worden daarna ingelezen in SAS. Het output bestand bevat een lijst van BSN's die in SAS bewaard wordt. Een backup is ook in een map op de Q-schijf bewaard. Het bestand is toegankelijk voor alle D.A.T. medewerkers.
- DTCheck:** als extra check willen we posten hebben die niet gecontroleerd zijn en die als goed getypeerd kunnen worden. Dit is nu alleen gedaan voor toeslagjaar 2017. Daarbij is gekeken naar het verschil tussen VT⁴⁰ en DT⁴¹ en meldingen die aanvragers hebben gedaan. Uit het bestand *dwat* zo verkregen is, is een gestratificeerde steekproef getrokken zodat de steekproef op een aantal achtergrondkenmerken overeenkomt met de populatie. Deze verwerking gebeurt in SAS. Gegevens uit het data-fundamenteel (beschikkingen) worden gebruikt. Het output bestand bevat een lijst van BSN's met het resultaat goed of

Met opmerkingen (25): Is het bestand dan nog voldoende representatief?

Met opmerkingen (26): Dit lijkt me bij uitstek een risico.

Met opmerkingen (27): Hoe kan zo'n beperkt setje gegevens een trainingscase vormen? (geldt ook voor de andere bullits waar dit genoemd is)

Met opmerkingen (28): Hoezo door de rechters?

Met opmerkingen (29): Gaat het hier nu om stappen of bronnen??

Met opmerkingen (30): Is een verschil altijd fout? Of zitten daar nog gradaties in? En zo ja, hoe worden de bandbreedtes bepaald waarbinnen het nog goed is?

Met opmerkingen (31): Relatie met de OGS-problematiek? Is de opzetbepaling betrouwbaar genoeg?

Met opmerkingen (32): Is het dan representatief?

Met opmerkingen (33): alleen beschikkingen of TVS en beschikkingen (zie onder werkopdrachten, daar staat het laatste).

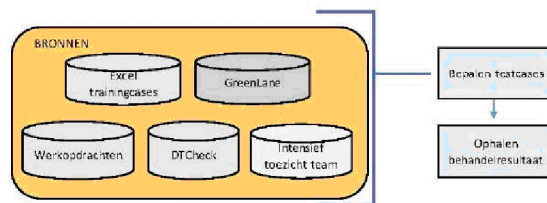
³⁹ GreenLane project uit 2016: Verschuiving van integraal toezicht naar onderbouwd selectief toezicht.

⁴⁰ Voorlopige Toekenning

⁴¹ Definitieve Toekenning

fout. De output is in SAS bewaard. Het bestand is toegankelijk voor alle D.A.T. medewerkers.

2. Bepalen behandelresultaat (goed / fout)



Voor elke bron zijn businessrules opgesteld wanneer een case 'fout' of 'goed' wordt genoemd. Per herkomst is de aanpak hiervan verschillend:

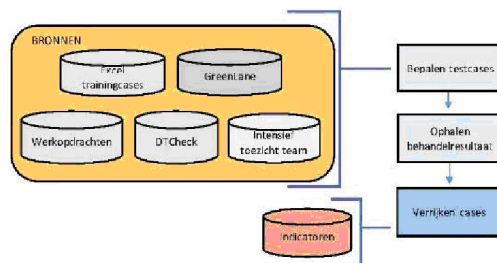
- **Excel:** bij opzet van het model en bij gebrek aan betere trainingcases is ervoor gekozen om een selectie van BSN's die over een langere periode geen mutaties hebben doorgevoerd allemaal als 'Goed' te labelen, ondanks dat hiervoor niet is vastgesteld of de case daadwerkelijk compliant was. Door steeds meer cases te toevoegen wordt de kans kleiner of dat een oneerlijk samengestelde trainingsset wordt gebruikt.
- **GreenLane:** alle posten die in deze iteraties als compliant beoordeeld zijn, kunnen als 'goed' case toegevoegd worden aan de trainingsset.
- **Werkopdrachten:** aan de werkopdrachten worden beschikkingen gekoppeld van vóór en na het toezicht. Afhankelijk van het soort toeslag en verschil tussen het bedrag voor en na de controle wordt bepaald of een post als goed of fout wordt geclassificeerd.
- **Intensief toezicht team:** alle posten die hieronder vallen, worden als 'fout' toegevoegd aan de trainingsset.
- **DTCheck:** alle posten die hieronder vallen, worden als 'goed' toegevoegd. Omdat we over weinig goede cases beschikken hebben we voor 2017 goede cases gehaald uit DT-beschikkingen die nauwelijks afwijken van de VT.

Met opmerkingen [34]: ??

Met opmerkingen [35]: Maar dit zal in het voordeel van de burger zijn toch? Want er is waarschijnlijk meer als 'goed' gelabeld dan daadwerkelijk goed was (want de premisse is 'geen nieuws (mutaties), goed nieuws')

Met opmerkingen [36]: ?

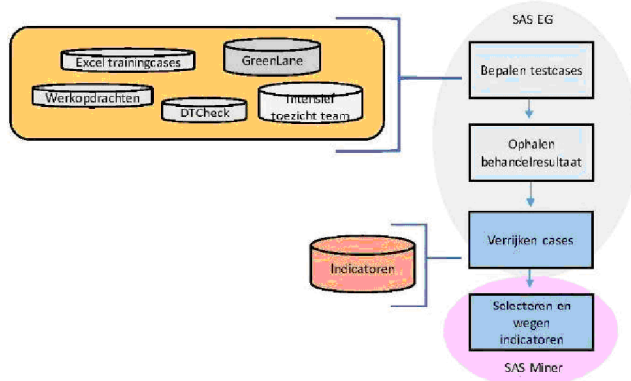
3. Verrijken cases



Aan de cases voegt een geautoriseerde medewerker van DAT gegevens over de aanvraag en aanvrager toe die pas in de vervolgstap worden geanalyseerd op voorspellende kracht. Dit zijn de indicatoren⁴² die gebruikt worden om het scoremodel te bepalen. De indicatoren worden toegevoegd vanuit verschillende bronnen (zie bijlage 3 en 4).

Deze gegevens worden gekozen op relevantie voor de toeslagsoort en op basis van heuristieken van medewerkers van DAT, HHR en Toezicht. De resulterende trainingsset en de indicatorenlijst worden opgeslagen in SAS en op de Q-schijf.

4. Selecteren en wegen indicatoren



De verrijkte gegevens worden met statistische-analyse-software SAS Enterprise Miner geanalyseerd op voorspellende kracht waarna hun weging in de voorspelling wordt bepaald. Door interactive grouping⁴³ wordt elke indicator onderverdeeld in bins, waarna bepaald wordt of de indicator verklarend genoeg is om te voorspellen of een beschikking goed of fout verklaard wordt (volgens de Gini coëfficiënt⁴⁴).

Door middel van een logistische regressiemodel wordt een scorecard⁴⁵ gemaakt, waaruit blijkt hoeveel weging elke bin heeft in het voorspellen van goed of fout. De som van de gescoorde indicatoren geeft de eindscore, waarvan een hogere score betekent dat de aanvraag een relatief hogere kans op fout heeft dan een lagere score. De resulterende scorecard wordt opgeslagen als code in SAS en als een Excel bestand op de Q-schijf.

⁴² De indicatoren zijn tijdens de bouwfase van het model gecreëerd tijdens een brainstormsessies tussen HHR en de behandelaars van het Toezicht team. Die selectie was gebaseerd op huidige en eerdere ervaringen uit ~~het~~ de toezichtpraktijk.

⁴³ Zie voor een beschrijving: <https://documentation.sas.com/?docsetId=emref&docsetTarget=p1qzvw7onopigcn11uc04j16urg7.htm&docsetVersion=15.1&locale=en>

⁴⁴ Gini coëfficiënt: ook wel Gini-index, is een statistische maatstaf van de ongelijkheid in een verdeling.

⁴⁵ Zie voor een beschrijving: <https://documentation.sas.com/?docsetId=emref&docsetTarget=n16jvl3wdvwn9mnlpfqgm3wGoaz5.htm&docsetVersion=14.3&locale=en>

Met opmerkingen (37): Tekst voetnoot in hoofdstekst opnemen. Dit lijkt me vrij essentieel.

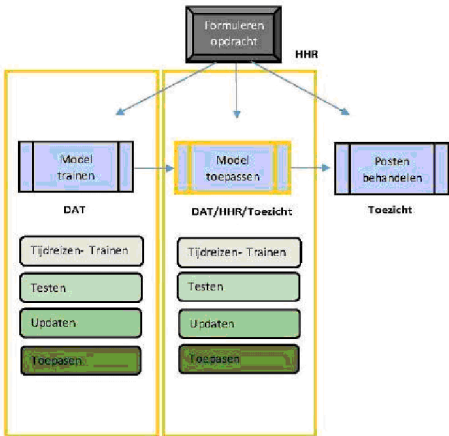
Met opmerkingen (38): Hoe voorkom je hierin bias?

Met opmerkingen (39): Wat zijn dat? Hoe kun je indicator splitsen? Of betekent in bins stoppen juist dat je ze samenvoegt? Dan anders formuleren.

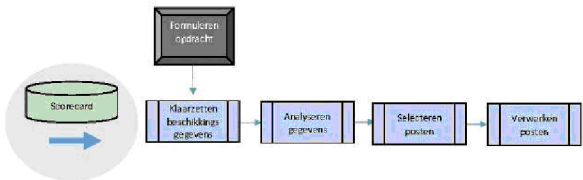
Met opmerkingen (40): Dus een stukje van een indicator? Of juist een setje?

De scorecard wordt daarna gebruikt om de risicoscore van de aanvragen te bepalen (zie onderdeel B. hieronder voor toepassing van het model).

B. Toepassen van het model



Hoofdproces



Proces	Omschrijving	Organisatie
1. Formuleren opdracht	Een intelligence analist van HHR geeft opdracht aan DAT een bepaald aantal toeslagaanvragen te selecteren met de hoogste kans op fout. Een controller van productieregie bepaalt wat de behandelnorm per type post is en een intelligence analist bepaalt hoeveel (afhankelijk van de	HHR, PR

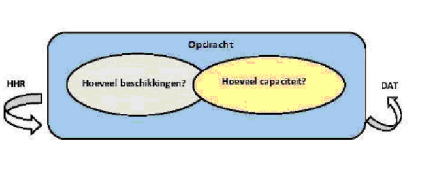
	behandelcapaciteit er beschikbaar is die maand) en waaraan aanvragen moeten voldoen om in aanmerking te komen voor selectie door het model. Een data analist voert de daadwerkelijke selectie uit.	
2. Voorbereiden scorecard	Geautoriseerde medewerkers van DAT bepalen met behandelde posten welke eigenschappen van een aanvraag voorspellend zijn voor het fout zijn van een aanvraagbeschikking, wat resulteert in een scorecard (zie trainen van het model)	DAT
3. Klaarzetten beschikingsgegevens	DAT ontvangt van verschillende bronnen gegevens in een beveiligde omgeving en koppelt deze gegevens aan de beschikkingen	DAT
4. Analyseren gegevens	Geautoriseerde medewerkers van DAT analyseren de beschikkingen aan de hand van gekoppelde gegevens en de scorecard, wat resulteert in gerangschikte beschikkingen	DAT
5. Selecteren posten	Geautoriseerde medewerkers van DAT selecteren te behandelen posten aan de hand van de gerangschikte beschikkingen en behandelcapaciteit. Een Toezicht medewerker voert een steekproef uit op de geselecteerde posten om te controleren of de verwachte risicoscore juist is. Geautoriseerde intelligence analisten van HHR controleren het resultaat en herinitieren het proces als nodig. Als alle stappen volgens verwachting zijn verlopen, biedt de intelligence analist de gewenste aantallen en welke posten niet meegenomen hoeven te worden aan, aan een data analist.	DAT, Toezicht, HHR
6. Verwerken posten	Geautoriseerde medewerkers van DAT leveren de beschikkingen aan het Toeslagen queryloket. Het queryloket maakt werkopdracht aan waaruit bundels volgen die worden aangeleverd aan CAP-UPG om ingelezen te worden in TVS voordat het formeel beschikken gaat draaien. Als ze in TVS staan, worden ze door Toeslagen werkvoorbereiding aangeboden aan medewerkers van IST om de geselecteerde posten te behandelen	DAT, Toezicht

Met opmerkingen (41): Welke bronnen? En wat voor gegevens? Of zijn dit de 'constragegevens' die in de vorige paragraaf beschreven zijn (dus uit de greenlane, de werkopdrachten etc.)?

Met opmerkingen (42): Hoezo koppelen? Het zijn toch al beschikingsgegevens?

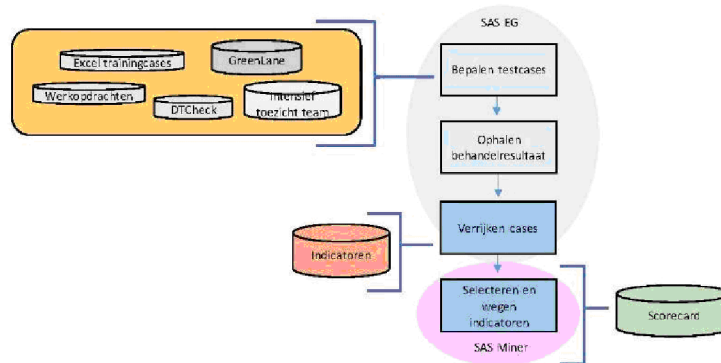
Met opmerkingen (43): Is dit handmatig? Het klonk tot nu toe of die analyse juist door het model wordt gemaakt in SAS.

1. Formuleren opdracht



Proces	Omschrijving	Organisatie
1.1 Behoeftte aan het classificeren van beschikkingen wordt uitgesproken	HHR bepaalt dat beschikkingen voor huurtoeslag, kinderopvangtoeslag en KGB dienen te worden ~~classifice~~ <u>re</u> ~~enge~~ <u>classifice</u> ~~erd~~ .	HHR
1.2 Beschikbare capaciteit wordt bepaald	HHR en PR ⁴⁵ bepalen hoeveel posten hoog risico KOT, hoog risico HT en hot-hor ⁴⁷ aangeboden worden voor handmatige behandeling. Het document waarin dit is vastgelegd wordt gedeeld met DAT per e-mail. Dit document bevat geen persoonsgegevens.	HHR, PR

2. Voorbereiden Scorecard



Het proces om de scorecard voor te bereiden is reeds beschreven in sectie A. Het trainen van het model.- Het doel van het voorbereiden van de scorecard is het bepalen welke indicatoren⁴⁶ voorspellen of een aanvraag goed_ of ~~af~~fout-gekeurd wordt en hoe zwaar deze indicatoren wegen. Het resultaat is een scorecard met te beoordelen indicatoren en hun weging. Randvoorwaarden voor een betrouwbare scorecard zijn een frequente toevoer van nieuwe ~~trainingstestcases~~ en voldoende gegevens over behandelde aanvragen en aanvragers. Hiermee blijft de trainingsset ~~een testcases~~-up-to-date, zodat de scorecard nieuwe types fouten kan blijven vaststellen. Aan de hand van de resulterende scorecard vindt in een vervolgstap de beoordeling plaats van de relatieve kans op het foutief bevinden van nieuwe aanvragen.

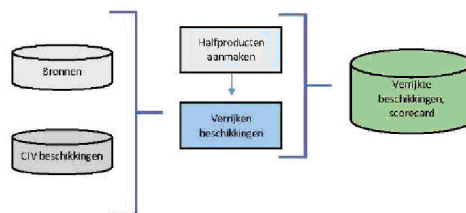
⁴⁶ Productieregie

⁴⁷ Hot Hor: hoge teruggaaf/hoog risico

Het model wordt sinds een jaar geleden niet meer getraind maar het proces wordt herhaald voor het voorbereiden van de scorecard. Steekproeven worden weinig uitgevoerd vanwege gebrek aan capaciteit.

Met opmerkingen (44): ?? Lijkt me een risico.

3. Klaarzetten beschikingsgegevens



Het doel van het klaarzetten van beschikingsgegevens is het verzamelen van gegevens over de aanvraag en aanvrager voor analyse en filtering in vervolgstappen. DAT ontvangt daarvoor ruwe gegevens van dataleveranciers (zie bijlagen 1) en zet deze om tot koppelbare data. De samenvoeging van deze data, al koppelbare gegevens uit het datafundament en beschikkingen, vormt een grote dataset. De gegevens in deze dataset maken het mogelijk een uitspraak te doen over de relatieve kans op het foutief beoordelen van de aanvraag. Ook kan met deze gegevens in een vervolgstap een filtering plaatsvinden op basis van criteria die HHR stelt.

Met opmerkingen (45): Wat wordt nu aan wat gekoppeld? Hoe kun je beschikingsgegevens aan een beschikking koppelen? Ze zitten er dan toch al in? (zie ook eerdere vraag hierover)

Met opmerkingen (46): Filtering waarvoor?

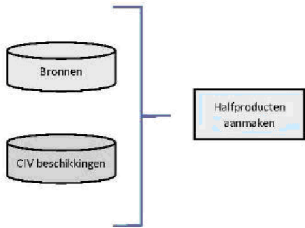
Proces	Omschrijving	Organisatie
3.1 Halfproducten aanmaken	DAT ontvangt van verschillende bronnen (zie bijlage 1) gegevens en bereidt deze voor om te koppelen aan beschikkingen	DAT
3.2 Verrijken beschikkingen	Geautoriseerde medewerkers van DAT koppelen gegevens uit halfproducten en het datafundament aan beschikkingen	DAT

Brondata is afkomstig van diverse afdelingen binnen Toeslagen. De manier waarop deze data aangevraagd en verkregen wordt verschilt per dataset. In alle gevallen is de inhoud van de aanvraag bij de leverende partij bekend (meestal is sprake van gestandaardiseerde queries). De verversing van brondata impliceert in de regel dus enkel een herhaling van bekende zetten. Per run wordt bepaald voor welke bronnen nieuwe data opgevraagd wordt. Voor iedere dataset die verversd moet worden, wordt aan de betreffende contactpersoon per email een nieuwe data-aanvraag gestuurd. Hiervoor wordt een opvraaglijstje bijgehouden.

Gegevens worden vaak per e-mail ontvangen als Excel of csv bestanden (zie bijlage 1 en 2 "ontvangstmethode" bij bronnen en halfproducten). In veel gevallen vindt levering van de data plaats via een gedeelde netwerkschijf. Gegevens uit CIV worden na het afronden van een officiële aanvraag en een WMK toets door CAP klaargezet in een uitwisselmap op de netwerkschijf. Vervolgens worden deze door een data-analist gekopieerd naar de SAS Grid van Toeslagen, waar alleen data-analisten die belast zijn met de risicoclassificatie bij kunnen. Na ontvangst wordt de brondata direct gecontroleerd op plausibiliteit en bruikbaarheid.

Deze controles worden bijgehouden in de risicoclassificatie controlematrix (zie bijlage 5). Bij bevindingen zal overleg plaatsvinden met de leverende partij, deze blijft in principe verantwoordelijk voor de kwaliteit van de data.

3.1 Halfproducten aanmaken



Na ontvangst van nieuwe brondata worden de hiervan afhankelijke halfproducten ververst. Elk halfproduct zet brondata om in een tabel die, vaak op het BSN, aan de (concept)beschikkingen gekoppeld kan worden. Het gaat in de halfproducten voornamelijk om 'contra-informatie': kenmerken van de aanvrager of aanvraag die in de vorm van indicatoren gebruikt worden als input voor het risicoclassificatiemodel.

Alle halfproducten bestaan uit een enkel SAS Enterprise Guide project. In de meeste gevallen zal het betreffende project hiervoor niet of nauwelijks aangepast hoeven te worden, en zal enkel het betreffende project opnieuw gedraaid moeten worden.

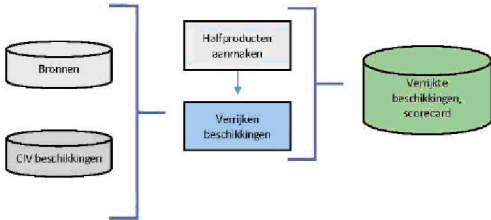
De verwerking wordt gecontroleerd op fouten en implausibele resultaten. De resultaten van deze controles worden bijgehouden in de risicoclassificatie controlematrix.

Nadat het proces gedraaid heeft worden de halfproducten in het model (HT_model en KOT_model) aan de beschikkingen gekoppeld.

Met opmerkingen (47): Wat houdt dit in? Het is ook nog steeds niet helemaal duidelijk wat halfproducten zijn, daarom is ook moeilijk voor te stellen wat deze zin betekent.

Met opmerkingen (48): Door wie? En op grond van welke criteria?

3.2 Verrijken beschikkingen



Koppeling van details uit de aanvraag vanuit de (concept)beschikkingen en contra-informatie vanuit de halfproduct. De gegevens uit de halfproducten worden gekoppeld aan de gegevens van de beschikking dataset die op het datafundament staat. Met deze data worden bepalingen gedaan zoals afstand KOO-aanvrager, hoogste maandbedrag, etc.

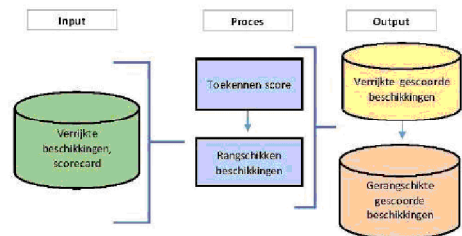
Met opmerkingen (49): Nu opeens concept?

Met opmerkingen (50): Op grond van?

Met opmerkingen (51): Hier staat mi. het omgekeerde van wat in de eerste zin staat. Hoe kan dat?

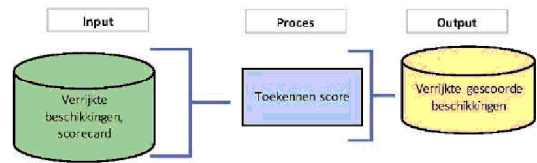
Met opmerkingen (52): Wat zijn dat voor bepalingen dan? Het lijken indicatoren, maar hoe verhouden ze zich dan tot de scorecard?

4. Analyseren Gegevens



Proces	Omschrijving	Organisatie
4.1 Toekennen score	Geautoriseerde medewerkers van DAT kennen per beschikking een score toe op basis van de scorecard	DAT
4.2 Rangschikken beschikkingen	Geautoriseerde medewerkers van DAT rangschikken gescoorde beschikkingen op hoogste kans van foutief	DAT

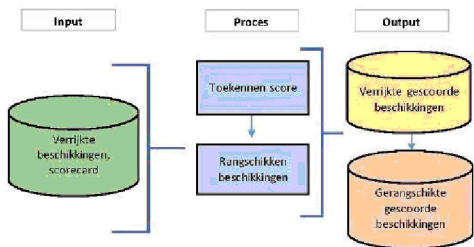
4.1 Toekennen score



Toekenning van risicoscores aan alle (concept)beschikkingen aan de hand van het getrainde risicoclassificatiemodel. De output van deze stap is de ruwe output van het model. Voor alle in het model gebruikte indicatoren is het aantal ScoreCard-punten bepaald. De scorecard wordt gebruikt om per BSN elke verklarende variabele een score te geven. De scores van verklarende indicatoren worden bij elkaar opgeteld. De hoogste score is de relatief hoogste kans op een foute aanvraag en vice versa.

Met opmerkingen [53]: idem als hiervoor

4.2 Rangschikken beschikkingen

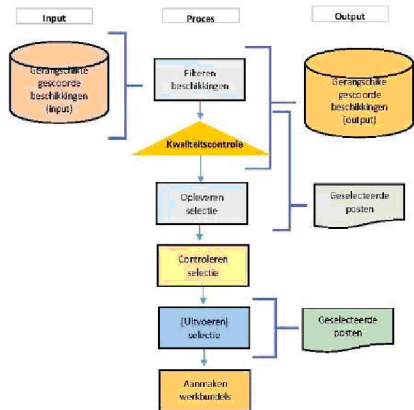


Door rangschikken is het mogelijk een selectie te maken van de groepsaanvragen waar de verwachte hitrate het hoogste is. Het resultaat wordt als een Excel bestand op de netwerkschijf opgeslagen. Om te finaliseren worden de gescoorde aanvragen nabewerkt zodat aan elke aanvraag een opvolgactie toebedeeld kan worden, en de diverse vormen van output samengesteld worden.

Met opmerkingen (54): Wat zijn dat? Of is bedoeld: groepen aanvragen?

Met opmerkingen (55): Wat houdt die nabewerking dan nog in? De te behandelen posten zijn toch al gerangschikt? Wat voor criteria worden bij het nabewerken gebruikt? En wat zijn 'diverse vormen van output'??

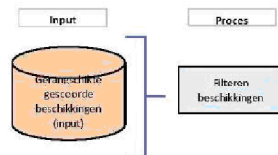
5. Selecteren Posten



Proces	Omschrijving	Organisatie
5.1 Filteren beschikkingen	Geautoriseerde medewerkers van DAT filteren beschikkingen op basis van criteria waaraan een beschikking moet voldoen om in aanmerking van selectie te komen	DAT

5.2 Kwaliteitscontrole	Het proces zelf en de output worden gecontroleerd aan de hand van de controlematrix	DAT
5.3 Opleveren selectie	Geautoriseerde medewerkers van DAT leveren aan een intelligence analist van HHR een lijst met geselecteerde beschikkingen	DAT, HHR, PR
5.4 Controleren selectie	Een Toezicht medewerker voert een steekproef uit op de geselecteerde posten om te controleren of de verwachte risicoscore juist is, geautoriseerde intelligence analisten van HHR controleren het resultaat en koppelt terug als een aanpassing gewenst is.	Toezicht, HHR, PR
5.5 Uitvoeren selectie	Geautoriseerde medewerkers van DAT leveren een definitieve selectie van beschikkingen aan een intelligence analist van HHR en stuurt de geselecteerde posten aan TQL ⁴⁹ om te verwerken	DAT, HHR, TQL
5.6 Aanmaken werkbundels	TQL zorgt ervoor dat werkbundels worden aangemaakt	TQL

5.1 Filteren beschikkingen



Behalve de risicoscore ~~zijn~~ **is** er nog een aantal kenmerken van de aanvraag of aanvrager die bepalen wat er met de conceptbeschikking gebeurt:

- Is de toeslagaanvraag 'op nul gezet'?
- Is voor het BSN van de aanvrager recent een bezwaar afgehandeld?
- Is de aanvrager een 'onbekende aanvrager'?
- Wordt de betreffende BSN voor de betreffende regeling al uitgesloten van beschikken (staat BSN op de 'uitsluitlijst') of was dit recent het geval?
- Is sprake van geen recht op toeslag vanwege 'bijvangst'?
- Is de laatste beschikking een conceptbeschikking?
- Is de aanvraag nieuw of leidt een mutatie tot een hogere toeslag dan voorheen?
- Vormt de aanvraag vanwege een hoog bedrag een hoog risico ('Hot Hor' – alleen voor KOT)?
- Is er een overledene bij de aanvraag betrokken?

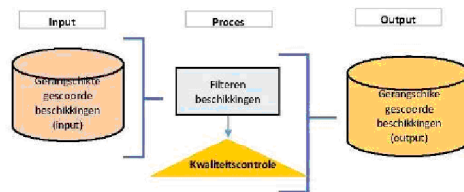
Deze kenmerken worden per toeslagaanvraag in het model vastgesteld.

Met opmerkingen (56): Maar die gaan mee in de selectie (zie vragen hiervoor over conceptbeschikkingen). Worden ze er dan later alsnog sowieso uitgepikt? Zie ook aanhef hierboven, het zijn toch allemaal conceptbeschikkingen (of aanvragen...)?

Met opmerkingen (57): De selectie is toch gericht op het er uit halen van deze? Hot-Hor is toch al onderdeel van risicoscore?

⁴⁹ Toeslagen Query Loket

5.2 Kwaliteitscontrole



Nadat het model gedraaid is wordt de inputdata, het proces zelf en de output gecontroleerd aan de hand van de controlematrix. Pas nadat alle controles afgerond zijn, worden de diverse vormen van output met de rest van de organisatie gedeeld.

Om de kans op fouten tijdens een risicoclassificatierun zo klein mogelijk te maken is een controlematrix samengesteld (zie bijlage 5).

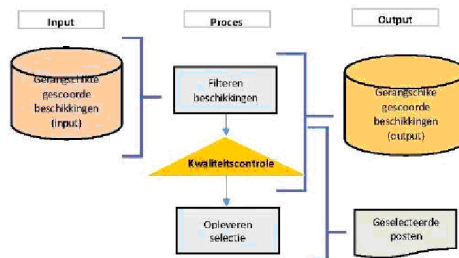
De matrix geeft een lijst met handmatig uit te voeren checks. Deze checks zijn open geformuleerd, er is in de meeste gevallen geen sprake van een norm of gekwantificeerd criterium waaraan voldaan moet zijn. Dit legt de verantwoordelijkheid voor de invulling van de controles bij de controleurs neer. Deze controleurs zullen dus vanuit hun expertise moeten beoordelen of aan de benodigde voorwaarden voldaan moet worden.

Met opmerkingen [58]: Hoe objectief is dit??

De controles voor tijdreizen en de modellen (HT en KOT) worden in principe onder een vierogenprincipe uitgevoerd om fouten te voorkomen die grote gevolgen voor de kwaliteit van het model kunnen hebben. De andere controles worden uitgevoerd door de collega (DAT) die het betreffende proces tijdens de risicoclassificatierun behandelt.

Met opmerkingen [59]: Dit komt een beetje uit de lucht vallen. Is wel ergens in plaatje genoemd, maar niet in de tekst uitgelegd.

5.3 Opleveren selectie



De definitieve run resulteert in verschillende vormen van output:

- Een inhoudelijke uitsluitlijst ('met reden'), die gebruikt wordt om aan de uitsluiting opvolging te geven (bijvoorbeeld in communicatie met de betreffende burger). De reden geeft aan op welke grond de post is geselecteerd.

- Een overzichtstabel, waarin de aantallen posten over een aantal categorieën gegevens worden getoond. Deze tabel vormt de basis voor een rapportage over de resultaten van de risicoclassificatie.
- Voor KOT: een lijst met alle Hot Hor posten die echt behandeld gaan worden.
- Een eindbestand waarin per BSN alle risicoscores zitten.

De bestanden worden in SAS bewaard en in csv formaat op de Q-schijf als back-up opgeslagen.

De resultaten van een run worden per e-mail opgeleverd naar de postbus van QueryLoket, Werkvoorbereiding en Kantoorregisseurs.

In de e-mail worden de volgende bestanden opgeleverd in csv formaat:

- De uitsluitlijsten (één per regeling)
- De overzichtstabel
- De lijst met alle Hot Hor posten

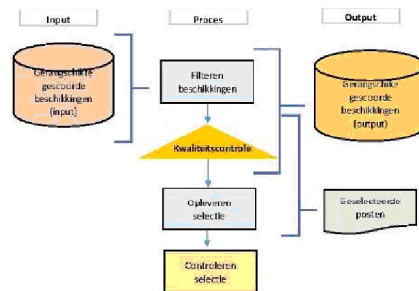
Deze bestanden heeft het Queryloket nodig om de posten uit te sluiten en Productie Regie om werkbundels aan te maken.

De manier hoe waarop de bestanden worden opgeleverd zal in Q1 2021 worden aangepast volgens de nieuwe richtlijnen voor veilig uitleveren van AU⁵⁰ (zie Risico 8 en 9 in Hoofdstuk 16 en 17 Risico's en Maatregelen).

Met opmerkingen (60): En als hiermee de 30 al gehaald wordt? Wat dan?

Met opmerkingen (61): Als dit zo belangrijk is voor het mitigeren van risico's, dan moet meer omschreven worden wat Analytics United inhoudt.

5.4 Controleren selectie



Iedere maand worden Hoog Risico posten uitgezet ter controle. Deze posten worden geselecteerd op basis van beschikbare toezichtcapaciteit en de risicoscore. Om de selectie te valideren wordt vóór het definitief selecteren van de posten, steekproeflijsten opgesteld door HHR en uitgezet bij de Toezicht teams. Vanuit de selectie worden meerdere lijsten/acties gedraaid, deze worden gecontroleerd door HHR & Toezicht.

Met opmerkingen (62): Is dit de hele uitkomst van toepassing van het model? Of weer een aparte categorie?

Proefrun en definitieve run

Voor het draaien van de proefrun bespreken de data en intelligence analisten met elkaar welke aantallen er opgeleverd gaan worden. Bij de proefrun wordt er een aantal Excel lijsten uitgezet ter

⁵⁰ Analytics United

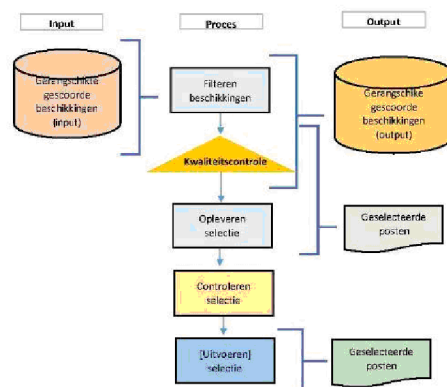
controle bij het Toezicht team. Als de ingevulde Excel lijsten terug gestuurd worden berekent HHR aan de hand van de hoog risico-lijsten het percentage posten dat door de steekproevers als risicovol wordt gezien.

- ✓ Er wordt met kwaliteitsregie afgestemd of er een kwaliteitscheck is gedaan.
- ✓ Het percentage risicovol per regeling wordt uitgerekend. De uitkomsten en redenen worden afgestemd met de data-analisten en er wordt een beslissing over de selectie bij de definitieve run genomen.
- ✓ De uitkomsten van het verslag worden besproken met de data-analisten en gedocumenteerd.

De steekproef heeft verschillende doeleinden. Het is een check om te beoordelen of een toeslagaanvraag terecht door het risicoclassificatiemodel als risicovol is aangemerkt. De behandelaar controleert of de aangeboden BSN's daadwerkelijk risicovol zijn. Dit doet hij o.b.v. enkele factoren (bijv. voor KOT: hoge opvanguren, verhouding inkomen - aantal opvanguren - kosten en voor Huur: verhouding inkomen - huurprijs, ronde huurprijzen, hoge servicekosten, evt. oppervlakte woning en aantal personen op de beschikking). Dit is belangrijk want het model selecteert de posten voor Risicoclassificatie en impliciet dus ook waar wij toezicht op inzetten, de controle door de behandelaar zorgt er voor dat die kan bevestigen of het model de juiste/risicovolle aanvragen eruit haalt. Daarnaast worden de resultaten van de steekproef gebruikt om de verdeling tussen Hoog Risico HUUR, Hoog Risico KOT en HoHoH te bepalen. Het kan namelijk voorkomen dat in de ene maand de Hoog Risico HUUR risicovoller is dan Hoog Risico KOT, met tot gevolg dat er naar verhouding meer HUUR uitgezet wordt.

Momenteel wordt de proefrun op de dag van de definitieve run uitgevoerd. Een verandering ten opzichte van de runs van voor november 2019 is dat er geen maandelijkse steekproef wordt uitgevoerd op de nieuwe aanvragen/mutatie Hoog Risico KOT en HUUR. Deze steekproef wordt enkel uitgevoerd als het model getraind wordt. Redenen die hebben geleid tot aanpassingen in steekproeven zijn o.a. de afnemende toezichtcapaciteit, risicogerichte keuzes, dat steekproeven veel tijd vergen en dat HHR meer kan met de beschikbare data.

5.5 Uitvoeren selectie

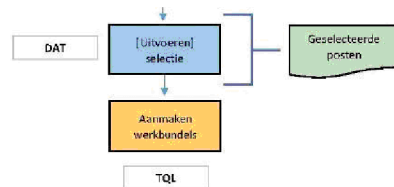


Met opmerkingen [63]: ??

Met opmerkingen [64]: En zo niet, wat dan?

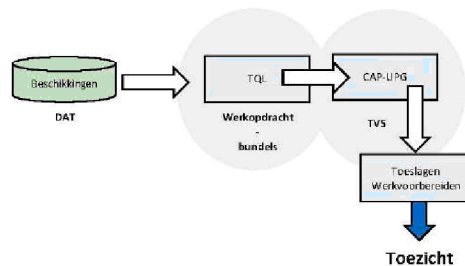
Na de validatie van de selectie levert DAT na een definitieve run een definitieve selectie van beschikkingen aan HHR. De geselecteerde posten worden per e-mail naar de postbus van TQL verstuurd om te verwerken.

5.6 Aanmaken werkbundels



TQL zorgt ervoor dat werkbundels worden aangemaakt zodat behandelaars de posten verder kunnen controleren.

6. Verwerken posten



DAT levert de beschikkingen (geselecteerde posten) aan het Toeslagen queryloket. Het queryloket maakt werkopdracht aan waaruit bundels volgen die worden aangeleverd aan CAP-UPG om ingelezen te worden in TVS voordat het formeel beschikken gaat draaien. Als ze in TVS staan, worden ze door Toeslagen werkvoorbereiding aangeboden aan medewerkers van Toezicht om de geselecteerde posten te behandelen.

Werkwijze:

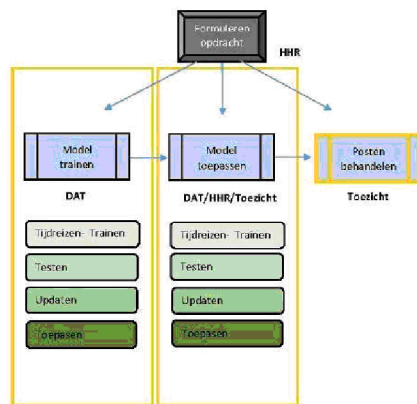
Werkbundels worden ontvangen middels een serviceaanvraag en begeleidende mail. Door de beperkte hoeveelheid bijlagen die ITSM aan kan worden deze ook via de mail aangeleverd. De serviceaanvragen zijn voor iedereen inzichtelijk die aan groep BCA12002 (=CAPHeffing Toeslagen Functioneel Beheer) is gekoppeld. De bijlagen in de postbus zijn voor iedereen toegankelijk die rechten heeft tot de postbus 'CAP Toeslagen Productie_Postbus'.

De werkbundels worden verwerkt conform de instructies in het opdrachtformulier. De bestanden worden vanuit de mail of serviceaanvraag verplaatst naar een directory op de Q-schijf. Vanuit het

Kantoorportaal wordt de opdracht gestart om de bestanden te verwerken. Alle leden van de vakgroepen 'Formeel Beschikken' en 'Productie Coördinatoren' hebben het recht om deze bestanden via het Kantoorportaal te verwerken. Na verwerking wordt de verwerkte werkbundel in een archiefmap op de Q-schijf geplaatst. Middels een query op de toeslagendatabase wordt bij een ongelijk aantal tussen aangeleverde en verwerkte items het verschil verklaard en teruggekoppeld aan Kantoor. Daar wordt vervolgens bepaald of er een nieuwe werkbundel aangeleverd wordt.

Het resultaat van de verwerkingen is het uitsluiten voor het formeel beschikken. Die wordt gedeeld via een reactie op de serviceaanvraag en op de gestuurde mail.

C. Het behandelen van de posten



De Toezicht team controleert als behandelaar de geselecteerd posten.
Dit proces wordt hier niet verder geanalyseerd omdat het binnen de scope valt van PIA M5777 Integraal Subjectgericht Toezicht en Toezicht.

Beveiliging – Autorisatiebeleid

De Belastingdienst (inclusief B/T) hanteert een niveau van beveiliging dat gericht is op de beveiliging van massale verwerking van persoons- en financieel-economische gegevens. De BIO (Baseline Informatiebeveiliging Overheid) vormt hiervoor het gemeenschappelijk normenkader voor de beveiliging van de informatie(systemen). Het niveau van beveiliging is gemiddeld genomen voor alle informatiesystemen tevens het basisniveau omdat voor wat betreft beveiliging de processen en gegevens voor een groot deel gelijkwaardig zijn. Voor alle informatiesystemen die geraakt worden in de verwerkingen is het basis beveiligingsniveau 2 (BBN2) van toepassing.

Autorisatiebeleid

In de autorisatiematrix IMS worden bedrijfsrollen aangemaakt. Aangewezen medewerkers krijgen deze bedrijfsrollen toegekend door de direct leidinggevende ter ondersteuning van de uitvoering van opgedragen werkzaamheden. Controle op uitgegeven bedrijfsrollen gebeurt via het LTB proces (logisch toegangsbeheer). Conform de BIO moeten alle teamleiders minstens 2x per jaar de autorisaties controleren. Centraal is door SSO F&MI in IMS geregeld dat als een medewerker doorstroomt naar een ander dienstonderdeel de autorisaties van het latende dienstonderdeel na 14 dagen automatisch worden ingetrokken.

Analyseomgeving: datafundament, SAS EG, SAS Miner

De analyseomgeving waar de B/T gebruik van maakt, wordt inclusief toegangsmanagement geleverd door het domein Analytics. Autorisaties verlopen via IMS en worden toegekend door de teamleider van het team Data Analytics Toeslagen (DAT) binnen B/T. De medewerker Informatiebeveiliging binnen B/T valideert vervolgens de autorisatie aanvraag. Op het moment van schrijven, 17 november 2020, hebben in totaal 13 medewerkers toegang tot het datagebied voor het risicoclassificatiemodel, het betreft 9 analisten en 4 beheerders/ontwikkelaars van het fundament. Afhankelijk van de opdracht kunnen deze medewerkers gegevens raadplegen, koppelen aan andere gegevens en gebruiken voor verdere analyse, waarbij de instructies uit de opdracht gevolgd worden. De gegevens worden geautomatiseerd door het domein Gegevens in de analyseomgeving geladen, of via een handmatig proces door de beheerders van B/T, zijnde medewerkers van DAT.

Ondanks het autorisatiebeleid is de toegang tot de gegevens in SAS m.b.t. het RCM niet voldoende beperkt. Op dit moment is er geen specifieke groep medewerkers die als enige geautoriseerd zijn om te werken aan het RCM model. Dit is ook de situatie op de netwerkschijf. Een accurate toewijzing van autorisaties en logging ontbreekt deels waardoor mutatie- en exportrechten ten onrechte gebruikt zouden kunnen worden. (zie Risico's en Maatregelen)

Logging en monitoring in het kader van Informatiebeveiliging

De Minister van Financiën, voor deze de [] Belastingdienst, voor deze de [] is verantwoordelijk voor het loggen en monitoren van het gebruik van de ICT-infrastructuur (verwerking M2854). Doel hiervan is het uitvoeren van het geldende beveiligingsbeleid en oplossen van problemen door het analyseren van (oneigenlijk) gebruik van ICT infrastructuur. Het gaat om de logging van het BSN en IP-adres van burgers, met datum/tijd en uitgevoerde actie, en het IP-adres en user ID van medewerkers, met datum/tijd en uitgevoerde actie. Dit heeft betrekking op alle ICT-systemen van de Belastingdienst en de uitvoering hiervan ligt bij het Security Operation Center (SOC) van de Belastingdienst.

Specifiek voor het datafundament is er monitoring op een aantal vooraf gedefinieerde handelingen. Als een van deze handelingen zich voordoet dan wordt de Business Security Officer van Toeslagen hierover geïnformeerd die een beoordeling doet en een mogelijk vervolgonderzoek instelt. De monitoring ziet toe op:

- specifieke zoekopdracht naar 1 burger
- specifieke zoekopdracht naar bankrekeningnummer
- exporteren van gegevens

Semi-geautomatiseerde besluitvorming

Er is sprake van semi-geautomatiseerde besluitvorming. Aan de hand van het RCM wordt bepaald of een aanvraag hoog risico is. Afhankelijk van de behandelcapaciteit worden de posten met de hoogste risico in behandeling genomen door de behandelteams van Toeslagen.

Door gebruik te maken van semi-geautomatiseerde besluitvorming is het mogelijk om voor Toeslagen op een efficiënte wijze de posten met de hoogste risico's te selecteren.

Het gebruik van het risicoclassificatiemodel door Toeslagen is een vorm van profileren omdat deze verwerking voldoet aan alle drie de voorwaarden genoemd in artikel 4, aanhef en onder 4, van de AVG. Ten eerste is sprake van een geautomatiseerde vorm van verwerking. Het risicoclassificatiemodel is namelijk een algoritme dat automatisch aanvragen selecteert waarop personele capaciteit wordt ingezet. Ten tweede heeft de verwerking betrekking op persoonsgegevens. De informatie die gebruikt wordt voor de indicatoren in het risicoclassificatiemodel zijn persoonsgegevens als bedoeld in artikel 4, aanhef en onder 1, van de AVG. Ten derde heeft de verwerking met het risicoclassificatiemodel ten doel om individuele kenmerken te evalueren en categoriseren. Aan de hand van persoonlijke aspecten van de aanvrager wordt geëvalueerd en op basis daarvan ingeschat hoe groot het risico is dat de aanvrager een onjuiste aanvraag heeft ingediend, en wordt het toezicht daarop afgestemd.

9. Juridisch en beleidsmatig kader

Benoem de wet- en regelgeving, met uitzondering van de AVG en de Richtlijn, en het beleid met mogelijke gevolgen voor de gegevensverwerkingen.

Algemene wet inkomensafhankelijke regelingen (AWIR): artikel 11. B/T voert de inkomensafhankelijke regelingen uit die haar zijn opgedragen, te weten: de Zorg-, Huur en kinderopvangtoeslag en het Kindgebonden budget. Belastingdienst/Toeslagen (B/T) is belast met het toekennen, uitbetalen en terugvorderen van tegemoetkomingen. Op grond van deze wetgeving kan dus worden gesteld dat B/T is belast met de publiekrechtelijke taak om uitvoering te geven aan het toekennen, uitbetalen en terugvorderen van ZT, HT, KOT en KGB. Bij de uitvoering van deze taken moet de B/Toeslagen de procedure volgen die is vastgelegd in hoofdstuk 2, paragraaf 2 Awir. Dat betekent o.a. dat B/T in het lopende jaar voorschotten verleent (artikel 16 Awir en definitieve toekenningen (art. 19, Awir, na afloop jaar).

Daarnaast past de gegevensverwerking in het vastgestelde beleid van B/T. Dit blijkt bijvoorbeeld uit de handhavingsvisie en het Klantsegment Toeslagen waarin behandeling op maat alsmede informatiegerecht risicodetectie ten doel zijn gesteld. Verdere informatie over de beleidsmatige uitwerking van de doelen staat beschreven in het Toeslagen Toezicht Werkplan 2018 (zie bijlage 6).

Bij de controle van ingediende aanvragen en bij de toekenning wordt gebruik gemaakt van de gegevens waarover de Belastingdienst reeds beschikt. Zo beschikt de Belastingdienst reeds over de inkomensgegevens en gegevens over de huishoudsituatie. Voor de huurgegevens bestaat in het kader van de uitvoering van de Huursubsidiwet reeds een automatische uitwisseling van huurgegevens met een aantal verhuurders (deels woningcorporaties). Op basis van deze gegevens kan een effectieve controle van de huurgegevens plaatsvinden. De controlemogelijkheden die deze automatische uitwisseling van huurgegevens met zich meebrengt, blijft ook in de toekomst, als de uitvoering van de huursubsidie in handen van de Belastingdienst Toeslagen komt, gewenst (Uit de MvT Awir).

Als een voorschot is verleend, is de aanvrager verplicht bepaalde wijzigingen door te geven (artikel 17 Awir jo. artikel 5 Uitvoeringsregeling Awir). Daarnaast is een aanvrager, partner of medebewoner (gegevens van medebewoners worden alleen voor huurtoeslag verwerkt) verplicht mogelijk relevante gegevens en inlichtingen te verstrekken, als B/T daarom vraagt (art. 18 Awir).

De hoogte van de kinderopvangtoeslag is onder andere afhankelijk van de draagkracht van de aanvrager en de eventuele toeslagpartner. Hiervoor wordt het toetsingsinkomen in aanmerking genomen (artikel 7, eerste lid, Awir, en artikel 1.7, eerste lid onderdeel a, Wko). Daarnaast is de



hoogte afhankelijk van de kosten van kinderopvang, waaronder het aantal uren afgenomen kinderopvang (artikel 1.7, eerste lid onderdeel b, Wko). Om die reden dienen relevante wijzigingen in deze omstandigheden binnen vier weken door aan de B/T te worden doorgegeven (artikel 5 Uitvoeringsregeling Awir). Als deze wijzigingen echter niet (tijdig) worden doorgegeven, dan kan er een te hoog voorschot zijn verleend, dat door de B/T moet worden teruggevorderd (artikel 26 Awir).

Publiekrechtelijke organisaties, zoals DUO en UWV, en aangewezen andere personen en organisaties zijn verplicht om gegevens en inlichtingen aan B/T te verstrekken waarvan de kennisneming van belang kan zijn voor de uitvoering van de Awir (artikel 38, eerste lid, Awir). In artikel 1a, eerste lid, onderdeel b, van het Uitvoeringsbesluit Awir zijn de kinderopvangorganisaties aangewezen voor de levering van de gegevens, bedoeld in artikel 11, eerste lid, onderdeel b, van de Regeling Wko.

Daarnaast verstrekken inspecteur en ontvanger aan de B/T gegevens en inlichtingen uit die nodig zijn voor de uitvoering van de Awir (artikel 39, eerste lid, Awir).

10. Bewaartermijnen

Bepaal en motiveer de bewaartermijnen van de persoonsgegevens aan de hand van de verwerkingsdoeleinden.

Op grond van artikel 12, tweede lid, Regeling Archiefbeheer Belastingdienst 2011 (RAB 2011) is voor B/T een selectielijst vastgesteld en gepubliceerd in Stcrt. 2015/31399 (zie ook artikel 5 van de Archiefwet 1995).

Het archiefbeherend onderdeel draagt zorg voor de vernietiging van de daarvoor in aanmerking komende archiefwaardige informatieobjecten na het verstrijken van in de selectielijst daarvoor vastgestelde termijn. Bij fysieke informatieobjecten gebeurt dit door Doc-Direct (zie artikel 16, eerste lid, RAB 2011). Bij digitale informatieobjecten gebeurt dit vanuit de Toeslagen Archiefvoorziening (TAV). Van de te vernietigen archiefwaardige informatieobjecten wordt een verklaring van vernietiging opgesteld. Deze wordt ondertekend door de B/CFD, unit Documenthuishouding (zie artikel 16, tweede lid, RAB 2011).

De bewaartermijn voor producten die voortkomen uit het verzamelen, analyseren en veredelen van gegevens en deze omzetten in relevante handavingsinformatie (intelligenceproces) bedraagt op grond van de Selectielijst Toeslagen 7 jaar na vervallen.

Het betreft hier producten (in- en output) zoals: Interne/ externe uitvoeringsgegevens, gedrag gerelateerde gegevens, fraude- en risicosignalen, (strategische) verkenningen, (evaluatie)onderzoeken, trendanalyses, monitoring, weegregels, werkbundels (toezichtslijsten), contra-informatie.

De bewaartermijn voor producten die voortkomen uit het uitvoeren van thematisch toezicht

bedraagt op grond van de Selectielijst Toeslagen 12 jaar na beëindiging van de toeslag.

Het betreft hier producten (in- en output) zoals: risicosignaal, fraudesignaal, melding, grondslag, beslissing, brief, boete, toeslagdossier ten behoeve van de overdracht aan de FIOD.

De bewaartermijn voor producten die voortkomen uit het handmatig, door een toeslagmedewerker, behandelen van gevallen die niet volledig geautomatiseerd kunnen worden verwerkt (handmatige klantbehandeling) bedraagt op grond van de Selectielijst Toeslagen 12 jaar na beëindiging van de toeslag.

Het betreft hier producten (in- en output) zoals: werkopdracht, melding, grondslag, beslissing, brief.

De bewaartermijn voor producten die voortkomen uit het nemen van besluiten over het te kiezen handavingsinstrumentarium bedraagt op grond van de Selectielijst Toeslagen 7 jaar na vervallen.



Het betreft hier producten (in- en output) zoals: kaders, advies, (concept) convenanten, handhavingsplannen, intelligenceopdrachten, werkbundels/werkvoorbereiding, effect- en mismatchrapportages, uitsluitlijsten.

Binnen Toeslagen wordt er naast reguliere applicaties ook gebruik gemaakt van gegevens in de vorm van lijsten en extracten. Deze gegevens worden door medewerkers op de persoonlijke omgeving, de e-mail of op samenwerkingsgebieden op de netwerkschijf bewaard. Hierbij blijft zorgvuldigheid van belang om onterecht gebruik, datalekken of onbedoelde vernietiging te voorkomen.

Ingevolge artikel 5 lid 1 sub e AVG mogen persoonsgegevens niet langer worden bewaard dan voor de doeleinden noodzakelijk is; persoonsgegevens mogen voor langere perioden worden opgeslagen voor zover de persoonsgegevens louter met het oog op archivering in het algemeen belang, mits passende technische en organisatorische maatregelen worden getroffen om de rechten en vrijheden van de betrokkene te beschermen. Momenteel is er geen formeel proces voor het vernietigen van gegevens in SAS en op de Q-schijf. Dit is echter belangrijk om te kunnen voldoen aan de bewaartermijn van de selectielijst. In Q1 2021 zullen er stappen genomen worden richting het vernietigen van de gegevens in SAS en de Q-schijf, om het risico van opslagbeperking te voorkomen (zie risico 9 in hoofdstuk 16 en 17 Risico's en Maatregelen). Tijdens deze implementatie zal ook geëvalueerd worden of een back-up van de output en de tussenbestanden van het risicoclassificatiemodel op de Q-schijf moet blijven worden opgeslagen. De reden daarvoor zou zijn dat SAS Grid een applicatie is van een derde partij en wij ons er van willen verzekeren dat wij kunnen voldoen aan inzageverzoeken, audits en uitlegbaarheid. Momenteel zijn back-up bestanden alleen via SAS toegankelijk voor geautoriseerde medewerkers van DAT.

B. Beoordeling rechtmatigheid gegevensverwerkingen

Beoordeel aan de hand van de feiten zoals vastgesteld in onderdeel A of de voorgenomen gegevensverwerkingen rechtmatig zijn. Het gaat hier om de beoordeling van de juridische rechtsgrond, noodzaak en doelbinding van de gegevensverwerkingen. Beoordeel tevens de wijze waarop invulling wordt gegeven aan de rechten van de betrokkenen. Voor dit onderdeel van de PIA is in het bijzonder juridische expertise nodig.

11. Rechtsgrond

Bepaal op welke rechtsgronden de gegevensverwerkingen worden gebaseerd.



In Hoofdstuk 2, paragraaf 5 van de Awir (artikelen 38 t/m 39a) is de informatieverstrekking aan en de informatie-uitwisseling door B/T geregeld.

Artikel 38, over de informatieverstrekking aan B/T, is er ook verder uitgewerkt in het uitvoeringsbesluit Awir (de artikelen 1a en 1b). Het laatste artikel bepaalt dat de gegevens worden verstrekt aan B/T onder vermelding van het BSN.

In de artikelsgewijze toelichting op artikel 38 Awir staat o.a. het volgende: Dit artikel regelt de kosteloze verstrekking aan de Belastingdienst Toeslagen van gegevens en inlichtingen waarvan de kennisneming van belang kan zijn voor de uitvoering van deze wet. Omdat, afhankelijk van de regeling waarop dit wetsvoorstel van toepassing is, de aanspraak op of de hoogte van een tegemoetkoming mede afhankelijk kan zijn van informatie die niet binnen de rijksbelastingdienst of de Belastingdienst Toeslagen voorhanden is, is in het eerste lid voor decentrale overheden, bepaalde semi-overheidslichamen en voor bij algemene maatregel van bestuur aangewezen natuurlijke personen, maat- en vennootschappen, verenigingen en andere rechtspersonen, instellingen en diensten een verplichting tot informatieverstrekking opgenomen. Het gaat dan vanzelfsprekend om gegevens en inlichtingen waarover die betrokken persoon of instantie redelijkerwijs geacht wordt te beschikken.

Voor de algemene persoonsgegevens geldt dat de rechtsgrond gebaseerd is op artikel 6, eerste lid, onderdeel e, AVG: de verwerking is noodzakelijk voor de vervulling van een taak van algemeen belang of van een taak in het kader van de uitoefening van het openbaar gezag dat aan de verwerkingsverantwoordelijke is opgedragen.

In artikel 11, tweede lid, van de Awir is immers bepaald dat de B/T het organisatiedeel van de rijksbelastingdienst is dat is belast met het toekennen, uitbetalen en terugvorderen van de tegemoetkomingen. In artikel 1.3, eerste lid, van de Wko is dit nog eens expliciet herhaald ten aanzien van de kinderopvangtoeslag.

Publieke taak Artikel 1.3, eerste lid, van de Wko bepaalt dat Toeslagen is belast met de uitvoering van het toekennen, uitbetalen en terugvorderen van de kinderopvangtoeslag. Op grond van het tweede lid van die bepaling is op de Wko de Awir van toepassing. Ingevolge artikel 43 van de Awir, gelezen in samenhang met artikel 1 van het Besluit aanwijzing toezichthouders Awir, is Toeslagen eveneens belast met het toezicht op de naleving van het bepaalde bij of krachtens de Awir. Gelet op de in paragraaf 3.1.2 beschreven positie van Toeslagen, als onderdeel van de onder de minister ressorterende rijksbelastingdienst, kunnen deze taken voor de toepassing van de AVG worden toegerekend aan de minister.

12. Bijzondere persoonsgegevens

Indien bijzondere of strafrechtelijke persoonsgegevens worden verwerkt, beoordeel of één van de wettelijke uitzonderingen op het verwerkingsverbod van toepassing is. Bij verwerking van een wettelijk identificatienummer beoordeel of dat is toegestaan.



De overige verwerkte gegevens en de context bij de verwerking van Nederlanderschap door het risicoclassificatiemodel geven geen aanleiding om een indirect verband tussen nationaliteit en ras of etnische afkomst aan te nemen. Het is in dit geval geen sprake van bijzondere persoonsgegevens in de zin van artikel 9, eerste lid, van de AVG. Ook heeft het gebruik niet tot doel om onderscheid te maken naar ras of etnische afkomst.

Nederlanderschap wordt niet meer als indicator in het model gebruikt. De gegevens Nederlanderschap ja/nee die nog onderdeel zijn van de datapreparatie zullen in Q1 2021 verwijderd worden uit het model.

Het gebruik van de BSN nummer: In de meeste bronbestanden zijn de gegevens gekoppeld aan het BSN en zijn die gegevens per definitie tot personen herleidbaar. Tegelijkertijd is het BSN bedoeld om gegevens/bestanden betrouwbaar te kunnen koppelen. Zonder een betrouwbare koppeling van gegevens is het vrijwel onmogelijk om tot een nauwkeurige risicoselectie te komen.

Voor het versleutelen van het BSN in de bronbestanden en de trainingcases is niet gekozen met als reden dat, gezien de omvang van het team dat toegang heeft tot de gegevens, dit een onevenredige inspanning zou zijn. Daar komt bij dat de bronbestanden los van het BSN ook tot personen herleidbare gegevens kunnen bevatten.

Supervised learning is alleen mogelijk met specifieke cases en dus niet met geaggregeerde gegevens. Voor het leren van het model worden dus de bronbestanden gekoppeld aan de BSN's van de trainingcases. Het BSN van een aanvraag wordt pas in de trainingsset opgenomen nadat een behandelaar de aanvraag heeft onderzocht en een beslissing heeft genomen. Er is dus altijd een menselijke tussenkomst voordat een aanvraag/beschikking als trainingcase wordt opgenomen. Als onderdeel van de maatregelen gepland voor Q1 2021 (zie Risico's en Maatregelen) zal ook de mogelijkheid voor pseudonimisatie onderzocht worden.

13. Doelbinding

Indien de persoonsgegevens voor een ander doel worden verwerkt dan oorspronkelijk verzameld, beoordeel of deze verdere verwerking verenigbaar is met het doel waarvoor de persoonsgegevens oorspronkelijk zijn verzameld.

We gebruiken de gegevens om toeslagen toe te kennen en dat moet rechtmatig gebeuren. De kerntaak van B/T is het rechtmatig, juist en tijdig toekennen van de toeslagen. Burgers verstrekken hun gegevens voor het aanvragen (of aanpassen c.q. stopzetten) van een toeslag. De in hoofdstuk 5 genoemde instanties en organisaties verstrekken gegevens voor het kunnen toetsen van de door de burgers verstrekte gegevens (contra-informatie). De oorspronkelijk verzamelde gegevens worden verwerkt voor dit hoofddoel. De verwerking "Risicoclassificatiemodel" is in overeenstemming met dit doel en is direct gekoppeld daaraan.

De Belastingdienst beoogt met 'handhaving' het gedrag van burgers en bedrijven zodanig te beïnvloeden dat zij structureel (fiscale) regels naleven (compliance). Dit moet zorgen voor de borging van de continuïteit van belastingopbrengsten en de rechtmatige betaling van toeslagen. De Belastingdienst handelt bij het handhaven binnen de juridische kaders van de wet en conform de beginselen van behoorlijk bestuur.



14. Noodzaak en evenredigheid

Beoordeel of de voorgenomen gegevensverwerkingen noodzakelijk zijn voor het verwezenlijken van de verwerkingsdoeleinden. Ga hierbij in ieder geval in op proportionaliteit en subsidiariteit.

- a. Proportionaliteit: staat de inbreuk op de persoonlijke levenssfeer en de bescherming van de persoonsgegevens van de betrokkenen in evenredige verhouding tot de verwerkingsdoeleinden?
- b. Subsidiariteit: kunnen de verwerkingsdoeleinden in redelijkheid niet op een andere, voor de betrokkene minder nadelige wijze, worden verwezenlijkt?

i

Het in artikel 6, eerste lid, aanhef en onder e, van de AVG neergelegde noodzakelijkheidsvereiste houdt in dat moet worden voldaan aan de beginselen van proportionaliteit en subsidiariteit. Het proportionaliteitsbeginsel houdt in dat de inbreuk op de belangen van de bij de verwerking van de persoonsgegevens betrokken personen niet onevenredig mag zijn in verhouding tot het met de verwerking te dienen doel. Op grond van het subsidiariteitsbeginsel mag het doel waarvoor de persoonsgegevens worden verwerkt in redelijkheid niet op een andere, voor de bij de verwerking van persoonsgegevens betrokken persoon, minder nadelige wijze kunnen worden verwezenlijkt. In het advies 06/2014 van de Groep gegevensbescherming artikel 29 (het huidige Europees Comité voor gegevensbescherming) staat dat de concrete toets is of minder inbreuk makende middelen beschikbaar zijn voor het bereiken van hetzelfde doel.

a. Proportionaliteit:

Ons doel als uitvoeringsorganisatie is om juist, tijdig en rechtmatig toeslagen uit te keren. Om met de beperkte middelen die we hebben zo efficiënt mogelijk om te gaan, is het RCM ontwikkeld om zo gericht mogelijk ons toezicht in te zetten. Om dit model te laten functioneren wordt er gebruik gemaakt van een set gegevens die niet elke keer volledig worden gebruikt. Per geval heb je niet altijd de volledige gegevensset nodig. Dit weet je echter pas nadat het model alle gegevens heeft gewogen vandaar dat je gegevens niet vooraf uit kunt sluiten. De reden daarvoor is dat het model inspeelt op risicopatronen uit eerdere casussen wat veroorzaakt dat de combinatie van gegevens altijd anders is. Als we van te voren gegevens uit zouden sluiten, zou de risicoscore niet meer zuiver zijn. Het is belangrijk dat het model goed werkt om de juistheid te waarborgen, wat belangrijk is voor het uitvoeren van onze wettelijke taak. De output van het model wordt altijd door een behandelaar gecheckt; op die manier wordt het risico voor de burger vermindert.

b. Subsidiariteit:

Andere oplossingen om rechtmatig toeslagen toe te kennen zouden de volgende kunnen zijn:

- Steekproefsgewijs controleren: maar dan worden ook burgers geraakt die veel minder risicovol zijn.
- Niks te doen vooraf en pas achteraf te controleren als we meer contra-informatie hebben om te gebruiken: maar dan kunnen burgers achteraf met hoge terugvorderingen worden geconfronteerd.

Vanuit uitvoeringsoogpunt is het niet haalbaar om alle aanvragen apart te behandelen en dan dus alleen de gegevens te verwerken die nodig zijn. Om toch zoveel mogelijk aanvragen in de toezichtbehandeling mee te nemen, helpt het om in te schatten welke aanvragen het meest risicovol zijn. Daarbij helpt het Risicoclassificatiemodel.

Om de inbreuk op de persoonlijke levenssfeer van de burgers te beperken zijn maatregelen geïmplementeerd zoals het uitvoeren van steekproeven om de nauwkeurigheid van het model te valideren en het gebruik van semi-geautomatiseerd besluitvorming met menselijk tussenkomst. Een andere mogelijke vorm van rechtsbescherming voor de burger is het indienen van een klacht/bezwaar bij B/T.

15. Rechten van de betrokkene

Geef aan hoe invulling wordt gegeven aan de rechten van betrokkenen. Indien de rechten van de betrokkene worden beperkt, bepaal op grond van welke wettelijke uitzonderingen dat is toegestaan.

Betrokkenen zijn de natuurlijke personen waarvan de persoonsgegevens worden verwerkt. Dit zijn de aanvragers van de kinderopvang en huurtoeslag, de eventuele toeslagpartner, medebewoners en de kinderen waarvoor kinderopvangtoeslag wordt verstrekt.

De rechten van betrokkenen moeten goed gewaarborgd worden. De B/T kent een geheimhoudingsplicht (artikel 2:5 Algemene wet bestuursrecht). Daarnaast worden betrokkenen op internet geïnformeerd over de wijze waarop de Belastingdienst uitvoering geeft aan de AVG (zie (https://www.belastingdienst.nl/wps/wcm/connect/bldcontentnl/niet_in_enig_menu/privie/privacy)).

"De doelen waarvoor de Belastingdienst gegevens verwerkt vloeien voort uit de wettelijke taken: heffen, uitkeren, innen, uitvoeren van goederentoezicht en opsporing van delicten die met al deze zaken samenhangen. We gebruiken de voor deze doelen verzamelde gegevens ook voor toezicht en dienstverlening.

Wij gebruiken gegevens voor het heffen en innen van Rijksbelastingen, premieheffing volksverzekeringen en andere belastingen die moeten worden betaald. Voorbeelden van Rijksbelastingen zijn loon- en inkomstenbelasting, motorrijtuigenbelasting, erf- en schenkelbelasting, overdrachtsbelasting, omzetbelasting, belastingen bij invoer van goederen, accijnzen. Dit doen we op basis van de Algemene wet inzake rijksbelastingen, de Invorderingswet en de verschillende belastingwetten.

De Belastingdienst gebruikt gegevens ook voor uitkeren van toeslagen. Het gaat om gegevens voor het berekenen van uw toeslagen en voor het uitbetalen daarvan. En als dat nodig is om teveel betaalde toeslagen van u terug te vragen. Het gaat om huurtoeslag, zorgtoeslag, kinderopvangtoeslag en kindgebonden budget. Dit doen we op basis van de Algemene wet inkomensafhankelijke regelingen en de verschillende toeslagenwetten.

De Belastingdienst neemt geen geautomatiseerde besluiten op basis van profilering in de zin van de AVG. De Belastingdienst maakt wel gebruik van persoonlijke aspecten om dossiers te selecteren voor toezicht op de juistheid van een aangifte of van een toegekende toeslag. Dossiers die op die manier zijn geselecteerd worden altijd door een mens beoordeeld en die persoon neemt ook het besluit."

Zo kunnen betrokkenen hier lezen dat zij op grond van de AVG bepaalde rechten hebben, en op welke wijze zij hiervan gebruik kunnen maken. Zij kunnen de B/T vragen om:

- a. inzage in de hem betreffende persoonsgegevens en een toelichting daarop (artikel 15 AVG)
- b. rectificatie van de hem/haar betreffende onjuiste persoonsgegevens (artikel 16 AVG); en
- c. onder voorwaarden verwijdering van gegevens (artikel 17 AVG).

Betrokkenen kunnen overigens ook hun gegevens inzien via Mijn toeslagen. Elke wijziging die wordt opgevoerd voor de toekenning van de toeslag is hier zichtbaar, en ook staan hier alle gegevens waarop de toeslag is gebaseerd. Verzoeken van betrokkenen op grond van de AVG komen centraal binnen bij B/CAP, waarbij generieke inzageverzoeken centraal worden afgehandeld met een algemene brief. Specifieke inzageverzoeken en verwijderings- en correctieverzoeken worden doorgezet naar de B/T. De reactie hierop is een voor bezwaar vatbaar besluit (artikel 34 Uitvoeringswet AVG). Het besluit wordt met ondersteuning van team Klacht genomen door (onder)gemandateerde medewerkers van team Vaktechniek, waarbij ook wordt beoordeeld of er beperkingen gelden op grond van artikel 23 van de AVG.

Specifiek over het Risicoclassificatiemodel:



- **Recht op inzage in de verwerkingen:** Dit is beperkt vanwege het feit dat als we het model prijsgeven is het niet meer mogelijk om effectief toezicht uit te voeren. Echter zouden betrokkenen informatie over de verwerking zonder alle details kunnen krijgen. Het inzagerecht kan in uitzonderingsgevallen geweigerd worden; bijvoorbeeld als er toezichtsbelangen spelen. (Zie Risico's en Maatregelen)
- **Recht op informatie:** de betrokkenen worden indirect én beperkt over het gebruik van het RCM geïnformeerd via de privacyverklaring die op de website van de belastingdienst staat. (Zie Risico's en Maatregelen)
Tekst:
De Belastingdienst neemt geen geautomatiseerde besluiten op basis van profilering in de zin van de AVG. De Belastingdienst maakt wel gebruik van persoonlijke aspecten om dossiers te selecteren voor toezicht op de juistheid van een aangifte of van een toegekende toeslag. Dossiers die op die manier zijn geselecteerd worden altijd door een mens beoordeeld en die persoon neemt ook het besluit.
- **Bezwaar maken tegen een besluit:** Aanvragers kunnen bezwaar maken tegen het resultaat van de behandeling van hun aanvraag.
Het transparantieprincipe hoeft niet zo uitgelegd te worden dat iedere betrokkene specifiek over de registratie van een risicosignaal geïnformeerd wordt (art. 67 AWR regelt geheimhouding). Dit wordt echter anders zodra er nadelige gevolgen voor de betrokkene kunnen optreden of zijn opgetreden. Betrokkenen kunnen over de informatie beschikken waarop deze beslissingen zijn genomen, zodat hij daartegen bezwaar kan maken.
- **Recht op beperking of beëindiging van de verwerking en recht op persoonsgegevens te verwijderen ('recht op vergetelheid');**
De inperking van deze twee rechten ontstaat als gevolg van de toeslagaanvraag. Zodra er een toeslag is toegekend kan de toeslagaanvrager niet meer verzoeken om beperking of eindiging van de verwerking. Dit ligt aan het feit dat B/T tot vijf jaar na afgifte definitieve beschikking een herziening mag afgeven op basis van nieuwe feiten. Daarnaast is het ook voor de toeslagaanvrager mogelijk om nog binnen vijf jaar na afgifte van een definitieve beschikking een verzoek om herziening aan te vragen. Ook het beperken van het recht op vergetelheid hangt hier mee samen.
- **Recht om aan uitsluitend geautomatiseerde besluitvorming te worden onderworpen:**
Het verwerken van persoonsgegevens in het risicoclassificatiemodel kwalificeert als profilering in de zin van artikel 4, aanhef en onder 4, van de AVG. Aanvragen die het risicoclassificatiemodel selecteert worden handmatig beoordeeld door een medewerker van Toeslagen. Er is door deze menselijke tussenkomst geen sprake van geautomatiseerde besluitvorming als bedoeld in artikel 22 van de AVG.

C. Beschrijving en beoordeling risico's voor de betrokkenen

Beschrijf en beoordeel de risico's van de voorgenomen gegevensverwerkingen voor de rechten en vrijheden van de betrokkenen. Houd hierbij rekening met de aard, omvang, context en doelen van de gegevensverwerking zoals in onderdeel A en B zijn beschreven en beoordeeld. Het gaat hierbij overigens niet om de risico's van de verwerkingsverantwoordelijke zelf.

16. Risico's

Beschrijf en beoordeel de risico's van de gegevensverwerkingen voor de rechten en vrijheden van betrokkenen. Ga hierbij in ieder geval in op:

- welke negatieve gevolgen de gegevensverwerkingen kunnen hebben voor de rechten en vrijheden van de betrokkenen;
- de oorsprong van deze gevolgen;
- de waarschijnlijkheid (kans) dat deze gevolgen zullen intreden;
- de ernst (impact) van deze gevolgen voor de betrokkenen wanneer deze intreden.

Nr.	Risico	Oorsprong	Kans ⁵¹ (1-5)	Impact ⁵² (1-5)
1.	Het risico kan ontstaan dat het resultaat onnauwkeurig en/of discriminerend is en dat het model niet meer representatief is.	Regressiemodellen moeten vaak getraind worden met actuele trainingcases, maar dit gebeurt niet altijd.	3	4 Als het model de verkeerde indicatoren gebruikt dan bestaat de kans dat er burgers een onterecht hoge score van het model krijgen, wat een vertraging van het ontvangen van toeslag kan betekenen.
2.	Het risico kan ontstaan dat gegevens onrechtmatig worden gebruikt en dat er onnodig gegevens worden verzameld (gebrek aan dataminimalisatie).	Verwerking van gegevens die niet nodig zijn voor het model en waar geen doelbinding bestaat.	3	4 Als er data onrechtmatig wordt gebruikt en dit is er de oorzaak van dat het een hogere score oplevert, dan zal de aanvraag naar Behandeling gaan, wat een vertraging van het ontvangen van toeslag kan betekenen.
3.	Het risico kan ontstaan dat de verdeling in het model niet evenredig is (selectiebias).	Wij onderzoeken alleen de hoge scores en er is geen toezichtproces dat ook andere kenmerken onderzoekt.	3	3 Impact is middelmatig. Bij het ontvangen van de definitieve toekenning kunnen burgers zien of ze te veel of te weinig voorschot hebben gekregen, of precies het juiste bedrag.
4.	Het risico kan ontstaan op het niet naleven van het transparantiebeginsel van de AVG.	De informatie in de privacyverklaring is beperkt	3	3 De burgers zouden geen duidelijke informatie over de verwerking krijgen.
5.	Het risico kan ontstaan dat het resultaat niet aan de eisen van de opdracht voldoet.	De opdracht is niet altijd duidelijk door o.a. het gebrek aan specifieke doelstellingen	1	4 Als het model – door onduidelijkheid in de opdracht – de verkeerde indicatoren gebruikt dan bestaat de kans dat de verkeerde burgers een onterecht hoge score van het model krijgen, wat een vertraging van het ontvangen van toeslag kan betekenen.
6.	Het risico kan ontstaan dat de gegevens die verwerkt worden in het model niet altijd correct zijn.	Gegevens worden uit verschillende bronnen verzameld. Het is mogelijk dat de gegevens niet correct of geüpdatet vanuit de	1	4 De burgers zouden een onterecht hoge score kunnen krijgen wat een vertraging

Met opmerkingen [65]: Zie citaten bij maatregelen, representativiteit is door niet geschiktheid/ouderdom gegevens ook een risico.

Met opmerkingen [66]: Er zitten ook veel handmatige afwegingen/keuzes in het proces, waarbij niet duidelijk is wat de eisen aan de medewerkers en de criteria zijn op grond waarvan dat gebeurt. Ook op dat punt maatregelen nemen lijkt mij.

Met opmerkingen [67]: Vage omschrijving. Duidelijk maken dat het hier gaat om het doel van de specifieke opdracht voor selectie van posten voor toezicht met behulp van het model. Anders is de laatste kolom ook niet duidelijk, want daarin lijkt het te gaan om fouten in het model. Maar het gaat dus om het kiezen van verkeerde indicatoren omdat de opdracht vaag is (daarom alvast tekstvoorstel opgenomen).

⁵¹ Kans op een schaal van 1 tot 5, waarbij 1 een lage kans is en 5 een hoge kans.

⁵² Impact op een schaal van 1 tot 5, waarbij 1 een lage impact is en 5 een hoge impact.

		oorspronkelijke bron zijn ontvangen.		van het ontvangen van toeslag kan betekenen.
7.	Het risico kan ontstaan dat de output van het model niet meer uitlegbaar is, wat een probleem zal vormen op het moment dat een burger zijn recht op inzage wil uitoefenen.	De logica van het model kan complex zijn. De risicoscore per indicator is niet altijd uitlegbaar.	2	3 De burgers zouden geen duidelijke informatie over de verwerking krijgen.
8.	Het risico kan ontstaan dat er per ongeluk gegevens naar een verkeerd e-mailadres verstuurd worden. Het gaat met name over een intern datalek binnen B/T.	Gegevens worden vaak per e-mail verstuurd. Op dit moment kan DAT niet extern mailen, maar andere afdelingen wel.	2	3 Als de gegevens lekken en/of worden misbruikt dan zou dat impact op de reputatie van de burger kunnen hebben. De impact is middelmatig omdat de output bestanden bevatten alleen BSN nummers en geen andere persoonsgegevens.
9.	Het risico kan ontstaan op een inbreuk op het beginsel van opslagbeperking.	Gegevens worden te lang bewaard en er is geen vernietigingsproces voor de Q-schijf en SAS.	3	2 Impact is laag voor de burger omdat de gegevens intern blijven.
10.	Het risico kan ontstaan dat medewerkers toegang hebben tot informatie die zij niet (meer) nodig hebben om hun werk te kunnen doen.	Autorisatiebeleid kan strenger zijn voor bepaalde producten binnen SAS en de Q-schijf.	2	2 De informatie blijft intern dus de impact voor de betrokkene is zeer laag.

Met opmerkingen ☐ **68]:** Maar dat geven we ze nu toch al niet? Uit oogpunt van handhavingsbeleid (strategie niet prijsgeven)? Dat wordt eerder opgemerkt in de PIA.

Met opmerkingen ☐ **69]:** Maar BSN's die op straat liggen zijn op zich al een groot risico (flinke reputatieschade BD met name, maar toch, daar zitten we ook niet op te wachten denk ik...).

D. Beschrijving voorgenen maatregelen

In onderdeel D wordt gezien welke maatregelen kunnen worden getroffen om de in onderdeel C erkende risico's te voorkomen of te verminderen. Welke maatregelen in redelijkheid worden getroffen is een belangenafweging van de wetgever of verwerkingsverantwoordelijke. Voor dit onderdeel van de PIA is, als het gaat om beveiligingsmaatregelen, expertise over informatiebeveiliging belangrijk.

17. Maatregelen

Beoordeel welke technische, organisatorische en juridische maatregelen in redelijkheid kunnen worden getroffen om de hiervoor beschreven risico's te voorkomen of te verminderen. Beschrijf welke maatregel welk risico aanpakt en wat het restrisico is na het uitvoeren van de maatregel. Indien de maatregel het risico niet volledig afdekt, motiveer waarom het restrisico acceptabel is.



Risico nr.	Maatregelen	Restrisico	Restrisico geaccepteerd
1	- Het model regelmatig trainen en testen: één keer voor dat het model gaat draaien en daarna iedere 2 maanden. Werk: +- 200 uur/jaar (DAT en HHR) - Oude trainingcases verwijderen. Statistisch checken of er verschillen zijn. In dat geval dmv steekproeven de nauwkeurigheid van het model meten. Werk: +- 8 uur (DAT) en 50 uur HH1 en Productieregie. Planning: Q1 2021	Capaciteit is niet altijd beschikbaar. Oplossing: minimaal 20% capaciteit posten nodig voor aselecte. Een team zou voor een keer per maand beschikbaar moeten komen.	☒
2	Gegevens periodiek reviewen en toetsen op dataminimalisatie en grondslag. Check op aanwezigheid van proxies. Waar nodig gegevens niet meer als bron gebruiken en het model hertrainen. Een audit proces vastleggen en goed documenteren. Werk: +-100 uur (DAT), +-50 uur (HHR) Planning: Q1 2021	Het model kan minder voorspellend worden. Actie: behandelen/wachten/trainen	☒
3	- Steekproeven gebruiken voor het trainen van het model en gestratificeerde aselecte steekproef uitvoeren voor het vinden van nieuwe cases. Ook naar de lagere scores kijken om evenredigheid te garanderen. - Iedere 6 maanden overwegen of de aselecte steekproef nodig is als er veranderingen ontstaan in het patroon. - Meer capaciteit inzetten voor het checken van de 20% groene posten. Hiermee wordt het model niet alleen getoetst op onderscheidend vermogen, maar ook getraind op rechtmatige cases. Werk: valt binnen de uren ingeschat voor risico 1.		n.v.t.
4	Vanuit B/T wordt het initiatief opgepakt om samen met IV&D te bekijken in welke mate het privacy statement van de Belastingdienst, zoals gepubliceerd op belastingdienst.nl, moet worden aangepast. Indien een aanpassing noodzakelijk blijkt wordt een tekstvoorstel met verzoek tot aanpassing ingediend. Daarnaast komt een addendum voor Toeslagen op de privacyverklaring van de B/T. Risico eigenaar: [] Toeslagen. Planning: Q1 2021		n.v.t.
5	Steekproef uitvoeren. Deze is onderdeel van de maatregelen uit risico 1. Uit de steekproef moet naar voren komen wat de doelstellingen zijn.		n.v.t.

Met opmerkingen [70]: Voor 5 maatregelen is Q1 21 genoemd als realisatiedatum. Dat betekent dat voor het overleg met de TK op 1 februari concrete tussenresultaten beschikbaar moeten zijn die aangeven dat dat ook haalbaar is.

Voor 5 maatregelen zijn geen of niet concrete data opgenomen. Deze moeten nog worden aangevuld.

Met opmerkingen [71]: citaat: Het model wordt sinds een jaar geleden niet meer getraind maar het proces wordt herhaald voor het voorbereiden van de scorecard. Steekproeven worden weinig uitgevoerd vanwege gebrek aan capaciteit.

Moet de maatregel niet zijn dat het proces van ontwikkelen en trainen opnieuw wordt ingericht met betere gegevens en een goed proces voor verversing ed.?

Met opmerkingen [72]: citaat: Ondanks de wisselende kwaliteit vormt deze verzameling posten een substantieel deel van de trainingsset. In december 2015 is besloten om cases geïdentificeerd in 2013 te bevroren, en geen nieuwe cases voor toeslagjaar 2013 meer toe te voegen. Daarom worden alle cases met toeslagjaar 2013 of eerder hier al uitgesloten van verdere verwerking. Met ingang van april 2019 zijn ook de jaren 2014 en 2015 bevroren. Sinds 2019 zijn geen cases meer toegevoegd.

Met opmerkingen [73]: Wb bias: ook verbeteringen voor de menselijke afwegingen/keuzes die in het proces worden gemaakt. Zie ook opmerking hierboven bij risico.

Met opmerkingen [74]: En dus ook in Q1 21 gerealiseerd? Datum nog opnemen.

Met opmerkingen [75]: Moet niet meer algemeen gekeken worden naar aanpassing van het privacy statement om betere informatie te geven over wat we doen met risicomodellen? Dan zou IV&D het voortouw bij deze actie moeten hebben lijkt mij.

Met opmerkingen [76]: Is één steekproef wel voldoende? En een steekproef kan de doelstellingen op niveau van de opdracht toch niet helder maken?? Dat moet de business toch doen? Daar wordt toch bepaald waarop posten moeten worden geselecteerd voor toezicht? Het lijkt mij dat de maatregel hier zou moeten zijn dat de opdrachtformulieren bekeken worden op geschiktheid voor het doel, en dat duidelijke afspraken worden gemaakt over de eisen die aan een ingevuld formulier worden gesteld, en wat het gevolg is als het niet duidelijk genoeg is (terugsturen, aanpassen, of misschien ook al vooraf meer toetsing, verstevigen van proces van vaststelling van de opdrachten).

Ook: datum voor realisatie noemen.

6	Er zijn afspraken met de leveranciers van de gegevens mbt de actualiteit van de gegevens. Toch is er altijd een risico dat de gegevens die niet uit de BRP (basis registratie) komen niet correct of geüpdatet vanuit de oorspronkelijke bron zijn ontvangen. Dit risico is heel laag en het wordt geaccepteerd.		n.v.t.
7	DAT heeft een Privacy by Design training gekregen (en er volgen nog een paar sessies) waar analisten leren om dit risico te identificeren en te behandelen. Daarnaast is DAT bezig om interne processen en checklists te creëren om o.a. dit soort risico's te voorkomen.		n.v.t.
8	De nieuwe werkwijze van Analytics United zal dit risico mitigeren. Naar verwachting wordt het in Q1 2021 geïmplementeerd.	Het risico blijft ontstaan tot de maatregel geïmplementeerd wordt.	☒
9	- Het implementeren van een proces om de gegevens te vernietigen gebaseerd op de Selectielijst. Hiervoor loopt het project informatie huishouding (voor e-mail en Q-schijf). - Maatregelen implementeren voor het vernietigen van gegevens in SAS: Werk: +- 24 uur (DAT). Sessies inplannen om te overleggen, maatregelen vastleggen, met [] afstemmen en acties ondernemen.		n.v.t.
10	Planning: Q1 2021 Een duidelijk en strenger autorisatiebeleid voor het RCM team implementeren. Audit-trail is al geïmplementeerd voor SAS. Risico eigenaar: DAT teamleader Deadline: het wordt geïmplementeerd vóór de volgende run.		n.v.t.

Met opmerkingen [77]: Is dat zo??
En waarom is dit anders bij de BRP?? Komen de gegevens van de BRP niet via BvR bij Toeslagen? Er is al gebleken dat BvR niet altijd voldoet (nationaliteit).
Hierop toch maatregelen formuleren? Minimaal afspraken met CAP over kwaliteit gegevens??

Met opmerkingen [78]: Documentatie van keuzes die gemaakt zijn bij bepalen van de indicatoren, de juridische onderbouwing daarvan (blijven ze binnen de wettelijke kaders) lijkt me hier een heel belangrijke maatregel. Dat zou je onder processen kunnen scharen, maar dat vind ik wel wat hoog over. Meer concretiseren dus wat er gaat gebeuren.

En wat is dat realisatiedatum? Ook opnemen.

Met opmerkingen [79]: Zie in hoofdstuk: meer uitleggen wat dat is als het zo belangrijk is als mitigerende maatregel.
En hoe zeker is het dat dit i.d. in Q1 21 ingevoerd wordt? Wat als het niet gehaald wordt? Is dan wat anders nodig?

Met opmerkingen [80]: Wanneer is dat? Datum concreet maken.

Bijlage 1:**Bronnen gebruikt in het risicoclassificatiemodel:**

Bronnen	Omschrijving	Oorspronkelijke bron	Ontvangsmethode	Opslag
6PP	Tabel met alle geldige postcodes in Nederland, inclusief lengte- en breedtegraad	d-centralize.nl	Csv-bestand gedownload	Q-schijf en SAS
BAG	Basisregistratie adressen en gebouwen	CIV	Csv-bestand per e-mail	Q-schijf en SAS, e-mail
Beschikkingen HT	Lijst met formele HT beschikkingen en de details hiervan	CIV	SAS-bestand	SAS
Beschikkingen KOT	Lijst met formele KOT beschikkingen en de details hiervan	CIV	SAS-bestand	SAS
Betrokkenen HT	Lijst met betrokken BSN's per formele HT beschikking	CIV	SAS-bestand	SAS
Bezwaren	Lijst met BSN's waarvoor recent door de Belastingdienst een bezwaar is afgehandeld (stellaposten)	Stella team	xlsx-bestand per e-mail	Q-schijf, SAS, e-mail
BVR	Beheer van relaties: adressenbestand van de Belastingdienst, mede gevoed door GBA en BRP, met persoonsdetails.	via het datafundament	SAS-bestand	SAS
COA	Centrale Ontvangers Administratie - BSN's met openstaande vorderingen bij de Belastingdienst	CAP (team Inning)	SAS-bestand via de Q-schijf	SAS, Q-schijf
Deeljaar beschikking HT	Details van HT beschikkingen per deel van het toeslagjaar.	CIV	SAS-bestand	SAS
Deeljaar beschikking KOT	Details van KOT beschikkingen per deel van het toeslagjaar.	CIV	SAS-bestand	SAS
Excepties	Lijst met BSN's die niet uitgesloten mogen worden.	HHR	xlsx-bestand per e-mail op de dag van de definitieve run	Q-schijf, SAS, e-mail
Holland2stay	Lijst met aanvragen HT van buitenlandse studenten. Doel: horizontaal toezicht (BSN's op de lijst worden niet gecontroleerd omdat de verhuurder heeft dat al gedaan).	Woningbouwvereniging Holland2stay	xlsx-bestand per e-mail een dag voor de definitieve run	Q-schijf, SAS, e-mail
Hot Hor	Lijst van BSN's waarvoor in TVS een actuele conceptbeschikking met een hoog bedrag (dus een hoog risico) klaarstaat en waarvoor nog geen handmatige controle is toegepast.	KPT_Werkopdracht view uit het datafundament	SAS-bestand	SAS
IGS	Lijst van BSN's die de IGS (ingebrekestelling) status hebben.	KPT_Werkopdracht view uit het datafundament	SAS-bestand	SAS
KOT conceptbeschikkingen	Lijst met formele KOT conceptbeschikkingen en de details hiervan	CIV	SAS-bestand	SAS
LRK	Register van Kinderopvang instellingen	DUO via CAP	Txt-bestand op de Q-schijf	Q-schijf en SAS
Meldingen FRS	Verzameling van door TVS gegenereerde meldingen.	CIV	SAS-bestand	SAS
Onbekende aanvragers	Lijst van BSN's die minder dan 1 jaar oud zijn, waarvoor nog nooit een formele beschikking met een bedrag > 0 is uitgereikt en waarvoor nu een	CIV	SAS-bestand	SAS

(nieuwe aanvragers)	conceptbeschikking met bedrag > 0 klaarstaat. De brondata wordt gebruikt in het halfproduct 'Uitsluitijst onbekende aanvragers Zorg'.			
Ouder Kind	Matrix met relaties tussen BSN's van ouders en BSN's van hun kinderen	CIV	SAS-bestand	SAS
Uitgesloten adressen	Lijst van fysieke adressen waarvoor geen recht op huurtoeslag mogelijk is.	Toeslagen verhuurder administratie_Post bus	xlsx-bestand per e-mail	Q-schijf, SAS, e-mail
Uitsluitijst	Actuele lijst van BSN's waarvoor toeslagrecht over een bepaalde periode is geblokkeerd.	CIV	SAS-bestand	SAS
Uitval	Lijst van in TVS ontbrekende informatie op een of meerdere toeslagaanvragen per BSN	TQL (TVS)	Txt-bestand op SAS Grid	SAS
VHA	Verhuurders Administratie - door Belastingdienst opgestelde lijst van huurbedragen per fysiek adres volgens verhuurders.	CIV	SAS-bestand	SAS
VIP	Lijst van binnen de Belastingdienst beschermde BSN's die niet voor elke behandelaar inzichtelijk zijn. De brondata wordt gebruikt in halfproducten HT_Model en KOT_Model.	CIV	SAS-bestand	SAS
Vluchtelingenwerk	Lijst met huurtoeslagaanvragen voor vluchtelingen, aangevraagd door de organisaties Vluchtelingenwerk en COA (Centraal Orgaan opvang Asielzoekers). Doel: horizontaal toezicht (BSN's op de lijst worden niet gecontroleerd omdat de COA heeft dat al gedaan).	CIV	SAS-bestand	SAS
Wanbetalers	Lijst met BSN's waarvan vanuit CVZ bekend is dat geen zorgpremie wordt afgedragen.	CIV	SAS-bestand	SAS
WOZ	Lijst van actuele onroerend goed bezittingen per BSN	CIV	SAS-bestand	SAS

Bijlage 2:**Halfproducten gebruikt in het risicoclassificatiemodel:**

Halfproducten	Omschrijving
GPP	Tabel met alle geldige postcodes in Nederland, inclusief lengte- en breedtegraad
BAG	Basisregistratie adressen en gebouwen
Bezwaren	Lijst met BSN's waarvoor recent door de Belastingdienst een bezwaar is afgehandeld.
BVR	Beheer van relaties: adressenbestand van de Belastingdienst, mede gevoed door GBA, met persoonsdetails.
COA	Tabel met BSN's die een openstaande vorderingen hebben bij de Belastingdienst.
DTM_HT	In deze flow worden de beschikkingen plat geslagen tot een tabel op deeljaar niveau. Hieraan worden de conceptbeschikkingen toegevoegd, en vervolgens gekoppeld aan de (inkomens) van betrokkenen.
DTM_KOT	In deze flow worden de beschikkingen plat geslagen tot een tabel op deeljaar niveau. Hieraan worden de conceptbeschikkingen toegevoegd, en vervolgens gekoppeld aan een aantal halfproducten.
Holland2stay	Lijst met aanvragen HT van buitenlandse studenten.
Hot Hor	Lijst van BSN's waarvoor in TVS een actuele conceptbeschikking met een hoog bedrag staat.
IGS	Selectie gemaakt op de werkitems mbt een ingebrekestelling en met status niet completed.
LRK	Tabel met kinderopvangstelling en een aantal details daarvan volgens het landelijk register.
Melding8	Tabel met meldingen 8 (nieuwe toeslagaanvraag) uit TVS per BSN.
MeldingFRS	Tabel met aantal opgegeven verhuizingen (melding 6), aantal veranderde betaaltwijzen per jaar (melding 15), aantal wijzigingen in aantal uren opvang per jaar (melding 18), aantal wijzigingen in de huurprijs (melding 19) per BSN.
Ouder Kind	Tabel met BSN's van de ouder(s) per kind.
Uitgesloten adres	Tabel met adressen waarvoor geen toeslagrecht bestaat.
Uitsluiten_onbekende_aanvragers_zorg	Lijst van BSN's met afgewezen aanvragen zorgtoeslag in de afgelopen jaren (geen recht vanwege buitenlands adres). Dit halfproduct wordt niet in het RCM gebruikt maar in het project Zorgverzekerd.
Uitsluitlijst	Tabel met BSN's die op een gegeven moment in de tijd uitgesloten worden van beschikkingen voor een bepaalde toeslagregeling.
Uitval	Lijst van in TVS ontbrekende informatie op een of meerdere toeslagaanvragen per BSN
VHA	Verhuurders Administratie - door Belastingdienst opgestelde lijst van huurbedragen per fysiek adres volgens verhuurders.
Vluchtelingenwerk	Lijst met aanvragen HT door vluchtelingen
WOZ	Lijst van actuele onroerend goed bezittingen per BSN

Bijlage 3:**Indicatoren HT:**

Label	Omschrijving	Bron
KALE_HUUR	Huurprijs exclusief de servicekosten	DTM_HT
aantal_huival53	Is er opgegeven of er een medebewoner is?	SAS
av_aant_vorderingen	Heeft de belastingdienst een vordering op deze persoon?	SAS
av_adr_aant_inschrijvingen	Hebben we het aantal ingeschrevenen kunnen bepalen?	SAS
av_groei_inkomen	Hebben we de groei van het inkomen kunnen vaststellen?	SAS
draagkracht_ht	Hebben we de draagkracht vast kunnen stellen?	SAS
opg_totale_huur	Is de totale huur bekend?	SAS
oppervlakte_per_persoon	Oppervlakte per persoon	HT_MODEL
oppervlakteprijs	Huurprijs per vierkante meter	HT_MODEL
perc_huur_draagkracht	Ratio huurprijs tov draagkracht	HT_MODEL
perc_huur_VHA_huur	Verschil huurprijs opgeven VHA en verhuurder	HT_MODEL
perc_huurprijs_huurprijs_omg	Huurprijs t.o.v. huurprijs in omgeving	HT_MODEL
perc_huur_tov_huurgrens	Huurprijs t.o.v. huurgrens	HT_MODEL
TERUGWERKENDE_KRACHT	Mate van terugwerkende kracht (obv melding 8) aangevraagd door (niet namens) burger zelf.	Melding8

Indicatoren HT die niet meer worden gebruikt: de volgende indicatoren zijn op reject gezet in het model, echter worden de brongegevens verwerkt in de data preparatie. Die zal in Q1 2021 structureel uit het model verwijderd worden. (zie Risico 2 in hoofdstuk 16 "Risico's" en 17 "Maatregelen")

Label	Omschrijving	Bron
av_adr_aant_kinderen	Aantal ingeschreven kinderen op adres - volgens GBA	BVR
av_adr_aant_volwassenen	Aantal ingeschreven volwassenen op adres - volgens GBA	BVR
av_eigenaar_pand	Of de aanvrager/toeslagpartner van een toeslag ook eigenaar is van dat pand	HT_MODEL
Huur_deelbaar_door_10	Huurprijs deelbaar door 10?	DTM_HT
Huur_deelbaar_door_50	Huurprijs deelbaar door 50?	DTM_HT
aant_wijz_huurprijs_burger	Indicatie of het bekend is hoe vaak de burger zijn huurprijs aangepast heeft	SAS
aantal_betrokkenen	Indicatie of het aantal betrokkenen is opgegeven	SAS
aantal_huival	Is er sprake van een onvolledig ingevulde aanvraag?	SAS
av_aant_verhuiz_opgegeven	Hoeveel verhuizingen zijn er door de burger opgegeven?	SAS
av_aant_verhuizingen_gba	Hoe veel verhuizingen zijn er volgens het GBA bij deze aanvrager?	SAS
av_inkomen_ht	Hebben we het inkomen kunnen vaststellen?	SAS
av_leeftijd_bsn	Hebben we de leeftijd van de aanvragers BSN kunnen vaststellen?	SAS
av_nationaliteit*	Hebben we de nationaliteit van de aanvrager kunnen vaststellen? Nederlandschap ja/nee (waarden 0 en 1). * Deze gegevens zullen in Q1 2021 worden verwijderd.	SAS
av_woont_in_nl	Hebben we vast kunnen stellen of de aanvrager in Nederland woont?	SAS
av_openstaande_vorderingen	Hebben we openstaande vorderingen gevonden?	SAS
av_status_vorderingen	Hebben we de status van vorderingen kunnen vaststellen?	SAS
overleden	Is persoon overleden?	HT_MODEL
perc_oppervlakteprijs_omg	Oppervlakteprijs t.o.v. oppervlakteprijs in de omgeving	HT_MODEL
totale_service_kosten	Hoe hoog zijn de totale service-kosten?	DTM_HT
wanbetalen_zorg	Indicatie of bij CVZ bekend is dat geen zorgpremie wordt afgedragen voor betreffende BSN. * Deze gegevens zullen in Q1 2021 worden verwijderd.	HT_MODEL

Bijlage 4:**Indicatoren KOT:**

Label	Omschrijving	Bron
aant_kinderen_in_opvang	Aantal kinderen waarvoor KOT is aangevraagd per BSN van de ouder.	DTM_KOT
afst_adr_opvang	Afstand in kilometers tussen huisadres aanvrager en adres van de kinderopvang volgens LRK.	DTM_KOT, LRK
av_aant_verhuizingen_gba	Aantal verhuizingen volgens de BVR.	BVR
av_aant_vorderingen	Aantal vorderingen per BSN.	COA
av_adr_aant_kinderen	Aantal ingeschreven kinderen per adres volgens de BVR.	BVR
av_adr_aant_volwassenen	Aantal ingeschreven volwassenen per adres volgens de BVR.	BVR
av_groei_inkomen	Groei van het jaarinkomen (inkomen TVS tov inkomen BVR). Dit ligt tussen -5 en 5.	KOT_Model.Merge_BS N, DTM_KOT, BVR
av_leeftijd_bsn	Leeftijd BSN (tov) gewenste ingangdatum (uit aanvraag). Om te bepalen hoe lang iemand al in Nederland is. Het zal in Q1 2021 structureel uit het model verwijderd worden. (zie Risico 2 in hoofdstuk 16 "Risico's" en 17 "Maatregelen")	DTM_KOT, BVR
av_missend_voorgaande_ink	Indicatie of inkomen missend is in BVR.	KOT_Model.Merge_BS N, BVR
av_nul_inkomen_opgegeven	Indicatie of opgegeven inkomen in TVS nul is.	KOT_Model.Merge_BS N, DTM_KOT
av_openstaande_vorderingen	Bedrag aan openstaande vorderingen per BSN.	COA
av_recht_op_kgb	Indicatie of aanvrager of partner recht heeft op KGB voor alle kinderen volgens informatie van SVB.	DTM_KOT
av_status_vorderingen	Status van de openstaande vordering(en) volgens COA.	DTM_KOT
draagkracht_KOT	Draagkracht van het huishouden volgens TVS waartoe BSN volgens TVS behoort.	DTM_KOT
Eigen_bijdrage_per_maand	Verschil tussen aangevraagde uren aangevraagd tarief - de toegewezen toeslag.	DTM_KOT
leeft_oudste_kind	Leeftijd van het oudste kind waarvoor KOT wordt aangevraagd volgens BVR.	DTM_KOT
Leeftijd_bij_geboorte	Leeftijd van de aanvrager volgens de BVR op het moment van de toeslagaanvraag.	DTM_KOT
Max_uren	Aantal uur dat de aanvrager (minst werkende partner) zou moeten werken per maand om aantal aangevraagde uren kinderopvang te rechtvaardigen	KOT_Model.Merge_BS N
ouder_kind_rel_mist	Indicatie of volgens BVR geen sprake is van ouder-kind relatie tussen aanvrager of partner van de aanvrager en het kind waarvoor KOT wordt aangevraagd.	DTM_KOT
perc_inkomen_eigen_bijdrage	Percentage van het door de aangever opgegeven inkomen volgens zelf uitgegeven wordt aan opvang	DTM_KOT
terugwerkende_kracht	Aantal dagen tussen datum waarop toeslag is aangevraagd en de ingangsdatum van het tijdvak waarop de aanvraag betrekking heeft.	
uitval_kot	Aantal uitvalmeldingen in TVS voor de actuele toeslagaanvraag KOT.	KOT_Model.Merge_BS N
heeftpartner	Geeft aan of de aanvrager een partner heeft of niet.	CONCEPTBESCHIKKIN GEN_KOT

Indicatoren KOT die niet meer worden gebruikt: de volgende indicatoren zijn op reject gezet in het model, echter worden de brongegevens verwerkt in de data preparatie. Die zal in Q1 2021 structureel uit het model verwijderd worden. (zie Risico 2 in hoofdstuk 16 "Risico's" en 17 "Maatregelen")

Label	Omschrijving	Bron
aantal_deelbaren	Aantal delen waaruit het toeslagjaar voor de betreffende BSN bestaat.	DTM_KOT
av_07_01_adr_niet_gelijk	Indicatie of '07 adres' (Belastingdienst) niet gelijk is aan '01 adres' (BVR): bij 1 zijn adressen niet gelijk, anders 0.	BVR
av_aant_verhuiz_opgegeven	Aantal verhuizingen opgegeven in TVS.	MeldingFRS
av_adr_aant_inschrijvingen	Aantal inschrijvingen per adres volgens de BVR.	BVR
av_adr_gelijk_aan_ov_adr	Indicatie of adres van opvanglocatie gelijk is aan adres	DTM_KOT, LRK
av_ht_uitgesloten_adres	Adres uitgesloten voor HT volgens halfproduct Uitgesloten adres.	Uitgesloten adres
av_leeftijd	Leeftijd van de aanvrager volgens BVR	BVR
av_nationaliteit	Nationaliteit cluster van de aanvrager (1 = Nederlander, 0 = niet Nederlander) volgens de BVR. * Deze gegevens zullen in Q1 2021 worden verwijderd.	KOT_Model.Merge_BSN, BVR
av_verschil_verhuizingen	Verschil tussen aantal in TVS opgegeven verhuizingen en aantal verhuizingen volgens de BVR.	KOT_Model.Merge_BSN, MeldingFRS, BVR
av_woont_niet_in_nlbedu	Indicatie of aanvrager in Nederland, België of Duitsland woont.	KOT_Model.Merge_BSN, BVR
ind_bso_go	Indicatie of toeslagaanvraag betrekking heeft op buitenschoolse opvang door een gastouder	DTM_KOT
ind_bso_kc	Indicatie of toeslagaanvraag betrekking heeft op buitenschoolse opvang door een kindercentrum	DTM_KOT
ind_do_go	Indicatie of toeslagaanvraag betrekking heeft op dagopvang door een gastouder	DTM_KOT
ind_do_kc	Indicatie of toeslagaanvraag betrekking heeft op dagopvang door een kindercentrum	DTM_KOT
kind_niet_in_NLBEDU	Indicatie of het kind waarvoor KOT wordt aangevraagd volgens de BVR woonachtig is in Nederland, België of Duitsland	DTM_KOT
kind_woont_bij_av	Indicatie of het kind waarvoor KOT wordt aangevraagd volgens de BVR woonachtig is bij de aanvrager	DTM_KOT
leeft_jongste_kind	Leeftijd van het jongste kind waarvoor KOT wordt aangevraagd volgens BVR.	DTM_KOT
leeftijd_BSN_kind	Tijd tussen het aanmaakmoment van de BSN van het kind waarvoor KOT wordt aangevraagd volgens de BVR en de ingangsdatum van de toeslagaanvraag.	DTM_KOT
leeftijd_ov_issues	Indicatie of de leeftijd van het kind volgens BVR waarvoor KOT is aangevraagd overeen komt met het type opvang (yes/no)	DTM_KOT
max_uren_bso	Aantal uur dat de aanvrager (minst werkende partner) zou moeten werken per maand om aantal aangevraagde uren kinderopvang te rechtvaardigen	DTM_KOT
max_uren_do	Het hoogste aantal uren dagopvang voor 1 kind in de toeslagaanvraag.	DTM_KOT
max_uurtarief_bso	Het hoogst opgegeven uurtarief voor buitenschoolse opvang in de toeslagaanvraag.	DTM_KOT
max_uurtarief_do	Het hoogst opgegeven uurtarief voor dagopvang in de toeslagaanvraag.	DTM_KOT
pa_dagtaaktype	Dagtaaktype van de toeslagpartner als bekend in TVS.	DTM_KOT

Te Jong Voor BSO	Indicatie of kind waarvoor toeslag voor buitenschoolse opvang is aangevraagd te jong is voor recht op toeslag voor buitenschoolse opvang.	DTM_KOT
Te Oud Voor BSO	Indicatie of kind waarvoor toeslag voor buitenschoolse opvang is aangevraagd te oud is voor recht op toeslag voor buitenschoolse opvang.	DTM_KOT
Te Oud Voor DO	Indicatie of kind waarvoor toeslag voor dagopvang is aangevraagd te oud is voor recht op toeslag voor dagopvang.	DTM_KOT
wanbetalen zorg	Indicatie of bij CVZ bekend is dat geen zorgpremie wordt afgedragen voor betreffende BSN. * Deze gegevens zullen in Q1 2021 worden verwijderd.	KOT_Model.Merge BSN
zelfde kind do en bso	Indicatie of voor hetzelfde kind over hetzelfde tijdvak zowel toeslag dagopvang als voor kinderopvang aangevraagd wordt.	DTM_KOT
Heeftbijstand (Hoogte)	Geeft aan of de aanvrager bijstand heeft of niet. * Deze gegevens zullen in Q1 2021 worden verwijderd.	FLG

Bijlage 5:

Control Matrix gebruikt voor kwaliteitscontrole van het model.

Voorbeeld van een deel van de control matrix:

Subproces	Controle		
Trainingscases CSV	CSV bestand met trainingcases is correct geüploadet met meest recente trainingcases. Denk aan: is datum beoordeeld juist ingevuld?		
initiaize	Bij toevoeging van een nieuwe maand is een nieuwe library aangemaakt in script A01.		
initiaize	Het bestand met trainingcases wordt correct ingeladen (script A02).		
initiaize	Bij toevoeging van een nieuwe maand is deze maand toegevoegd aan samenvoeging gescoorde cases in script A03.		
Create testcases	Bij toevoeging van een nieuwe maand is een nieuwe proces flow toegevoegd.		
Create testcases	Bij toevoeging van (een) nieuwe indicator(en) zijn deze voor alle historische runs op correcte wijze toegevoegd.		
Append	Bij toevoeging van een nieuwe maand is deze maand toegevoegd aan samenvoeging trainingcases in script Z01.		
Append	Alle overbodige kolommen worden gedropt in script Z01.		
Append	Lijst de controle in script Z02 volgen geen missings op indicatoren die niet missing mogen zijn.		
Miner - Data Partition	Alle aangeleverde trainingcases worden gebruikt voor het trainen of het valideren van het model. Vergelijk aantal records HT_training en KOT_training met aantal regels in Enterprise Miner.		
Miner - Scorecard	De door het model geselecteerde indicatoren, hun groepering en de hieraan gekoppelde scorecard points zijn plausibel.		
Miner - Scorecard	De samenstelling van het model vertoont geen onverklaarbare of implausibele verschillen met het model van de vorige run.		
#	Proces	Controle	Wie
1.1.	Brondata	Het aanleverformaat (postcodes, scheidingsteken, tekst, kwaliteits) is conform de in documentatie gegeven specificaties.	
1.2.	Brondata	Alle benodigde kolommen zijn beschikbaar.	
1.3.	Brondata	Het regelformaat is conform verwachting en lijft met eerdere dataleveringen.	
1.4.	Brondata	Aantal regels in bronbestand = aantal regels in SAS	
1.5.	Brondata	In SAS zijn dataformaten (denk aan: veldlengte, decimalen, datumformat) goed geladen	
2.1.	Halfproduct	Brondata is overgevoerd	
2.2.	Halfproduct	Er zijn na het draaien van het halfproduct geen errors/functiefouten/ontbrekende waarden in log.	
2.3.	Halfproduct	De output is conform verwachting: verslote kolommen zijn beschikbaar en hebben plausibele waarden, regelaantal is conform verwachting.	
2.4.	Halfproduct	Geen missings op numerieke indicatoren die niet missing mogen zijn	
2.5.	Halfproduct	Waar op kolomniveau na halfproduct bewaren is correct gegaan (in het algemeen: met voor vorige run van betreffende toeslagjaar)	
Algemeen			
3.1.	Productie	Voor dat producties run wordt gestart zijn alle historische halfproducten vervuld.	
3.2.	Productie	Voor dat producties run wordt gestart zijn alle actuele halfproducten naar de laptop gekopieerd	

Bijlage 6:

Toeslagen Toezicht Werkplan 2018

1. Inleiding

In het Handavingsplan Toeslagen 2018 is beschreven welke nieuwe activiteiten we dit jaar ontplooiën en welke activiteiten we verder ontwikkelen om de doelstellingen van onze handhaving te realiseren.

Handhaving bij Toeslagen bestaat uit dienstverlening, toezicht, fraudebestrijding, sancties en strafrecht. Met behulp van data-analyse zijn we steeds beter in staat om deze instrumenten op het juiste moment én in samenhang in te zetten. We richten ons op het vooraf invullen van gegevens in de aanvraag, het contactloos verwerken van wijzigingen door de grondslagen voor de burger te muteren zodat de burger dit niet meer hoeft te doen, en attenderen de burgers op een noodzakelijke mutatie wanneer wij die niet voor de burger kunnen doen. Waar nodig voeren we handmatig toezicht uit in de actualiteit. Onze fraudebestrijding is gericht op het vroegtijdig detecteren en voorkomen van fraude.

Met onze handavingsaanpak richten we ons op het vergroten van de toekenningszekerheid én op het vergroten van het vermogen en de bereidheid van burgers tot naleving van regelgeving. Hiermee beogen we terugvorderingen zowel in aantallen als in hoogte te beperken.

De mate waarin de voornemens te realiseren zijn, wordt onder meer beïnvloed door:

- de tussentijdse beoordeling van de uitkomsten van uitgevoerde handavingswerkzaamheden,
- de autonome ontwikkeling van het overige werkaanbod voor de handmatige verwerking, wat tot meer/minder capaciteit voor uitvoering van toezicht kan leiden,
- de beschikbaarheid van (IV-)resources,
- de uitkomst van afstemming met partijen buiten Toeslagen,
- de resultaten van proeftrajecten,
- de beschikbaarheid en de kwaliteit van gegevens uit externe bronnen,
- wijzigingen in wet- en regelgeving en nieuwe ontwikkelingen in de jurisprudentie,
- de voortgang van de implementatie van de topstructuur
- de besluitvorming met betrekking tot het wetsvoorstel Nieuw financieringsstelsel kinderopvang

2. De doelstellingen voor 2018 uit het Handavingsplan

Data-analyse

Het verkrijgen van een beter inzicht in de mogelijkheden van data- analyse voor de afnemers van informatieproducten en het bepalen aan welk inzicht behoefte bestaat.

De mogelijkheid om goede en relevante informatieproducten te kunnen ontwikkelen is volledig afhankelijk van de gegevenspositie van Toeslagen. Daarom zal in 2018 niet alleen worden

geïnvesteed op het verkrijgen van meer gegevens, maar ook op het verkrijgen van betere gegevens.

Dienstverlening

In 2018 onderzoeken we laaggeletterdheid en de gevolgen daarvan met betrekking tot toeslagen.

Daarnaast stellen we in 2018 voor de volgende onderwerpen vast hoe we onze communicatie met burgers vormgeven:

-Wel of geen wijzigingen doorgeven?

Duidelijke communicatie moet deze onduidelijkheid en daarmee vaak ook onzekerheid bij burgers voorkomen.

-eHerkenning

Aansluiting van dienstverleners op onze diensten en services, waardoor deze dienstverleners burgers goed kunnen helpen.

-Ontvangstbevestiging

In 2018 onderzoeken we welke wijze van het bevestigen van ontvangst het beste werkt om het aantal klant geïnitieerde contacten verder terug te brengen.

We investeren in 2018 ook in het dienstverleningsnetwerk. Zowel wat betreft de kwaliteit van de door de dienstverleners ingediende aanvragen en mutaties als in de ondersteuning aan en voorzieningen voor deze dienstverleners. Met Belastingdienst Particulieren werken we aan de verdere ontwikkeling van de website Kennisnetwerk voor dienstverleners en de mogelijkheid tot het inzetten van een mobiel team van Belastingdienst-medewerkers op locatie in specifieke gevallen. Ook de samenwerking met bewindvoerders worden geoptimaliseerd.

Toezicht

In 2018 herijken we het risicoclassificatie model om te bepalen welke indicatoren niet relevant genoeg (meer) zijn in het bepalen van het risico. Hierbij kijken we ook naar de betrouwbaarheid en actualiteit van de indicatoren. De herijking van het model kan leiden tot aanpassingen in de set indicatoren.

Daarnaast stellen we door middel van kort-cyclische pilots in lemniscaatteams vast voor welke grondslagen en welke (groepen) burgers we handmatig thematisch toezicht kunnen vervangen door uitbreiding van geautomatiseerd toezicht, contactloze verwerking proactieve interactie of een combinatie hiervan. Op basis van resultaat- en effectmetingen wordt voor iedere pilot vastgesteld of deze succesvol was. Succesvolle werkwijzen worden vervolgens geïmplementeerd in het reguliere proces.

Bij de inzet van attenderingen naar de burger toe, zoals attentiebrieven en SMS-berichten, experimenteren we met het inzetten van technieken uit de gedragswetenschap. Hierbij kan worden gedacht aan het vaststellen van verschillende persona's, het testen van verschillende nudging technieken en het onderzoeken van het beste moment voor het versturen van attentie- of herinneringsbrieven.

We herijken in 2018 onze samenwerkingsverbanden en daar waar nodig ontwikkelen we (aanvullend) beleid. In dit jaar doe we dit voor voor samenwerkingsverbanden ten behoeve van gegevensleveringen en ook voor de externe (overheids-)samenwerkingsverbanden.

In 2018 willen we het datafundament verder uitbreiden met generieke onderdelen, waardoor effectmetingen steeds gemakkelijker opgezet kunnen worden.

Naast het meten van het korte- en lange termijn effect van interventies, zetten we in 2018 de ontwikkeling voort van een indicatorenmodel, op basis waarvan we – op termijn – de effectiviteit van ons totale toezicht kunnen meten. Deze set aan indicatoren moet, samen met de reeds bestaande indicatoren, inzicht geven in de mate waarin het instrument toezicht als onderdeel van handhaving bijdraagt aan de strategische intenties van Toeslagen. In 2018 voeren we een dergelijke meting voor de eerste keer uit voor toeslagjaar 2016, waarna het model kan worden aangepast om eind 2018 nogmaals te gebruiken voor toeslag jaar 2017.

De acties voor TEOS:

Going concern/korte termijn:

- LSI projecten effectmeting
Het uitstralingseffect van LSI projecten wordt gemeten. Hierbij wordt gekeken naar de uitstraling van het project op de wijk/gemeente/netwerk (te denken aan familie) erom heen. Als de uitstralingseffecten in beeld zijn gebracht, kan er van daaruit toezicht worden ingericht of verbeterd. Er loopt een onderzoek naar de uitstralingseffecten. Vervolgactie is mogelijk aanpassen van het toezicht/dienstverlening.
- Bold move
In dit project is contact met gemeenten over derdengegevens van de sociale recherche die mogelijk interessant zijn voor fraude of misbruik meldingen. Dit project loopt al een tijdje. Het is lastig om met meerdere gemeenten te schakelen en dezelfde soort informatie terug te krijgen.
- Fenomenen
Jaarlijks zien we fenomenen zich voor doen. Deze fenomenen worden (soms samen met het Intensief toezicht team) benoemd en onderzocht vaak met behulp van buiten toezicht om te zien wat er daadwerkelijk speelt.

Midden lange termijn:

- Leercirkel uit projecten inrichten
Leercirkel inrichten naar aanleiding van resultaten uit projecten zodat dit verwerkt kan worden in en meegenomen bij toezichtacties of bijvoorbeeld het RCM. Dit geldt ook voor de bevindingen die volgen uit de projecten van de LAA.
- Meerdere huurtoeslagen op één adres
TEOS krijgt signalen dat er meerdere huurtoeslagen op één adres lopen, vaak terecht, behalve dat de gemeente het adres niet wil splitsen/aanpassen. TEOS is al in gesprek met gemeenten en wil verder verkennen welke mogelijkheden er zijn gemeenten te houden aan hun taak om BAG juist te vullen.
- Data huurcommissie uitspraken
De data van de huurcommissie uitspraken waarbij de huurprijs wijzigt zijn beschikbaar, alleen de ingangsdatum van de nieuwe huurprijs ontbreekt vaak. Vooral nog wordt er niks met deze data gedaan en de verwachting is dat de burger ook vaak nalaat de wijziging door te geven. Onderzoeken of en hoe we gebruik kunnen maken van deze data.
- Samenwerken/kennis ophalen bij gedrag beïnvloedende instanties
Idee is om de partijen waar we al mee samen werken te bevragen over of en hoe ze gedragswetenschappelijke kennis inzetten voor hun handhaving. Een mogelijkheid hierbij zou ook zijn om onderzoekers te vragen onze processen onder de loep te nemen. Vanuit het LSI is er veel aandacht voor gedrag beïnvloedende mogelijkheden. Hierop zouden we mee kunnen liften.
- Schuldenproblematiek

Onderzoeken of het hebben van schulden invloed heeft op het extra/oneigenlijk aanvragen van toeslagen. Zo ja, TEOS heeft contacten met gemeenten waarbij onderzoeken lopen naar schulden. Vervolgonderzoek hoe we vanuit toeslagen daarop in kunnen springen.

Lange termijn:

- Metatoezicht op kwaliteit van gegevens (MKB)

TEOS is vanuit MKB betrokken bij het metatoezicht op kwaliteit van gegevens. Dit is ook voor toeslagen relevant dus willen onderzoeken hoe we dit groter op kunnen zetten en ons voordeel mee kunnen doen.

Fraudebestrijding

In het 2018 wordt nog meer de nadruk gelegd op analyse van fraudepatronen en -fenomenen, de contra-informatie en nog meer gebruik gemaakt van data-analyse. Dit doen we onder andere door resultaten van onderzoeken naar individuele fraudesignalen en georganiseerde fenomenen te verzamelen, samen te voegen en te analyseren.

Daarnaast onderzoeken we in 2018 of het mogelijk is om naast de risicoclassificatie op aanvragen een risicomodel te ontwikkelen om fraudemeldingen te classificeren. Deze classificatie zorgt ervoor dat de meest risicovolle meldingen als eerste worden opgepakt. Op basis van de uitkomsten van de behandeling worden indicatoren ontwikkeld waarmee vergelijkbare fraude voorspeld kan worden.

Waar nodig bekijken we of de wettelijke mogelijkheden voor gegevensuitwisseling uitgebreid kunnen worden.

Met gebruikmaking van resultaten uit onderzoek willen we onze communicatie zodanig gaan inzetten dat we bewustwording creëren over de manier waarop wij ons toezicht uitvoeren en over de consequenties van misbruik met toeslagen.

In 2018 meten we ook het effect van sancties. We willen weten of het opleggen van sancties het gedrag van burgers positief of negatief beïnvloed. Met andere woorden, geven burgers die een verzuimboete opgelegd hebben gekregen in het vervolg wel netjes wijzigingen op tijd door, of juist niet.

3. Monitoren activiteiten 2018

Per activiteit wordt door de trekker van het betreffende handavingsonderdeel een factsheet opgesteld waarin de volgende gegevens worden vastgelegd:

- wie is de trekker
- wie zijn de verdere deelnemers
- beschrijving van de acties die nodig zijn
- stappenplan
- tijdpad

In het eerste overleg in maart 2018 waar de trekkers van de verschillende onderdelen bij elkaar komen worden de factsheets met elkaar besproken en zo nodig aangepast. Daarna worden ze voorgelegd aan de verantwoordelijke managers en vormen ze na goedkeuring het uitgangspunt voor de daadwerkelijke realisatie in 2018. Maandelijks komen de trekkers bij elkaar om elkaar te informeren over de stappen die zijn gerealiseerd voor de verschillende acties.

Ministerie van Financiën, Belastingdienst/Toeslagen - Toeslagen

De managers worden ook maandelijks geïnformeerd over de voortgang van de voorgenomen activiteiten. Een keer per vier maanden voegen we de factsheets bij de viermaandsrapportage.

