



Gemeente Rotterdam

Data Protection Impact Assessment (DPIA) Rapportage

**Ontwikkeling algoritme heronderzoeken
(Pilotfase)**

Van

[Redacted]

Datum

16 januari 2023



Inhoudsopgave

1	Inleiding	4
1.1	Versiebeheer	4
2	Noodzaak van de PIA	<u>65</u>
3	Omschrijving van de verwerking	
3.1	Doel	<u>86</u>
3.2	Relatie met register	<u>117</u>
3.3	Omschrijving van de verwerking	<u>117</u>
3.4	Uitgesloten van de scope	<u>1913</u>
3.5	Betrokken partijen	<u>2013</u>
3.6	Beschrijving van de persoonsgegevens die worden verwerkt	<u>2013</u>
4	Risico analyse privacy	<u>2116</u>
4.1	Rechtsgrond van de verwerking (artikel 6)	<u>2116</u>
4.2	Doelbinding (artikel 5)	<u>3019</u>
4.3	Juist, actueel en volledig	<u>3120</u>
4.4	Dataminimalisatie	<u>3220</u>
4.5	Bewaartermijn	<u>3622</u>
4.6	Verwijdering	<u>3722</u>
4.7	Locatie	<u>3722</u>
4.8	Anonimiseren en/of pseudonimiseren	<u>3823</u>
4.9	Informatieplicht	<u>3923</u>
4.10	Rechten van betrokkenen	<u>4024</u>
4.11	Training en instructie van medewerkers	<u>4226</u>
5	Risico Analyse beveiligingsmaatregelen	<u>4427</u>
5.1	Fysieke beveiliging	<u>4427</u>
5.2	Logische toegangsbeveiliging	<u>Fout! Bladwijzer niet gedefinieerd.28</u>
5.3	Logging	<u>Fout! Bladwijzer niet gedefinieerd.30</u>
5.4	Beheer	<u>Fout! Bladwijzer niet gedefinieerd.32</u>
5.5	Incident management	<u>Fout! Bladwijzer niet gedefinieerd.33</u>
5.6	Continuïteit	<u>Fout! Bladwijzer niet gedefinieerd.36</u>
5.7	Verwerker	<u>Fout! Bladwijzer niet gedefinieerd.37</u>
6	Overige risico's	<u>5138</u>



7	Overzicht maatregelen op basis van de risico's	5239
8	Planning maatregelen	5640
9	Akkoord	5741
10	Bijlage - Bronnen	5942
	A.1. Geïnterviewden	5942
	A.2. Documentatie	5942



1 Inleiding

De Gemeente Rotterdam is volgens de Algemene Verordening Gegevensbescherming (AVG)¹ verplicht een Data Protection Impact Assessment (DPIA) uit te voeren voor verwerkingen die waarschijnlijk een verhoogd risico opleveren voor de betrokkenen van wie de persoonsgegevens worden verwerkt.

Het doel van de DPIA is het verkrijgen van inzicht in risico's van de verwerking en in de maatregelen die genomen zijn of nog genomen zouden kunnen worden om deze risico's weg te nemen of te verminderen.

Een DPIA moet ten minste aan de volgende vereisten te voldoen²:

1. Beschrijving te geven van de beoogde verwerking en de doeleinden daarvan;
2. Beoordeling van de rechtmatigheid van de beoogde verwerking: beoordeling van de rechtsgrond, noodzaak en evenredigheid van de verwerking in relatie tot het beoogde doel;
3. Beschrijving en beoordeling van de risico's voor betrokkenen;
4. Beschrijving en beoordeling van de maatregelen om de risico's voor betrokkenen tegen te gaan of te beperken.

De proceseigenaar is verantwoordelijk voor het proces en is dit ook voor de initiatie en realisatie van de DPIA. De PO (Privacy Officer) en de DISO (Decentrale Information Security Officer) beoordelen de DPIA en voorzien hem van advies. De DPIA wordt daarna voor voorgelegd aan de Functionaris Gegevensbescherming (FG) van de gemeente Rotterdam. Het proces van de PIA is te vinden op RIO.

Indien de maatregelen niet of onvoldoende zorgen voor een inperking van de risico's voor betrokkene moet hierover voor aanvang van de verwerking overleg plaats te vinden met de Autoriteit Persoonsgegevens (AP). Dit wordt 'voorafgaande raadpleging' genoemd. De betreffende afdeling/ proceseigenaar die dit verzoek doet, stemt dit eerst af met de Privacy Officer van het cluster en de FG.

1.1 Versiebeheer

Versie	Datum	Door	Aanpassing	Opmerking
0.1 – 0.4	01-03-2022 – 12-08- 2022	[REDACTED] [REDACTED] [REDACTED] advies	Inhoudelijk en tekstueel	Adviezen [REDACTED] [REDACTED]
0.5	24-08-2022	[REDACTED] [REDACTED] [REDACTED]	Inhoudelijk en tekstueel	Adviezen [REDACTED] [REDACTED]

¹ Artikel 35 lid 1 AVG

² Artikel 35 lid 7 AVG



Versie	Datum	Door	Aanpassing	Opmerking
0.6	05-09-2022	[REDACTED] [REDACTED] [REDACTED]	Inhoudelijk en tekstueel	Adviezen [REDACTED] [REDACTED]
0.8	17-10-2022	[REDACTED]	Algoritme Expert	
0.9	18-10-2022	[REDACTED] [REDACTED]	Inhoudelijk en tekstueel	
1.1	24 -10 - 2022	[REDACTED] [REDACTED] [REDACTED]	Inhoudelijk en tekstueel	
1.2	27-10-2022	[REDACTED]	Inhoudelijk en tekstueel, hoofdstuk 6 en 7	Advisering
1.2	01-11-2022	[REDACTED] [REDACTED] [REDACTED] adviseur Functionaris Gegevensbe- scherming, [REDACTED] [REDACTED]	Inhoudelijk	Advisering
1.3	2-12-2022	[REDACTED] [REDACTED] [REDACTED]	Inhoudelijk en tekstueel	Adviezen FG en PO
1.4	12-01-2023	[REDACTED] [REDACTED]	Inhoudelijk en tekstueel	Adviezen [REDACTED] [REDACTED]



2 Noodzaak van de DPIA

De AVG bepaalt sinds 25 mei 2018 dat voor reeds bestaande verwerkingen met een waarschijnlijk hoog privacy risico, er elke 3 jaar opnieuw een DPIA moet worden uitgevoerd of eerder bij een inhoudelijke wijziging. Bij nieuwe verwerkingen moet de DPIA uitgevoerd zijn, voordat de verwerking start.

Om de rechtmatigheid van verstrekte uitkeringen te onderzoeken, maakt het cluster W&I onder meer gebruik van heronderzoeken. Jaarlijks wordt het aantal te verrichten heronderzoeken met de gemeenteraad afgestemd. Het heronderzoek heeft naast de controle op rechtmatigheid als doel dat samen met de uitkeringsgerechtigde wordt gekeken naar diens situatie. Daarbij wordt ook onderzocht of behoefte bestaat aan ondersteuning op bepaalde leefgebieden. De uitkeringsgerechtigden die voor een heronderzoek worden uitgenodigd, worden op verschillende wijze geselecteerd. Een van die methoden is selectie op basis van het risico-inschattingsmodel. In dit risicomodel wordt onder andere een algoritme ontwikkeld en toegepast. Het proces om van het vastgestelde aantal heronderzoeken uit het collegeakkoord te komen tot een lijst met actieve werkzoekenden met een risico inschatting op onrechtmatigheid, noemen we het hoofdproces "Toepassen algoritme voor heronderzoeken". Dit hoofdproces bevat twee fasen.

I. Pilotfase: Ontwikkelen van het algoritme voor heronderzoeken

In deze fase wordt het algoritme getraind en getest op historische data, om zo te komen tot een algoritme dat voldoende in staat is om de kans op (on)rechtmatigheid te voorspellen. In deze fase wordt een normenkader ontwikkeld. Op basis waarvan door de wethouder en proceseigenaar een weloverwogen besluit kan worden genomen om het algoritme wel of niet te gaan gebruiken.

II. Gebruiksfase: Het feitelijk in gebruik nemen van het algoritme heronderzoeken

In deze fase wordt het algoritme toegepast op de data van actuele uitkeringsgerechtigden, om zo te bepalen bij welke uitkeringsgerechtigden er (volgens het algoritme) sprake is van een verhoogde kans op onrechtmatigheid en die daarom voor een heronderzoek zouden moeten worden uitgenodigd

Een DPIA is noodzakelijk, omdat het proces/ verwerking 'Ontwikkeling algoritme heronderzoeken' waarschijnlijk een hoog risico meebrengt voor betrokkenen. Voor zowel de pilotfase (I) als de gebruiksfase (II) wordt separaat een DPIA opgeleverd.

Deze DPIA betreft de pilotfase. Een DPIA voor de gebruiksfase wordt na afsluiting van de ontwikkel-fase en voorafgaande aan de feitelijke ingebruikname van het algoritme opgeleverd. In hoofdstuk 3 is de scope van deze DPIA nader uitgewerkt.

De verwerking staat op de lijst van de Autoriteit Persoonsgegevens (AP) over DPIA-plichtige verwerkingen.

- ☐ Heimelijk onderzoek
- ☐ Zwarte lijsten
- ☒ Fraudebestrijding
- ☐ Creditscores
- ☒ Financiële situatie
- ☐ Genetische persoonsgegevens
- ☐ Gezondheidsgegevens
- ☐ Samenwerkingsverbanden

Artikel 5.2 lid 1 Woo



- ☐ Cameratoezicht
- ☐ Flexibel cameratoezicht
- ☐ Controle werknemers
- ☐ Locatiegegevens
- ☐ Communicatiegegevens
- ☐ Internet of things
- ☒ Profileren

(Het algoritme is een risicoclassificatiemodel. Het gebruik van het risico-classificatiemodel is een vorm van profileren, omdat deze verwerking mogelijk voldoet aan alle drie de voorwaarden genoemd in artikel 4, aanhef en onder 4, van de AVG)

- ☐ Observatie en beïnvloeding van gedrag
- ☐ Grootschalige verwerkingen en/of stelselmatige monitoring van biometrische gegevens met als doel een natuurlijk persoon te identificeren

Daarnaast voldoet de verwerking aan twee of meer criteria van de European Data Protection Board (EDPB).

- ☒ Beoordelen van mensen op basis van persoonskenmerken
- ☐ Geautomatiseerde beslissingen
- ☐ Stelselmatige en grootschalige monitoring
- ☒ Gevoelige gegevens
- ☒ Grootschalige gegevensverwerkingen
- ☒ Gekoppelde databases
- ☒ Gegevens over kwetsbare personen (o.a. verplichting tot het aanleveren van inkomensgegevens, werkzoekenden)
- ☒ Gebruik van nieuwe technologieën

Daarnaast is middels een Algorithm Risk Assessment (ARA) vastgesteld dat het te ontwikkelen algoritme een hoog risicovol model betreft. De werking ervan is moeilijk volledig én beknopt uit te leggen aan de gemiddelde Rotterdammer. Daarbij is er een risico dat de algoritmetoepassing uitmondt in een onrechtvaardige handeling van burgers. Het derde aandachtspunt in de ARA is niet aanwezig. Er is bij de uitvoering van heronderzoeken (ongeacht of die op basis van output van het model worden uitgevoerd) altijd sprake van een menselijke handeling; de uitkomst van het algoritme heeft geen automatische rechtsgevolgen voor de burger.

Aangezien uit de ARA is gebleken dat het om een hoog risicomodel gaat, is volgens de concernbrede Governance betreffende Algoritmen 3 IAMA (Impact Assessment Mensenrechten en Algoritmes) workshops gehouden op de oude versie van het algoritme. De definitieve IAMA-rapportage is nog in wording. Zowel in de ARA als in de IAMA is aandacht besteed aan de doelstellingen achter de ontwikkeling van het algoritme. Een IAMA op het algoritme moet nog uitgevoerd worden. Zie hierover meer in paragraaf 3.



3 Doel

Het doel van dit hoofdstuk is om lezers van deze rapportage welke geen voorkennis hebben van het proces te laten begrijpen wat er exact met de persoonsgegevens gebeurt.

3.1 Doel verwerking

Cluster Werk & Inkomen (W&I) van de gemeente Rotterdam voert jaarlijks, op verzoek van de gemeenteraad een vastgesteld aantal heronderzoeken uit. Deze onderzoeken hebben tot hoofddoel om de rechtmatigheid van de verstrekte uitkering op grond van de Participatiewet (PW) te kunnen beoordelen. Het heronderzoek is ook het moment om in contact te blijven met de burger en door te praten over diens situatie. Mocht in de gesprekken in het kader van een heronderzoek naar voren komen dat behoefte is aan aanvullende dienstverlening, dan kan hier aandacht aan worden besteed, bijvoorbeeld in de vorm van een doorverwijzing naar een andere instantie of onderdeel van de gemeente.

Juridische grondslag

Gemeenten zijn de uitvoerders van de Participatiewet. Voor een eerlijke en op solidariteit gebaseerde uitvoering van de Participatiewet is het voor gemeenten van belang om onrechtmatig gebruik van bijstandsuitkeringen op te sporen. Een ten onrechte uitgekeerde uitkering kan worden teruggevorderd en eventueel kan een boete worden opgelegd. Het tijdig signaleren van onjuistheden in de uitkering, is daarom ook in het belang van de burger die een uitkering ontvangt.

Op basis van artikel 53a Participatiewet is het aan gemeenten om te bepalen hoe het onderzoek naar de rechtmatige verstrekking wordt verricht en welke gegevens en bewijsstukken daarbij moeten worden verstrekt.

De Centrale Raad van Beroep heeft in een aantal uitspraken overwogen dat met de uitvoering van de bijstandswetgeving aanzienlijke publieke middelen gemoeid zijn. Controle en handhaving zijn nodig, aldus de raad, voor een juiste besteding van die middelen, het voorkomen van misbruik en het in stand houden van het maatschappelijk draagvlak. Om redenen van efficiëntie en kostenbeheersing mogen gemeenten gerichte vormen van onderzoek daarom beperken tot bepaalde groepen uitkeringsgerechtigden bij wie een grotere kans bestaat dat de controle bijdraagt aan de juiste uitvoering van de bijstandswetgeving. Daarbij mogen zij onder voorwaarden ook risicoprofielen opstellen en data koppelen. De uitkomst daarvan mag echter niet discriminerend zijn. De grondslag hiervan is te vinden in art. 53a lid 1 en 6 PW. De grenzen aan het opstellen en hanteren van risico-profielen volgen met name uit jurisprudentie. In "[Gelijk recht doen Deelrapport domein Sociale zekerheid \(staten-generaal.nl\)](#)" is hier aandacht aan besteed. Mr. Drs. Marije Otte heeft in "Rechtmatigheidsonderzoek op basis van risicoprofielen" hier een uitwerking aan gegeven (<http://szweb.adlibhosting.com/De-tails/.%5CPDF%5C200086901.pdf>)

Met name gelet op art. 53a lid 1 en 6 PW, heeft de gemeente de bevoegdheid tot het doen van onderzoek naar de rechtmatigheid van de op grond van de Participatiewet verstrekte uitkering. Het verwerken van persoonsgegevens met dat doel heeft als wettelijke grondslag artikel 6, lid 1, sub e AVG en luidt:

Artikel 5.2 lid 1 Woo



‘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’.

Risico-inschattingmodel

Gezien het aantal van rond 33.000 uitkeringsgerechtigden is het praktisch niet mogelijk om de gehele doelgroep met enige regelmaat over de rechtmatigheid van de uitkering op het stadhuis te spreken. Een selectie uit deze doelgroep is daarom noodzakelijk. Momenteel worden de uitkeringsgerechtigden via een a-selectie steekproef geselecteerd voor een uitnodiging voor een heronderzoek.

Een methodiek die in voorgaande jaren is ingezet en nu opnieuw in ontwikkeling is, is selectie op basis van het risico-inschattingmodel. Het risico-inschattingmodel werkt met een algoritme, dat uitkeringsgerechtigden op basis van een zogenaamd risico-inschattinggetal inschat. Door resultaten van rechtmatigheidsonderzoeken uit het verleden te analyseren, schat het algoritme de kans op rechtmatige of onrechtmatig verstrekking van bijstand in voor de huidige uitkeringsgerechtigden. Bij de ontwikkeling van het nieuwe algoritme dat naar verwachting in 2023 wordt ingezet, wordt alleen nog gebruik gemaakt van de resultaten van eerdere rechtmatigheidsonderzoeken bij burgers die a-select zijn uitgenodigd. Dit voorkomt doorwerking van eventuele bias in eerder gehanteerde onderzoeksmethoden in het nieuwe algoritme.

Het risico-inschattingmodel helpt bij het aanbrengen van focus voor selectie voor heronderzoeken. Het risico-inschattingmodel selecteert de uitkeringsgerechtigden waarbij het risico op onrechtmatigheid het grootste is.

Het gaat hierbij om een kansbepaling. Het risico-inschattingmodel was in eerdere jaren met een bepaalde mate van nauwkeurigheid in staat om te voorspellen bij welke uitkeringsgerechtigden een kans op onrechtmatigheid aanwezig was. In die zin is gebleken dat het algoritme ons helpt om kostbare tijd effectiever in te zetten en daarmee ook tijdig problemen bij een aantal werkzoekenden te voorkomen. Een heronderzoek is voor de burger vrij impactvol, het risico-inschattingmodel draagt er ook aan bij dat vooral mensen aan een heronderzoek worden onderworpen als er een goede aanleiding voor is.

De output van het algoritme wordt alleen gebruikt om te bepalen aan welke uitkeringsgerechtigden het algoritme het hoogste risico-inschattinggetal verbindt (en waarbij volgens het algoritme dus kans op onrechtmatigheid het hoogst is). Vervolgens worden de uitkeringsgerechtigden met het hoogste risicogetal voor een heronderzoek uitgenodigd.

Het algoritme is alleen bepalend voor de vraag wie voor een heronderzoek zou moeten worden uitgenodigd. Het algoritme neemt niet zelfstandig/ geautomatiseerd besluiten over het recht op of de hoogte van de uitkering.

Het inhoudelijke onderzoek naar de rechtmatige verstrekking van bijstand wordt altijd door een medewerker verricht. Daarnaast wordt het onderzoek voor alle geselecteerde uitkeringsgerechtigden op gelijke wijze uitgevoerd, ongeacht of de uitkeringsgerechtigde voor een heronderzoek is uitgenodigd op basis van uitkomsten van het algoritme of een andere selectiemethode, zoals een aselecte steekproef. De ambtenaar die het heronderzoek uitvoert heeft geen wetenschap van hoe de wijze waarop de uitkeringsgerechtigde is geselecteerd voor een heronderzoek.

De doelstellingen van het algoritme zijn (samengevat):



- Efficiënte inzet van middelen
- Effectieve inzet van capaciteit
- Waarborging rechtmatigheid
- Beperking van het aantal uitkeringsgerechtigden dat rechtmatig een uitkering ontvangt en wordt geconfronteerd met relatief zware onderzoeken.
- Draagvlak socialezekerheidsstelsel overeind houden: zorgen dat mensen de juiste uitkering krijgen, waar ze ook recht op hebben. En dat mensen die er geen recht op hebben, het niet krijgen.

Ontwikkeling algoritme 2023

Sinds 2018 wordt er gebruik gemaakt van een risico-inschattingsmodel. Sindsdien zijn meerdere versies van het model gebruikt. De laatste versie van het risico-inschattingsmodel is ontwikkeld in 2020 en de onderzoeken op basis van de output van dit model zijn in de loop van 2021 afgerond.

Voor het uiteindelijke algoritme wordt steeds die techniek geselecteerd die de meest betrouwbare voorspelling van de kans op (on)rechtmatigheid oplevert. Er is dus telkens sprake van een volledig nieuw model, waarbij de onderliggende (modelleer)techniek telkens kan verschillen. Daarbij zal de in gebruik name van elk niet ontwikkeld model gepaard gaan met een privacy toetsing van de gebruikte data in het model. Het algoritme wordt ontwikkeld door gespecialiseerde data-scientists van afdeling Onderzoek en Business Intelligence, Cluster BCO van de gemeente Rotterdam (OBI).

Om door te gaan met het risico-inschattingsmodel, wordt nu een nieuw algoritme ontwikkeld en getraind voor gebruik in 2023.

Voorgeschiedenis: periode tot heden

Begin 2021 is meer aandacht gekomen voor de ethische aspecten van het werken met een algoritme. In februari 2021 publiceerde Concern Auditing van de gemeente Rotterdam op verzoek van de directeur Uitvoering van W&I het 'Advies Analytics Uitkeringsonrechtmatigheid'. In april 2021 volgde het rapport 'Gekleurde technologie' van de Rekenkamer Rotterdam.

Beide rapporten geven aan dat de transparantie rondom (de totstandkoming van) het algoritme moet worden vergroot en duidelijk inzichtelijk moet worden gemaakt hoe eventuele vooringenomenheid (bias) in het model wordt voorkomen. Weliswaar trekt het algoritme geen conclusies over de rechtmatige of onrechtmatige verstrekking van bijstand, maar wanneer er gebruik wordt gemaakt van variabelen en features die kunnen leiden tot bias, kan wel degelijk sprake zijn van ongelijke behandeling van uitkeringsgerechtigden. Bepaalde uitkeringsgerechtigden hebben immers een grotere kans om voor een heronderzoek te worden uitgenodigd. Met name de Rekenkamer Rotterdam heeft aandacht gevraagd voor het bestaan van zogenaamde 'proxy-variabelen' (variabelen die indirect discriminatoir kunnen werken) in het model.

Eind 2021 is in opdracht van de gemeenteraad gestart met voorbereiden van een nieuw risico-inschattingsmodel, waarvan het de bedoeling is dat het in 2023 in gebruik kan worden genomen. Cruciale doelstellingen bij de ontwikkeling van het risico-inschattingsmodel zijn:

1. Een transparante ontwikkeling van het algoritme, waarbij potentiële vooringenomenheden - en daarmee het risico van discriminatie - worden voorkomen.

Artikel 5.2 lid 1 Woo



2. Ontwikkeling van een uitlegbaar risico-inschattingsmodel, onder meer door het gebruik van glass box technieken. Hiermee kan ook op het niveau van een individuele burger beter worden uitgelegd hoe diens score (het risico-inschattingsgetal) tot stand is gekomen.
3. Transparant communiceren over doel en werking van het algoritme (onder meer door opname in het algoritmeregister) en het gebruik ervan.

Bij de ontwikkeling van het algoritme wordt, naast ethische normen vanuit de IAMA-assessment en Privacy wet- en regelgeving, gebruik gemaakt van de door de Europese Commissie in april 2019 gepubliceerde 'ETHICS GUIDELINES FOR TRUSTWORTHY AI' (de Ethische Richtsnoeren voor Betrouwbare Algoritmen): <https://www.aepd.es/sites/default/files/2019-12/ai-ethics-guidelines.pdf>.

Het algoritme wordt getraind op basis van resultaten van eerder uitgevoerde rechtmatigheidsonderzoeken die middels de a-selectie methode zijn geselecteerd³.

3.2 Relatie met register, Informatiebeheerplan en applicaties

Het algoritme wordt ingezet voor het proces "Uitvoeren heronderzoeken", nummer P1389 in het informatiebeheerplan en nummer W1060 in het Register van verwerkingen.

Het algoritme zelf en alle documentatie betreffende de ontwikkeling van het algoritme betreft het proces "data science projecten", nummer P3693 in het informatiebeheerplan en BCO021 in het Register van verwerkingen.

Het algoritme wordt op een projectmatige wijze en onder verantwoordelijkheid van portefeuille Inkomen van het cluster Werk en Inkomen ontwikkeld. De projectdocumentatie betreft het proces "uitvoeren opgave", nummer P1459 van het Informatiebeheerplan. Het systeem DIVA wordt gebruikt om de betreffende projectdocumentatie te archiveren.

Het algoritme wordt ontwikkeld op het Azure DAP'R platform. Het Datalake en concern Datawarehouse is voor toepassing van de Baseline Informatiebeveiliging Overheid geclassificeerd als BBN2+.

Het algoritme maakt gebruik van de data uit de systemen Socrates en RMW/Raak. Zie verder paragraaf 3.3.

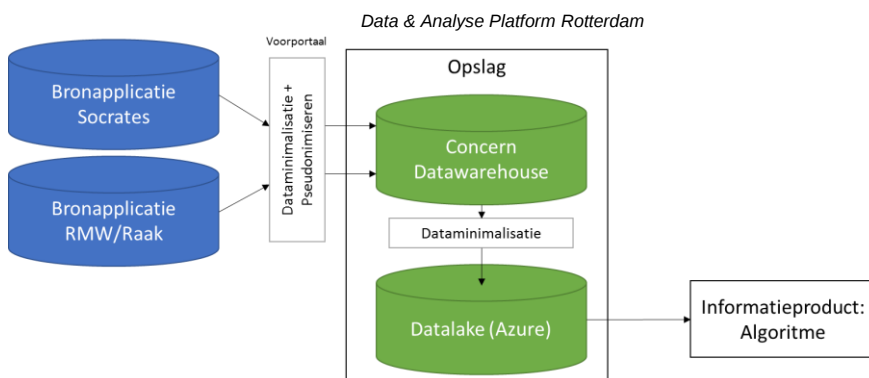
3.3 Omschrijving van de verwerking

Zoals beschreven in paragraaf 3.1, wordt met behulp van het model de kans op onrechtmatigheid voorspeld. Dat gebeurt op basis van statistische analyses van historische data, die zijn vastgelegd in het kader van de uitvoering van de Participatiewet. Het algoritme maakt gebruik van historische data betreffende de afgeronde heronderzoeken die de afgelopen 5 jaar zijn uitgevoerd. Hiervoor wordt de peildatum gebruikt van 5 jaar voor de peildatum van de ingebruikname van het algoritme die middels de a-selectie methode zijn geselecteerd. Deze gegevens worden vastgelegd in de bronsystemen: Socrates en RMW/Raak. Socrates is de uitkeringsadministratie, waarin rechtmatigheid van uitkering wordt vastgesteld. RMW/Raak is het systeem waarin de gegevens over de re-integratie van de uitkeringsgerechtigde (doelmatigheid) worden vastgelegd.

³ Zie voetnoot 3.



Het algoritme wordt ontwikkeld op de omgeving Datalake (Azure) DAP'R. DAP'R (Data & Analyse Platform Rotterdam) is de centrale faciliteit voor datagedreven werken en bestaat uit de onderdelen: concern Datawarehouse, Datalake (Azure), Front-end tooling (Cognos, Power BI, SAS Viya) en Advanced analytics m.b.v. Azure Databricks. Er vindt gegevenstransport plaats van de bronsystemen Socrates en RMW/Raak naar DAP'R, specifiek eerst naar concern Datawarehouse en daarna naar Datalake (Azure). In de toekomst (2023) is er een direct gegevenstransport vanuit de bronsystemen Socrates en RMW/RAAK naar Datalake-Azure.



Procesbeschrijving "Ontwikkelen algoritme voor heronderzoeken"

Voor de ontwikkeling van het algoritme wordt als leidraad de procesbeschrijving "Ontwikkelen algoritme voor heronderzoeken" gevolgd. Deze procesbeschrijving is door directeur Inkomen vastgesteld (d.d. 25.02.2022). In de bijlage is deze procesbeschrijving toegevoegd.

De data-scientists ontwikkelen in nauwe samenwerking met het projectteam het algoritme voor heronderzoeken.

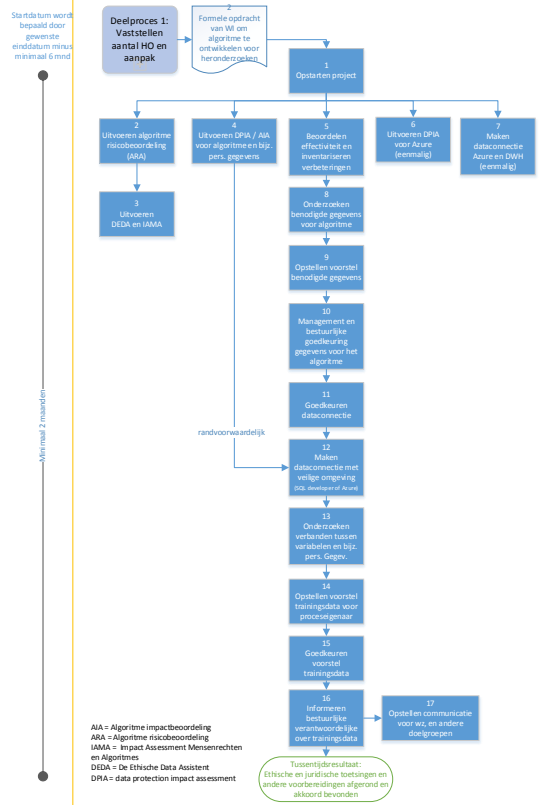
Het proces 'Ontwikkelen algoritme voor heronderzoeken' bestaat uit 2 subfases. In subfase 1 worden de zaken uitgevoerd die randvoorwaardelijk zijn voor het ontwikkelen van een algoritme. In de 2e subfase wordt het algoritme daadwerkelijk getraind c.q. ontwikkeld.

Subfase 1 start met een formele opdracht van de opdrachtgever (Directeur uitvoering, Cluster W&I) aan Cluster OBI om een algoritme te ontwikkelen voor heronderzoeken en hiervoor data-scientists beschikbaar te stellen. Subfase 1 eindigt als de ethische en juridische toetsen en andere voorbereidingen, zoals de benodigde dataset zijn afgerond en zijn goedgekeurd door de proceseigenaar.



Proces Ontwikkelen algoritme voor heronderzoeken

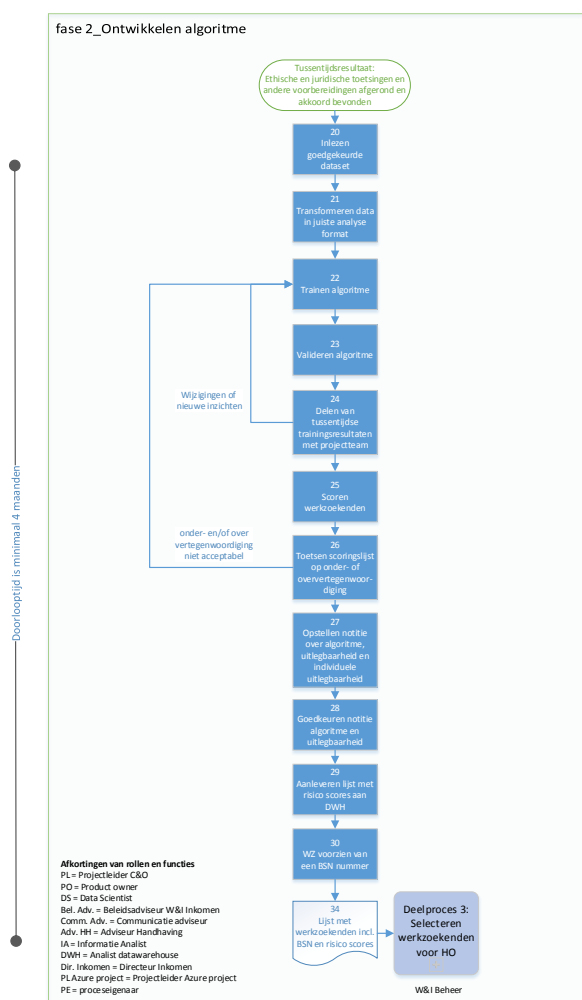
fase 1 Inregelen randvoorwaardelijke zaken





Subfase 2 start als de randvoorwaardelijke zaken in subfase 1 zijn afgerond en goedgekeurd. Subfase 2 start met het inlezen van de dataset die is goedgekeurd door de proceseigenaar en eindigt met de validatie van het algoritme (stap 23) en het delen van de resultaten met het projectteam (stap 24).

Hierna wordt de pilotfase afgerond en wordt er een keuze gemaakt over het al dan niet inzetten van het algoritme. Het feitelijk inzetten van het algoritme door deze toe te passen op actuele uitkeringsgerechtigden (stap 25 en verder) valt onder het DPIA voor de gebruiksfase.





Vanuit het perspectief van de verwerking van persoonsgegevens zijn de volgende activiteiten in subfase 1 van het proces relevant.

Stap 7 Maken dataconnectie Azure en concern Data Warehouse

Het algoritme wordt ontwikkeld op het Azure DAP'R platform. Deze stap betreft het aanleggen van de infrastructuur met een werkende dataconnectie tussen het concern Data Warehouse en Azure. Deze infrastructuur is de basis waarop het algoritme wordt ontwikkeld.

Omdat de dataset versleutelde BSN's bevat, wordt de dataset afgezonderd opgeslagen met een eigen autorisatie voor toegang tot gebruik.

Stap 12 Maken dataconnectie veilig omgeving (Azure)

Op de in stap 7 aangelegde infrastructuur met een werkende, beveiligde dataconnectie tussen concern Datawarehouse en Azure, wordt vervolgens van deze infrastructuur gebruik gemaakt. Op dagelijkse basis wordt de actuele data uit Socrates en RMW/Raak ingelezen ten behoeve van de ontwikkeling van het algoritme.

Het verschil tussen stap 7 en stap 12 van subfase 1 in de procesbeschrijving is:

- Stap 7, een eenmalige dataconnectie ten behoeve van het aanleggen van de benodigde infrastructuur.
- Stap 12, het dagelijks inlezen van de actuele data ten behoeve van de inzet van het algoritme.

Stap 8 Onderzoeken benodigde gegevens voor algoritme

Het algoritme maakt gebruik van historische data betreffende de afgeronde heronderzoeken die de afgelopen 5 jaar zijn uitgevoerd. Hiervoor wordt de peildatum gebruikt van 5 jaar voor de peildatum van de ingebruikname van het algoritme die middels de a-selecte methode zijn geselecteerd. Zoals eerder uiteen is gezet⁴, worden voor het te ontwikkelen algoritme alleen de data uit heronderzoeken gebruikt waarbij sprake was van a-select uitgenodigde uitkeringsgerechtigden. Door het algoritme te baseren op de resultaten van eerdere a-select uitgevoerde heronderzoeken, heeft het algoritme per definitie een hogere effectiviteit dan de a-selecte methode.

Effectiviteit:

Tijdens de pilotfase kan alleen iets worden geroepen over de te verwachten effectiviteit van het getrainde algoritme. Hiertoe wordt het algoritme getest op een afgescheiden deel van de historische data (de 'testdata'). Op deze wijze kan worden getoetst in welke mate het algoritme correct gevallen weet aan te wijzen waarin er feitelijk sprake is geweest van onrechtmatigheid. Hierbij wordt o.a. gekeken naar hitrate en accuracy. Hoe groot de effectiviteit van het algoritme daadwerkelijk is, is pas na te meten na afronding van heronderzoeken op basis van het algoritme. Dan kunnen we de resultaten van de selectiemethoden a-select, thema-onderzoek en algoritme met elkaar vergelijken. In de jaarrapportage Handhaving 2023 vindt deze effectiviteitsmeting plaats.

Wat we vanuit voorgaande jaren over de effectiviteit van het algoritme weten is dat dit algoritme per definitie effectiever is dan de a-selecte methode. Het algoritme wordt namelijk getraind met de data uit de a-selecte onderzoeken die in de afgelopen 5 jaar zijn uitgevoerd.

⁴ Zie voetnoot 3.



Uit de onderzoeksresultaten over de jaren 2018 tot en met 2021 blijkt dat het verrichten van heronderzoek op basis van het risico-inschattingsmodel, effectiever is dan het verrichten van heronderzoeken bij a-select gekozen uitkeringsgerechtigden of op basis van een vooraf bepaald thema. Dit uit zich in een hoger percentage beëindigde uitkeringen en terugvorderingen bij onderzoeken die in deze jaren zijn uitgevoerd op basis van het risico-inschattingsmodel. Dit betekent dat het model een goede voorspellende waarde heeft. Bij de uitkeringsgerechtigden die we voor een rechtmatigheidsonderzoek uitnodigen op basis van dit model, hebben we dus een grotere kans om onrechtmatigheden op te sporen en tijdig te verhelpen.

Voor de cijfers van 2021 wordt verwezen naar de in bijlage 7 toegevoegde 'rapportage Handhaving cluster Werk en inkomen 2021'. Voor de cijfers van de jaren 2018 t/m 2020 wordt verwezen vraag 4 in de bijlage 6 toegevoegde 'Beantwoording van de schriftelijke vragen van het raadslid A.J. van Zevenbergen (SP) over 'Gemeentelijke algoritmes onder de loep' d.d. 14-12-2021.

Gebruikte data en voorkomen van vooringenomenheden:

Niet alle gegevens uit de gebruikte administraties in de systemen Socrates en RMW/Raak worden gebruikt. Incompleteheid, identieke invulling of samenhang tussen variabelen kan leiden tot het buiten beschouwing laten van bepaalde variabelen. Dit is het resultaat van de datapreparatie die door het team Advanced Analytics wordt uitgevoerd. Mogelijke direct of indirect (proxy) discriminatoire variabelen worden niet opgenomen in het algoritme. Hierbij wordt uitgegaan van de gronden op basis waarvan volgens het College van de Rechten van de Mens niet gediscrimineerd mag worden. Deze gronden zijn: godsdienst, levensovertuiging, politieke gezindheid, ras, geslacht, nationaliteit, seksuele gerichtheid, burgerlijke staat, handicap of chronische ziekte en leeftijd. Om vast te stellen of er een verband bestaat tussen de variabelen uit de systemen Socrates en RMW/Raak en discriminatoire variabelen, wordt door het team Advanced Analytics correlatieonderzoeken uitgevoerd (zie stap 13). De uitkomsten van de correlatieonderzoeken leiden tot een voorstel om bepaalde variabelen uit de systemen Socrates en RMW/Raak in het algoritme worden meegenomen of juist bewust uit te sluiten. Het gaat dan om variabelen die (in)direct correleren met discriminatoire gronden en om deze reden niet in het algoritme worden opgenomen. Uiteindelijk besluit de gemeenteraad over dit voorstel. Dit is stap 20 in subfase 1 van het ontwikkelproces.

Stap 13 Onderzoeken verbanden tussen variabelen en bijzondere persoonsgegevens

Om inzicht te krijgen in mogelijke (in)direct discriminerende werking van variabelen die potentieel in het algoritme worden meegenomen worden correlatieonderzoeken uitgevoerd. De vraag daarbij is steeds of er een verband zit tussen:

- de variabelen uit de systemen Socrates en RMW/Raak en
- de hierboven genoemde gronden op basis waarvan niet gediscrimineerd mag worden.

Met uitzondering van ras/afkomst, worden deze correlatieonderzoeken uitgevoerd met gebruikmaking van de data uit RMW/Raak en Socrates. Aangezien gegevens over ras/afkomst niet direct of indirect wordt geregistreerd in deze systemen, is het gebruik van de data uit het systeem BRP (Basis Registratie Persoonsgegevens) voor een correlatieonderzoek betreffende ras/afkomst een vereiste. Dit is de enige wijze waarop op objectieve wijze een eventuele correlatie met ras/afkomst kan worden aangetoond.



De data uit de BRP worden enkel gebruikt voor de uitvoering van het correlatie-onderzoek (validatie) met betrekking tot ras/afkomst. Met als doel om mogelijke (indirecte) discriminatie van het algoritme uit te sluiten. De BRP-data worden dus niet gebruikt in het model of om nieuwe modellen op te trainen.

In samenspraak met data experts van cluster Dienstverlening wordt nog nader bepaald welke exacte extractiekolommen vanuit de BRP voor dit doel beschikbaar zouden moeten worden gesteld. Het team Advance Analytics heeft aangegeven voor deze correlatie-check behoefte te hebben aan de volgende extractiekolommen vanuit de BRP: herkomstland, geboorteland ouders en historische nationaliteit.

Op 18 juni 2021 heeft Juridische Diensten een positief advies gegeven over het gebruik van de BRP-data voor de uitvoering van het correlatieonderzoek. Citaat advies Juridische Diensten: *“De wettelijke grondslag voor het gebruik van gegevens over nationaliteit als controlemiddel bij het algoritme moet worden gezocht in artikel 53a, van de Participatiewet waarin de algemene bevoegdheid voor het college voor het verrichten van (her)onderzoeken naar de rechtmatigheid van verleende bijstandsuitkeringen is geregeld.”* “Het gebruik van gegevens over nationaliteit en afkomst als controlemiddel bij dit algoritme lijkt vooralsnog geen juridische onvolkomenheden op te leveren, mits volstrekt inzichtelijk wordt gemaakt dat deze gegevens niet worden gebruikt om het algoritme te ‘voeden’.” In de bijlage is het advies van Juridische Diensten toegevoegd.

Vooraf, gedurende en na afronding van het correlatieonderzoek op ras/afkomst zijn duidelijke beheermaatregelen getroffen om deze gegevens te beschermen (zie uitwerking bij norm 15):

- De BRP-data worden ter beschikking gesteld en op twee duidelijk vooraf aangegeven momenten gebruikt puur ter validatie, dus niet om te verwerken in een model
- De BRP-data worden enkel vrijgegeven aan en gebruikt door één data-scientist van afdeling OBI. Deze data-scientist is in loondienst van de gemeente Rotterdam.
- De BRP-data worden na gebruik volgens de kaders en richtlijnen van de AVG en Archiefwet vernietigd.
- De beschikbaarstelling van de BRP-data wordt indien nodig in een afzonderlijke gebruikers-overeenkomst vastgelegd.

Gedurende de correlatieonderzoeken wordt per variabele/feature in Socrates en RMW/Raak door de leden van de projectgroep:

1. Bekeken of er een statistische correlatie is tussen de variabele/feature en een correlatievariabele. Hierbij wordt de regel gehanteerd dat alle variabelen/features met een gemiddelde of hoge correlatie niet in het algoritme worden meegenomen. De R-waarde van 0,3 en hoger is de grens die hierbij wordt gehanteerd.
2. Een nadere statistische analyse uitgevoerd per variabele/feature met een R-waarde tussen 0,25 en 0,3. Met als doel om variabelen/features in het grensgebied diepgaander te analyseren op eventuele discriminerende werking.
3. Op basis van de naamgeving en onderliggende beschrijving van een variabele/feature bekeken of er een verband met een correlatievariabele lijkt te bestaan.

Nadat de correlatieonderzoeken zijn afgerond, wordt de lijst met variabelen/features die in het algoritme worden meegenomen voor besluitvorming aan de wethouder voorgelegd.



Vanuit het perspectief van de verwerking van persoonsgegevens zijn de volgende activiteiten in fase 2 van het proces relevant.

Stap 22. Trainen algoritme

Het algoritme wordt ontwikkeld op het Azure DAP'R platform. Het algoritme maakt gebruik van de data uit de systemen Socrates en RMW/Raak. Door twee data-scientists van afdeling OBI van de gemeente Rotterdam wordt het algoritme getraind. Met een iteratief proces van algoritme training wordt de kans op rechtmatige of onrechtmatige verstrekking van bijstand voorspeld. Tijdens een groot deel van het proces zijn de gegevens van uitkeringsgerechtigden gepseudonimiseerd.

Stap 23. Valideren algoritme

Valideren van het algoritme houdt in dat er op meerdere vooraf bepaalde meetwaarden getoetst wordt hoe bruikbaar het algoritme is.

Meetwaarden zijn bijvoorbeeld:

- Technische meetwaarden
- Uitlegbaarheid: Is het algoritme uitlegbaar?

Overige meetwaarde om het algoritme te valideren worden in het ontwikkelproject in gezamenlijkheid bepaald. Dat betekent in samenspraak met de data-analist, beleidsadviseur, proceseigenaar en Privacy Officer van cluster Werk & Inkomen en de data-scientists van team Advanced Analytics. Het algoritme zal meerdere validatieslagen doorlopen, voordat het 'goed' genoeg is om uiteindelijk voor te leggen aan proceseigenaar ter goedkeuring.

Stap 24. Delen van tussentijdse trainingsresultaten met projectteam

Met de projectgroep wordt gedurende het trainingstraject, meerdere keren overlegd of de ontwikkeling van het algoritme op een juist spoor is of dat er nog zaken moeten veranderen. Deze informatie wordt dan weer meegenomen om het algoritme te verbeteren.

Onderwerpen die besproken worden in dit overleg zijn:

- Prestaties modellen;
- Uitlegbaarheid algoritme globaal;
- Uitlegbaarheid algoritme op het niveau van de individu (waarom deze score per individu).

Stap 25. Scoren werkzoekenden

Het vastgestelde algoritme wordt gebruikt om de risico onrechtmatigheid te beoordelen op de huidige groep werkzoekenden. Het scoren betekent dat alle werkzoekenden een inschatting krijgen van het nieuwe model. Het scoren op de totale groep werkzoekenden geeft een beeld hoe de verdeling van de scores is over de populatie. Maar ook geeft dit een beeld welke groep werkzoekenden mogelijk uitgenodigd worden voor een heronderzoek. Het resultaat van deze stap is een lijst met geanonimiseerde werkzoekenden met een risicoscore op onrechtmatigheid.

Stap 26. Toetsen scoringslijst op onder- of oververtegenwoordiging (discriminatie toetsing)

De groep die in aanmerking komt voor heronderzoek wordt hier getoetst aan de hand van bijzondere persoonsgegevens op over- of ondervertegenwoordiging. Het is niet wenselijk meer of juist minder mensen met bepaalde bijzondere karakteristieken te selecteren dan we zouden mogen verwachten in deze groep. Deze toets is een extra toets om te bepalen of het model zonder gebruik te maken van discriminerende variabelen tot een vooringenomen resultaat komt. Als de proceseigenaar de

Artikel 5.2 lid 1 Woo



Artikel 5.2 lid 1 Woo

over- en/of ondervertegenwoordiging nog niet acceptabel vindt dan wordt het algoritme of data aangepast. De Data-scientist levert de informatie en analyse met betrekking tot over- en/of ondervertegenwoordiging aan de projectgroep. Na beoordeling in de projectgroep, neemt de proceseigenaar een besluit.

Gehanteerd normenkader:

Bij de afweging of het algoritme wel of niet in gebruik kan worden genomen, wordt in de pilotfase een normenkader ontwikkeld. Dit normenkader bestaat uit de volgende onderdelen:

1. Geen onder- of oververtegenwoordiging

Gehanteerde norm is dat er op de discriminatoire normen van de rechten van de mens geen verschil zit tussen de aard van de selectie door het algoritme en de gehele uitkeringspopulatie. Voorbeeld: als de uitkeringspopulatie voor 40% uit vrouwen bestaat, dan moet het algoritme 40% vrouwen als uitkomst hebben.

2. Effectiviteit

De effectiviteit van het algoritme is hoger dan tenminste 1% beter dan de a-selecte methode, omdat het algoritme getraind wordt op basis van de uitkomsten van historische a-selecte heronderzoeken. Een kwantitatieve onderbouwing van de effectiviteit volgt aan het einde van de pilotfase, als het algoritme ontwikkeld is. Hierover is in paragraaf 3.3. (onderdeel stap 8) nader aandacht besteed.

3. Ethisch normenkader

Het algoritme voldoet aan het gestelde normenkader van de ethische IAMA-toets. Zie hierover meer in hoofdstuk 2.

Stap 27. Opstellen notitie over algoritme, uitlegbaarheid en individuele uitlegbaarheid

De werking van het algoritme en de uitlegbaar beschrijven ten behoeve van besluitvorming door proceseigenaar en wethouder. Als er na de toetsing op over- en ondervertegenwoordiging geen bijzondere afwijkingen geconstateerd zijn en/ of de onder- of oververtegenwoordiging is acceptabel bevonden door de product owner, dan is het algoritme compleet. Het algoritme moet goedgekeurd worden door de proceseigenaar en de wethouder voordat scoringslijst in gebruik wordt genomen. De product owner stelt een notitie op met een uitleg over de werking en uitlegbaarheid van het algoritme en legt deze ter accordering voor aan de proceseigenaar.

Stap 30. Werkzoekende voorzien van een BSN

Elke werkzoekende wordt weer voorzien van een BSN. Alleen bij het Datawarehouse kunnen bevoegde medewerkers de pseudonimisering weer terugdraaien, zodat duidelijk wordt welke risicoscore bij welk individu hoort. Zo kan worden geïdentificeerd welke personen moeten worden uitgenodigd voor heronderzoek. Het resultaat van deze stap is een lijst met uitkeringsgerechtigden, inclusief BSN's en risico-scores.

3.4 Uitgesloten van de scope

Zoals eerder aangegeven betreft deze DPIA-stap 1 t/m 24 van het deelproces II 'De ontwikkeling van het algoritme voor de uitvoering van heronderzoeken uitkeringen' binnen cluster Werk & Inkomen. In



de bijlage 2 'De ontwikkeling van het algoritme voor de uitvoering van heronderzoeken uitkeringen' is de betreffende procesbeschrijving opgenomen.

De DPIA voor de gebruiksfase (fase II) wordt opgeleverd na de ontwikkeling van het algoritme en voorafgaande aan de ingebruikname van het algoritme. De DPIA voor de gebruiksfase betreft stap 25 t/m 30 van het deelproces 2 'De ontwikkeling van het algoritme voor de uitvoering van heronderzoeken uitkeringen'; En deelproces 3 'Selecteren werkzoekenden voor heronderzoek'. In de bijlage 3 'Selecteren werkzoekenden voor heronderzoek' is de betreffende procesbeschrijving opgenomen.

Het primaire W&I proces 'Uitvoeren heronderzoeken' valt buiten de scope van beide DPIA's. Daarmee maakt deelproces 1 'Vaststellen aantal heronderzoeken en aanpak' geen onderdeel uit van deze DPIA. Voor het inzicht in het brede kader waarin het algoritme past is het deelproces 1 'Vaststellen aantal heronderzoeken en aanpak' als bijlage toegevoegd.

Het algoritme wordt ontwikkeld op het Azure DAP'R platform. Hierop is in 2021 en in mei 2022 een DPIA ('Consolideren en bewaren van gegevens in het Datalake') uitgevoerd. Ook op concern Datawarehouse is in 18 juni 2021 een DPIA ('Centrale datawarehouse') uitgevoerd.

Voor de analyse van de risico's en maatregelen op het gebied van informatiebeveiliging wordt verwezen naar de DPIA van Datalake en concern Data Warehouse. De vastgestelde beheersmaatregelen worden vanuit het project SPIDER (Cluster BCO) geïmplementeerd.

De bronapplicaties Socrates en RMW/ RAAK vallen buiten de scope van onderhavige DPIA.

Betrokken partijen

- In dit proces is het college van Burgermeesters en Wethouders van de gemeente Rotterdam de verwerkingsverantwoordelijke.
- Directeur uitvoering (portefeuille Inkomen) van cluster W&I is op clusterniveau proceseigenaar van het proces 'Ontwikkelen algoritme heronderzoeken'.
- Afdelingshoofd Beheer Inkomen van cluster W&I is op afdelingsniveau proceseigenaar van het proces heronderzoeken en gebruiker van het algoritme.
- Teamleider Data Science van afdeling OBI, cluster BCO is verantwoordelijk voor de ontwikkeling en training van het model.
- Teamleider Data Warehouse van afdeling OBI, cluster BCO is verantwoordelijk voor de verkrijging, opslag en bewerking van data binnen Concern Data Warehouse en Azure Datalake.
- Directeur uitvoering (portefeuille Inkomen) van cluster W&I is verantwoordelijk voor de inhoudelijke keuzes die in het ontwikkelproces worden gemaakt.

3.5 Beschrijving van de persoonsgegevens die worden verwerkt

In de nog uit te werken documenten worden de persoonsgegevens weergegeven. Deze documenten worden aan het einde van de ontwikkelfase van het algoritme opgeleverd. Het betreft:

- De bronlijst van variabelen uit Socrates en RMW/Raak.
- De uitkomsten van de uitgevoerde correlatieonderzoeken.
- Het voorstel en besluit welke variabelen wel en welke niet in het algoritme worden meegenomen met onderbouwing

Artikel 5.2 lid 1 Woo

Artikel 5.2 lid 1 Woo



In het algoritme worden alle discriminatoire- en proxyvariabelen gehaald, waardoor de kans op vooringenomenheid/ bias in het algoritme wordt voorkomen. Het betreft variabelen met betrekking tot de gronden waarop volgens het College van de rechten van de mens niet gediscrimineerd mag worden. Deze gronden zijn: godsdienst, levensovertuiging, politieke gezindheid, ras, geslacht, nationaliteit, seksuele gerichtheid, burgerlijke staat, handicap of chronische ziekte en leeftijd.

Artikel 5.2 lid 1 Woo

In de bijlage 5 'Correlatie corvar en features' (d.d. 12-12 is toegevoegd waarmee inzicht wordt verschaft in de bronlijst van variabelen uit Socrates en RMW/Raak en de uitkomsten van de correlatiecheck op de gronden: leeftijd, geslacht, burgerlijke staat, en handicap of chronische ziekte. Dit kan gezien worden als een tussenstand. Indien de BRP-data die we nodig hebben voor het onderzoeken van correlatie met ras/afkomst (zie stap 13) beschikbaar worden gesteld, dan wordt de uitkomsten van deze correlatiecheck verwerkt. Dat zal leiden tot een definitief voorstel en bestuurlijk besluit welke variabelen wel en welke niet in het algoritme worden meegenomen.

Risicoanalyse privacy

In dit hoofdstuk wordt per onderwerp een aantal vragen gesteld op basis van de norm die door de AVG is bepaald. Per onderwerp worden op basis van de norm risico's bepaald, welke maatregelen genomen zijn/ worden en wat de restrisico's zijn na het nemen van de maatregelen.

3.6 Rechtsgrond van de verwerking (artikel 6)

Onderstaande normen hebben betrekking op de rechtmatigheid van de verwerking van de persoonsgegevens.

Norm 1	Persoonsgegevens mogen alleen rechtmatig worden verwerkt. Dit betekent dat de verwerking plaats dient te vinden op grond van een van de grondslagen genoemd in de AVG. Voor onderhavige verwerking is artikel 6, eerste lid AVG relevant. Op grond van artikel 6, eerste lid, van de AVG is de verwerking van persoonsgegevens alleen rechtmatig indien en voor zover aan ten minste een van de in die bepaling vermelde voorwaarden is voldaan. De in die bepaling, onder e, vermelde voorwaarde is dat de verwerking noodzakelijk is 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. 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. De concrete toets is of minder inbreuk makende middelen beschikbaar zijn voor het bereiken van hetzelfde doel.
Hoe wordt voldaan aan de	1. Ten aanzien van persoonsgegevens die uit de bronsystemen Socrates en RMW/Raak als trainingsdata gebruikt worden voor het ontwikkelen van het algoritme heronderzoeken zodat uitkeringsgerechtigden met een hoge score op onrechtmatigheid door de mens geselecteerd worden voor een heronderzoek: Cluster Werk & Inkomen van de gemeente Rotterdam heeft de wettelijke taak om de Participatiewet (PW) uit te voeren en moet daarom persoonsgegevens verwerken. Werkzoekenden zijn verplicht om



norm ?	alle inlichtingen te verschaffen die van invloed kunnen zijn op de uitkering (art. 17 PW). Met name gelet op art. 53a PW, heeft de gemeente de bevoegdheid tot het doen van onderzoek naar de rechtmatigheid van de op grond van de Participatiewet verstrekte uitkering. Ter bestrijding van onjuist gebruik van sociale voorzieningen is in jurisprudentie erkend dat inzet van risicoprofielen met oog op effectiviteit, efficiëntie en kostenbesparing is toegestaan. Wel moet duidelijk zijn op basis van welke criteria mensen geselecteerd worden en waarom juist deze criteria zijn gekozen, zodat beoordeeld kan worden of een risicoprofiel binnen de grenzen van het discriminatieverbod blijft. Deze activiteiten vallen onder het uitvoeren van een taak van algemeen belang conform artikel 6, lid 1, sub e AVG (juncto art. 53a PW). Het risico-inschattingsmodel is een vorm van 'risicoprofiel'. Het risico-inschattingsmodel wordt vanuit oog op effectiviteit, efficiëntie en kostenbesparing ingezet en om te voorkomen dat uitkeringsgerechtigden onnodig worden belast met een heronderzoek. Het risico-inschattingsmodel kan in beginsel als 'noodzakelijk' worden gezien. Proportionaliteit inzet algoritme Uit de onderzoeksresultaten over de jaren 2018 tot en met 2020 blijkt dat het verrichten van heronderzoek op basis van het risico-inschattingsmodel, effectiever is dan het verrichten van heronderzoeken bij a-select gekozen uitkeringsgerechtigden of op basis van een vooraf bepaald thema. Dit uit zich in een hoger percentage beëindigde uitkeringen en terugvorderingen bij onderzoeken die in deze jaren zijn uitgevoerd op basis van het risico-inschattingsmodel. Zie hiervoor tevens de 'beantwoording van de schriftelijke vragen van het raadslid A.J. van Zevenbergen (SP) over Gemeentelijke algoritmes onder de loop, d.d. 14.12.2022'. Door het algoritme te baseren op resultaten van eerder uitgevoerde heronderzoeken bij uitkeringsgerechtigden die a-select (op basis van registratiegegevens uit de systemen Socrates en RMW/Raak) zijn uitgenodigd heeft het algoritme per definitie een hogere effectiviteit dan de a-selectie methode. Hoe groot daadwerkelijk de effectiviteit van het algoritme is, is pas na ingebruikname van het algoritme te meten. Dan kunnen we de selectiemethoden a-select, thema-onderzoek en algoritme met elkaar vergelijken. In de jaarrapportage Handhaving 2023 vindt deze effectiviteitsmeting plaats. In de huidige procesbeschrijving 'Ontwikkelen algoritme voor heronderzoeken' bij 'stap 5' is aangegeven dat de effectiviteit van het model onderzocht zal worden: <table><tr><td>Stap 5</td><td>Beoordelen effectiviteit en inventariseren verbeteringen</td></tr><tr><td>Doel</td><td>Beoordelen effectiviteit algoritme en identificeren verbeteringen o.b.v. geleerde ervaringen</td></tr><tr><td>Omschrijving</td><td> In de evaluatie worden de resultaten van heronderzoeken van voorgaande jaren verzameld en geëvalueerd. Het is belangrijk om te onderzoeken voordat je begint met het project of het inderdaad waardevol is gezien voorgaande jaren en of er wellicht zaken zijn die we kunnen leren en mee kunnen nemen in het ontwikkelen van een nieuw algoritme (n.b. het algoritme van 2020 wordt niet hergebruikt/doorontwikkeld). De evaluatie omvat o.a. de volgende elementen: • Evaluatie op de statistiek, voorspelbaarheid en afwijkende bevindingen. • Hoe verhoudt het rendement van algoritme zich tot andere selectiemethoden (zoals A-select en woonprofielen) die we gebruiken, m.a.w. wat is de toe- </td></tr></table>	Stap 5	Beoordelen effectiviteit en inventariseren verbeteringen	Doel	Beoordelen effectiviteit algoritme en identificeren verbeteringen o.b.v. geleerde ervaringen	Omschrijving	In de evaluatie worden de resultaten van heronderzoeken van voorgaande jaren verzameld en geëvalueerd. Het is belangrijk om te onderzoeken voordat je begint met het project of het inderdaad waardevol is gezien voorgaande jaren en of er wellicht zaken zijn die we kunnen leren en mee kunnen nemen in het ontwikkelen van een nieuw algoritme (n.b. het algoritme van 2020 wordt niet hergebruikt/doorontwikkeld). De evaluatie omvat o.a. de volgende elementen: • Evaluatie op de statistiek, voorspelbaarheid en afwijkende bevindingen. • Hoe verhoudt het rendement van algoritme zich tot andere selectiemethoden (zoals A-select en woonprofielen) die we gebruiken, m.a.w. wat is de toe-
Stap 5	Beoordelen effectiviteit en inventariseren verbeteringen						
Doel	Beoordelen effectiviteit algoritme en identificeren verbeteringen o.b.v. geleerde ervaringen						
Omschrijving	In de evaluatie worden de resultaten van heronderzoeken van voorgaande jaren verzameld en geëvalueerd. Het is belangrijk om te onderzoeken voordat je begint met het project of het inderdaad waardevol is gezien voorgaande jaren en of er wellicht zaken zijn die we kunnen leren en mee kunnen nemen in het ontwikkelen van een nieuw algoritme (n.b. het algoritme van 2020 wordt niet hergebruikt/doorontwikkeld). De evaluatie omvat o.a. de volgende elementen: • Evaluatie op de statistiek, voorspelbaarheid en afwijkende bevindingen. • Hoe verhoudt het rendement van algoritme zich tot andere selectiemethoden (zoals A-select en woonprofielen) die we gebruiken, m.a.w. wat is de toe-						

Artikel 5.2 lid 1 Woo



	gevoegde waarde van algoritme? Hiervoor stelt de adviseur HH een rapportage op waarin de effectiviteit van algoritme wordt vergeleken met andere selectiemethoden. • Wat zijn handhavingsresultaten versus risico inschatting; Kun je op basis van de feedback van heronderzoeken, het voorspelvermogen van het algoritme vergroten/ beter inschatten welke variabelen je moet gebruiken. • Welke variabelen zullen bijdragen aan een betere voorspelling • Reflectie op ontwikkelproces van afgelopen jaar
Naast het toepassen van een algoritme risico-inschattingsmodel moet de input aan persoonsgegevens ook noodzakelijk zijn om het doel te bereiken. Met andere woorden: persoonsgegevens die niet strikt noodzakelijk zijn voor het beoogde doel, mogen niet gebruikt worden. Voor nu is nog onbekend om welke exacte persoonsgegevens het gaat. Deze zijn nog 'in onderzoek' en staan beschreven in de procesbeschrijving onder stap 8, 9, 10, 13, 14 en 15:	
Stap 8	Onderzoeken benodigde gegevens voor algoritme
Doel	Inventariseren welke gegevens gebruikt mogen worden voor dit specifieke algoritme.
Omschrijving	De product owner is verantwoordelijk voor het onderzoeken welke gegevens gebruikt gaan worden voor het ontwikkelen van het algoritme. De product owner bepaalt met o.a. de adviseur handhaving welke gegevens van Socrates en/of Raak gebruikt moeten/ mogen worden bij het ontwikkelen van het algoritme. Deze gegevens moeten ethisch en privacy verantwoord zijn. <u>Dit betreft in beginsel alle data uit de systemen RMW/Raak en Socrates, met uitzondering van de potentieel discriminatoire variabelen / proxy's (op basis van eerdere besluitname, dan wel de correlatiechecks). Na de uitvoering van de correlatiechecks zullen indirect discriminerende variabelen uit het model worden gehaald. Vervolgens kunnen in de fase van training meerdere variabelen afvallen als deze geen effect hebben op de performance van het algoritme. Door in beginsel zoveel mogelijk variabelen uit de systemen RMW/Raak en Socrates in de ontwikkeling van het algoritme mee te nemen, wordt het komt dit het voorspellende karakter van het algoritme vergroten daarmee de performance van het algoritme ten goede.</u> De product owner kan op basis van de DWH-tabellen van Socrates, nagaan welke gegevens (kolommen) nodig zijn voor het algoritme, en welke vanwege ethische en privacy bezwaren niet meegenomen kunnen worden. Op verzoek van de product owner kan de data scientist gevraagd worden hierover advies te geven. De data scientist kan bijvoorbeeld adviseren welke gegevens er zijn. Het is de uitvoering (product owner en proceseigenaar) die bepalen welke gegevens zij willen gebruiken voor het ontwikkelen van het algoritme.
Stap 9	Opstellen voorstel benodigde gegevens
Doel	Motiveren welke gegevens nodig zijn voor het algoritme
Omschrijving	Nadat bepaald is welke gegevens nodig zijn voor het algoritme, moet een voorstel opgesteld worden voor de proceseigenaar en andere belangrijke stakeholders, met een motivatie van de keuze van bepaald gegevens.

Artikel 5.2 lid 1 Woo

Artikel 5.2 lid 1 Woo



Stap 10	Management en bestuurlijke goedkeuring gegevens voor het algoritme
Doel	Goedkeuring data die benodigd zijn voor het ontwikkelen van algoritme voor heronderzoeken
Omschrijving	De product owner stelt een voorstel op m.b.t. de benodigde data voor het ontwikkelen van het algoritme. De data scientist levert input hiervoor aan de product owner. Het voorstel wordt ter goedkeuring voorgelegd aan de proceseigenaar, directeur Inkomen en wethouder.
Stap 13	Onderzoeken verbanden tussen variabelen en bijzondere persoonsgegevens
Doel	Zorgen voor een ethisch verantwoorde dataset
Omschrijving	Het doel is om variabelen die een sterk verband hebben met variabelen die wijzen op vooringenomenheid, niet mee te nemen in het algoritme. Om deze reden nemen we bijzondere persoonsgegevens niet op. In deze stap wordt voorkomen dat zgn proxy variabelen worden meegenomen in het ontwikkelen van het algoritme. Hierbij wordt uitgegaan van de discriminatoire normen van de rechten van de mens. Na deze stap is duidelijk welke data gebruikt gaat worden om het algoritme te ontwikkelen.
Stap 14	Opstellen voorstel trainingsdata voor proceseigenaar
Doel	Het motiveren van de keuze voor de benodigde dataset voor het ontwikkelen van het algoritme.
Omschrijving	Na het onderzoek op het ethisch toepassen van de data voor het algoritme, moeten de resultaten en conclusies gemotiveerd worden. De motivatie wordt samen met het advies ter goedkeuring voorgelegd aan de proceseigenaar. Hierbij wordt uitgegaan van de gronden op basis waarvan volgens het College van de Rechten van de Mens niet gediscrimineerd mag worden. Zonder formele goedkeuring van de dataset wordt niet gestart met het ontwikkelen van het algoritme. De onderbouwing voor de trainingsdata en advies wordt vastgelegd in een notitie voorstel en voorgelegd ter goedkeuring aan de proceseigenaar. In het voorstel moet ook gemotiveerd worden waarom bepaalde (relevante/nuttige) variabelen en kenmerken NIET worden meegenomen om het algoritme te ontwikkelen.
Stap 15	Goedkeuren voorstel trainingsdata
Doel	Het beschikken over de goedkeuring van de proceseigenaar om de voorgestelde dataset te gebruiken voor het ontwikkelen van het algoritme.
Omschrijving	De proceseigenaar neemt op basis van het voorstel m.b.t. de gewenste dataset voor het ontwikkelen van het algoritme, een besluit. De data scientist, Adviseur HH en product owner hebben vanwege hun kennis een belangrijke adviserende rol richting de proceseigenaar, zodat deze de impact en consequenties van het besluit goed kan doorgronden.

Wel staat vast dat de input voor de ontwikkeling van het algoritme gepseudonimiseerde persoonsgegevens zijn die oorspronkelijk verzameld zijn voor de uitvoering van de PW (afkomstig uit de bron-systemen Socrates en RMW/RAAK).

In het algoritme worden alle (proxy)variabelen weggehaald, waardoor die een discriminerende werking zouden kunnen hebben. Het betreft variabelen met betrekking tot de gronden waarvan volgens

Artikel 5.2 lid 1 Woo



het College van de rechten van de mens gediscrimineerd zou kunnen worden. Deze gronden zijn: godsdienst, levensovertuiging, politieke gezindheid, ras, geslacht, nationaliteit, seksuele gerichtheid, burgerlijke staat, handicap of chronische ziekte en leeftijd.

In de procesbeschrijving 'Ontwikkelen algoritme voor heronderzoeken' zijn expliciet stappen opgenomen waarin kritisch wordt gekeken naar de gebruikte variabelen. Er zijn waarborgen ingebouwd om vooringenomenheid binnen het algoritme te voorkomen. Zie procesbeschrijving bij de stappen '3, 4, 8, 13, 14, 15, 24 en 26':

Stap 13	Onderzoeken verbanden tussen variabelen en bijzondere persoonsgegevens
Doel	Zorgen voor een ethisch verantwoorde dataset
Omschrijving	Het doel is om variabelen die een sterk verband hebben met variabelen die wijzen op vooringenomenheid, niet mee te nemen in het algoritme. Om deze reden nemen we bijzondere persoonsgegevens niet op. In deze stap wordt voorkomen dat zgn proxy variabelen worden meegenomen in het ontwikkelen van het algoritme. Na deze stap is duidelijk welke data gebruikt gaat worden om het algoritme te ontwikkelen.

Stap 26	Toetsen scoringslijst op onder- of oververtegenwoordiging (discriminatie toetsing)
Doel	De groep met de hoogste score op onrechtmatigheid toetsen op discriminatie en vooringenomenheid.
Omschrijving	Scoringslijst = lijst met alle actieve werkzoekenden en een risico inschatting. Het is niet wenselijk meer of juist minder mensen met bepaalde bijzondere karakteristieken te selecteren dan we zouden mogen verwachten in deze groep. Deze toets is een extra check om te bepalen of het model zonder gebruik te maken van discriminerende variabelen tot een vooringenomen resultaat komt. De toetsing is tweeledig; de data scientist zal voornamelijk een cijfermatige toetsing uitvoeren en de product owner zal samen met de Adviseur Handhaving de toetsing uitvoeren o.b.v. de variabelen. Deze toetsing is ook om na te gaan of er toch sprake is van gebruik van proxy variabele, ondanks dat deze bij de selectie van data zijn uitgesloten. Als de product owner de over- en/of ondervertegenwoordiging nog niet acceptabel vindt dan wordt het algoritme of data aangepast, of wordt een nieuw algoritme ontwikkeld. <u>Gehanteerde norm is dat er op de discriminatoire normen van de rechten van de mens geen verschil zit tussen de aard</u>

Artikel 5.2 lid 1 Woo



	<u>van de selectie door het algoritme en de gehele uitkeringspopulatie. Voorbeeld: als de uitkeringspopulatie voor 40% uit vrouwen bestaat, dan moet het algoritme 40% vrouwen als uitkomst hebben.</u> De Data scientist levert de informatie en analyse m.b.t over- en/of ondervertegenwoordiging aan de product owner. De product owner beoordeelt dit en neemt al dan niet in samenspraak met de proceseigenaar een besluit. Noot: als dit proces voor het eerst wordt uitgevoerd, zal pas duidelijk worden hoe dit proces concreet uitgevoerd moet worden.
	Het is een voorspellend algoritme (machine learning) en geen voorschrijvend algoritme. Het uiteindelijke heronderzoek wordt door een mens uitgevoerd en er zijn waarborgen in het primaire proces aanwezig voor het nemen van een zorgvuldig besluit. 2. Ten aanzien van persoonsgegevens die gebruikt worden om vooraf het algoritme te toetsen op mogelijke discriminatie op nationaliteit: De wettelijke grondslag voor het gebruik van gegevens over nationaliteit als controlemiddel bij het algoritme moet worden gezocht in artikel 53a, van de Participatiewet waarin de algemene bevoegdheid voor het college voor het verrichten van (her)onderzoeken naar de rechtmatigheid van verleende bijstandsuitkeringen is geregeld. Op 18 juni 2021 heeft Juridische Diensten een positief advies gegeven over het gebruik van de BRP-data voor de uitvoering van het correlatieonderzoek. In de bijlage is het advies van Juridische Diensten toegevoegd. Zie verder de onderbouwing hierboven onder punt 1. Na een second-opinion adviesaanvraag over de wettelijke grondslag van onderhavige verwerking bij Juridische Diensten is de wettelijke grondslag bevestigd Zie bijlage advies juridische diensten
Wat zijn de risico's die overblijven?	De volgende risico's blijven aanwezig ook ingeval alle processtappen conform het proces 'Ontwikkelen algoritme voor heronderzoeken' zijn uitgevoerd. Risico's voor betrokkene 1. Discriminatie van een werkzoekende 2. Financieel verlies werkzoekende (terugwerkende kracht + toekomstige uitkeringen) 3. Reputatieschade werkzoekende Risico's voor organisatie 1. Onrechtmatige verwerking als ondanks de getroffen waarborgen het risico-inschattingsmodel onbewust discriminerend werkt, waardoor bepaalde groepen/ werkzoekenden sneller uitgenodigd worden voor een heronderzoek dan andere. 2. Reputatieschade, verlies van vertrouwen, sancties toezichthouders, klachten en gerechtelijke procedures, financiële schade.



Hoe groot is de impact voor de betrokkene ?	Hoog.
Hoe groot is de kans dat het risico zich voordoet voor de betrokkene ?	Laag, want diverse waarborgen zijn aanwezig om discriminatie, bias te voorkomen (zie procesbeschrijving). Als blijkt dat er desondanks discriminerende werking zou kunnen zijn, dan worden de (indirecte) discriminatoire variabele verwijderd. Mocht het model dan niet meer bruikbaar zijn, dan wordt ermee gestopt.
Aanvullende maatregelen risico's betrekken ?	1. Voer de IAMA op het nieuwe algoritmemodel z.s.m. uit en de daarin voorgesteld processtappen (en documenteer). 2. Zorg dat het proces 'Ontwikkelen algoritme voor heronderzoeken' aansluit met nieuwe ontwikkelingen rondom AI-wetgeving, zoals de AI-verordening. 3 a. Voer alle processtappen conform het proces 'Ontwikkelen algoritme voor heronderzoeken' uit (en documenteer). 3 b. Controleer periodiek of alle processtappen conform het proces 'Ontwikkelen algoritme voor heronderzoeken' zijn gevolgd. 4. Laat het risico-inschattingsmodel toetsen door een onafhankelijk algoritme toetsingscomité. 5 Doe een juridische toets op de definitieve bronlijst.

Artikel 5.2 lid 1 Woo

Norm 2	Indien de verwerking is gebaseerd op toestemming, dient aan de eisen in de AVG te zijn voldaan.
Hoe wordt voldaan aan de norm?	De verwerking is niet gebaseerd op toestemming. De verwerking is gebaseerd op een taak van algemeen belang, omdat de Participatiewet wordt uitgevoerd.
Wat zijn de risico's die overblijven?	Niet van toepassing



Hoe groot is de impact voor de betrokkene?	Niet van toepassing
Hoe groot is de kans dat het risico zich voordoet?	Niet van toepassing
Aanvullende maatregelen?	Niet van toepassing

Norm 3	Indien de verwerking plaats vindt op basis van toestemming heeft de betrokkene het recht om toestemming te allen tijde in te trekken.
Hoe wordt voldaan aan de norm?	Zie norm 2.
Wat zijn de risico's die overblijven?	Niet van toepassing
Hoe groot is de impact voor de betrokkene?	Niet van toepassing
Hoe groot is de kans dat het risico zich voordoet?	Niet van toepassing
Aanvullende maatregelen?	Niet van toepassing

Norm 4	Het verwerken van bijzondere persoonsgegevens is alleen toegestaan, indien wordt voldaan aan een van de 10 uitzonderingen onder de AVG.
Hoe wordt voldaan aan de norm?	1. Uit de DPIA en het ingewonnen juridisch advies blijkt dat er bijzondere persoonsgegevens verwerkt gaan worden om achteraf het algoritme te valideren op mogelijke discriminatie op de norm nationaliteit en mogelijk andere discriminatiegronden. In de DPIA is aangegeven dat het gebruik van de data afkomstig uit de BRP enkel wordt gebruikt voor de uitvoering van de interne validatie betreffende een mogelijke correlatie met nationaliteit. Met als doel om mogelijke discriminatie van het algoritme objectief en zoveel als mogelijk uit te sluiten. De BRP-data worden op geen enkele wijze meegenomen in het model van algoritme. Juridisch advies luidt: <i>"De wettelijke grondslag voor het gebruik van gegevens over nationaliteit als controlemiddel bij het algoritme moet worden gezocht in artikel 53a, van de Participatiewet waarin de algemene bevoegdheid voor het college voor het verrichten van (her)onderzoeken naar de rechtmatigheid van verleende bijstandsuitkeringen is geregeld."</i> Uitgaande van het juridisch advies kan gesteld worden dat een wettelijke grondslag aanwezig is. De wettelijke grondslag is artikel 9, lid 2, sub a AVG juncto artikel 53a PW. De verwerking is noodzakelijk in het kader van de uitvoering van regels op het



	gebied van arbeids- en socialezekerheidsrecht (in casu artikel 53a PW). 2. Gesteld wordt dat ook bijzondere persoonsgegevens uit de bronsystemen Socrates en RMW/Raak als historische data gebruikt gaan worden voor het ontwikkelen van het algoritme heronderzoeken. Voor nu is nog onbekend om welke bijzondere persoonsgegevens het gaat. Zodra deze bekend zijn en de noodzakelijkheid ervan onderbouwd is, kan zeer waarschijnlijk dezelfde uitzonderingsgrond als hierboven onder punt 1 worden gehanteerd. In de DPIA is aangegeven dat het ontwikkelde algoritme in de praktijk toegepast wordt als het objectief is aangetoond dat het geen discriminerende werking heeft.
Wat zijn de risico's die overblijven?	1. Geen 2. Verlies controle over eigen gegevens bij onrechtmatige verwerkingen: het ontwikkelde algoritme wordt in de praktijk toegepast als het objectief is aangetoond dat het geen discriminerende werking heeft. Daarom is het uitgangspunt dat vanuit het perspectief van betrokkene het opgeleverde algoritme geen risico's voor wat betreft vooringenomenheid zal hebben.
Hoe groot is de impact voor de betrokkene?	2. Hoog
Hoe groot is de kans dat het risico zich voordoet?	2. Laag
Aanvullende maatregelen?	1. Onderbouw waarom en welke bijzondere persoonsgegevens noodzakelijk zijn voor het bereiken van het doel en verwerk geen bijzondere persoonsgegevens in het algoritme als deze niet noodzakelijk zijn voor het bereiken van het doel. 2. Voer de processtappen in de Procesbeschrijving 'Ontwikkelen algoritme voor heronderzoeken' en de processtappen conform de 'IAMA' uit (en documenteer) om continu risico's te mitigeren.

Norm 5	Het BSN mag alleen worden verwerkt als dit expliciet is toegestaan in de wet op het BSN of in een andere wet
Hoe wordt voldaan aan de norm?	De gebruikte gegevens komen via een voorportaal uit de bronsystemen RMW/Raak en Socrates. In Socrates en RMW/Raak wordt informatie over uitkeringsgerechtigden met behulp van het BSN vastgelegd. De persoonsgegevens uit deze bronsystemen worden bij de start van de levering door een geautoriseerde en gemandateerde medewerker in het voorportaal gepseudonimiseerd. Het voorportaal houdt geen data vast en heeft alleen als doel om bron van DAP'R af te schermen (dataminalisatie) en privacygevoelige data te versleutelen. Vanuit het voorportaal wordt het concern Datawarehouse gevoed en via Azure Datalake aangeleverd bij het team Advanced Analytics van de afdeling OBI. Pseudonimiseren houdt in dat in plaats van alle BSN's een nieuw nummer zichtbaar wordt. Op geen enkele manier is uit dit nieuwe nummer het oude BSN te herleiden (zonder oorspronkelijke sleutel). Dit nieuwe nummer gebruiken twee collega's van



	het team Advanced Analytics op een wijze, zodat uitsluitend zij weten welke informatie bij welk nummer (persoon) hoort zonder dat zij weten wie deze persoon daadwerkelijk is. Nadat de vanuit de systemen Socrates en RMW/Raak beschikbaar is gesteld en in de Azure omgeving is verwerkt, is een model vastgesteld waarmee de uitkeringspopulatie gescoord is. Hierbij wordt de oorspronkelijke sleutel gebruikt om de resultaten terug te zetten naar BSN's. Hiermee kan cluster W&I aan de slag met het daadwerkelijk uitnodigen van uitkeringsgerechtigden voor een heronderzoek.
Wat zijn de risico's die overblijven?	Geen
Hoe groot is de impact voor de betrokkene?	Niet van toepassing
Hoe groot is de kans dat het risico zich voordoet?	Niet van toepassing
Aanvullende maatregelen?	Niet van toepassing

3.7 Doelbinding (artikel 5)

Voor achtergrondinformatie zie ook: <https://autoriteitpersoonsgegevens.nl/nl/over-privacy/persoonsgegevens/verstrekken-van-persoonsgegevens>

Norm 6	Indien de gegevens uit bestaande bronnen komen (een andere applicatie) of binnen een ander proces zijn verzameld, dient het doel van deze verwerking verenigbaar te zijn met het oorspronkelijke doel.
Hoe wordt voldaan aan de norm?	Er is sprake van doelbinding. Er worden in het algoritme uitsluitend gegevens gebruikt die in de administratie zijn vastgelegd in het kader van de uitvoering van de Participatiewet. En in het bijzonder ten behoeve van het uitvoeren van heronderzoeken conform art.53a PW). Het algoritme ondersteunt het proces van het uitvoeren van heronderzoeken, omdat er scores op onrechtmatigheid uitrollen, waarna de organisatie een selectie maakt van werkzoekenden met hoge scores om vervolgens uit te nodigen voor een heronderzoek van de uitkeringsrechtmatigheid. Ten aanzien van de persoonsgegevens die gebruikt worden om vooraf het algoritme te toetsen op mogelijke discriminatie op de norm nationaliteit, is het gebruik van BRP-data een vereiste. Het doel van dit correlatieonderzoek is verenigbaar met het oorspronkelijke doel: het uitvoeren van de Participatiewet. De rechtmatigheid van de verwerking van persoonsgegevens uit de BRP is bij norm 1 verantwoord.
Wat zijn de risico's die overblijven?	Geen



Hoe groot is de impact voor de betrokkene?	Niet van toepassing
Hoe groot is de kans dat het risico zich voordoet?	Niet van toepassing
Aanvullende maatregelen?	Niet van toepassing

Norm 7	Gegevens die in deze verwerking worden verwerkt, mogen alleen in andere verwerkingen worden gebruikt (gedeeld worden voor andere activiteiten) als het doel van die verdere verwerking (die andere activiteit) verenigbaar is met het doel van deze verwerking.
Hoe wordt voldaan aan de norm?	Niet van toepassing. Er is geen sprake van een verdere verwerking.
Wat zijn de risico's die overblijven?	Niet van toepassing
Hoe groot is de impact voor de betrokkene?	Niet van toepassing
Hoe groot is de kans dat het risico zich voordoet?	Niet van toepassing
Aanvullende maatregelen?	Niet van toepassing

3.8 Juist, actueel en volledig

Voor achtergrondinformatie zie: <https://eur-lex.europa.eu/legal-content/NL/TXT/HTML/?uri=CELEX:32016R0679&from=EN> Art. 5, sub 1d

Norm 8	Persoonsgegevens zijn juist, actueel en volledig.
Hoe wordt voldaan aan de norm?	Er wordt gebruik gemaakt van de laatst beschikbare data uit de bronsystemen RMW/Raak en Socrates. In de uitvoeringsprocessen binnen Werk & Inkomen is de kwaliteit van de persoonsgegevens gewaarborgd door interne controles door kwaliteitsmedewerker (4-ogen principe). En in de werkprocessen en -instructies zijn het doorvoeren van mutaties door daartoe bevoegde medewerkers (zoals Administratief Ondersteuners en Inkomensconsulenten) belegd. Ook vinden er kwaliteitscontroles plaats op reguliere processen. Zie verder onder norm 1 hoe data verder voorbereidt wordt ten behoeve van het risico-inschattingmodel. Bij de ontwikkeling van het model wordt uitgegaan van artikel 10 van de WAI (Wet op de Artificiële Intelligentie -concept-). In het ontwerp is gekozen voor ontwikkeling van het model op basis van historische random samples van heronderzoeken. Er is gekozen voor de ontwikkeling van een



	glass box model vanwege uitlegbaarheid en transparantie. De data worden gecontroleerd op geschiktheid, volledigheid en representativiteit d.m.v. diverse controles (bijvoorbeeld op aantallen en voorkomen van combinaties van variabelen). De data worden geprepareerd en opgeschoond, waarbij variabelen vanuit diverse tabellen (afkomstig uit Socrates en RMW) worden omgezet tot een trainingsdataset met features. Mogelijke vertekeningen (bias) worden geanalyseerd d.m.v. correlatiechecks met bijzondere persoonsgegevens, zoals bedoeld in artikel 9, lid 1, van Verordening (EU) 2016/679, artikel 10 van Richtlijn (EU) 2016/680 en artikel 10, lid 1, van Verordening (EU) 2018/1725. De uitkomsten van de correlatiechecks en de keuze van variabelen worden besproken tijdens diverse werksessies en voorgelegd aan de opdrachtgever.
Wat zijn de risico's die overblijven?	Bias (onwenselijke systematische afwijkingen van de uitkomsten van het algoritme voor de gehele groep of juist voor specifieke personen of deelgroepen)
Hoe groot is de impact voor de betrokkene?	Hoog
Hoe groot is de kans dat het risico zich voordoet?	Laag (gezien de aangegeven maatregelen)
Aanvullende maatregelen?	Voer periodiek controles uit (tijdens ontwerp en uitvoering) om de juiste werking van het algoritme te garanderen.

4.4 Dataminimalisatie

Voor achtergrondinformatie zie: <https://www.privacy-regulation.eu/nl/5.htm>

Norm 9	Er worden niet meer persoonsgegevens gebruikt dan noodzakelijk voor het bereiken van het doel van de verwerking. Van elk persoonsgegeven is vastgesteld dat het noodzakelijk is.						
Hoe wordt voldaan aan de norm?	Persoonsgegevens die niet strikt noodzakelijk zijn om het doel te bereiken, mogen niet gebruikt worden. Voor nu is nog onbekend om welke exacte persoonsgegevens het gaat en waarom. Deze zijn nog 'in onderzoek'. In het proces 'Ontwikkelen algoritme voor heronderzoeken' zijn de processtappen voor het selecteren van persoonsgegevens beschreven onder stap 8, 9, 10, 13, 14 en 15: <table><tr><th>Stap 8</th><th>Onderzoeken benodigde gegevens voor algoritme</th></tr><tr><td>Doel</td><td>Inventariseren welke gegevens gebruikt mogen worden voor dit specifieke algoritme.</td></tr><tr><td>Omschrijving</td><td>De product owner is verantwoordelijk voor het onderzoeken welke gegevens gebruikt gaan worden voor het ontwikkelen van het algoritme. De product owner bepaalt met o.a. de adviseur handhaving welke gegevens van Socrates en/of Raak gebruikt moeten/ mogen worden bij het ontwikkelen van het algoritme. Deze gegevens moeten ethisch en privacy verantwoord zijn.</td></tr></table>	Stap 8	Onderzoeken benodigde gegevens voor algoritme	Doel	Inventariseren welke gegevens gebruikt mogen worden voor dit specifieke algoritme.	Omschrijving	De product owner is verantwoordelijk voor het onderzoeken welke gegevens gebruikt gaan worden voor het ontwikkelen van het algoritme. De product owner bepaalt met o.a. de adviseur handhaving welke gegevens van Socrates en/of Raak gebruikt moeten/ mogen worden bij het ontwikkelen van het algoritme. Deze gegevens moeten ethisch en privacy verantwoord zijn.
Stap 8	Onderzoeken benodigde gegevens voor algoritme						
Doel	Inventariseren welke gegevens gebruikt mogen worden voor dit specifieke algoritme.						
Omschrijving	De product owner is verantwoordelijk voor het onderzoeken welke gegevens gebruikt gaan worden voor het ontwikkelen van het algoritme. De product owner bepaalt met o.a. de adviseur handhaving welke gegevens van Socrates en/of Raak gebruikt moeten/ mogen worden bij het ontwikkelen van het algoritme. Deze gegevens moeten ethisch en privacy verantwoord zijn.						

Tabel met opmaak



Artikel 5.2 lid 1 Woo

	De product owner kan op basis van de DWH-tabellen van Socrates, nagaan welke gegevens (kolommen) nodig zijn voor het algoritme, en welke vanwege ethische en/of privacy bezwaren niet meegenomen kunnen worden. Op verzoek van de product owner kan de data scientist gevraagd worden hierover advies te geven. De data scientist kan bijvoorbeeld adviseren welke gegevens er zijn. Het is de uitvoering (product owner en proceseigenaar) die bepalen welke gegevens zij willen gebruiken voor het ontwikkelen van het algoritme.
Stap 9	Opstellen voorstel benodigde gegevens
Doel	Motiveren welke gegevens nodig zijn voor het algoritme
Omschrijving	Nadat bepaald is welke gegevens nodig zijn voor het algoritme, moet een voorstel opgesteld worden voor de proceseigenaar en andere belangrijke stakeholders, met een motivatie van de keuze van bepaald gegevens.
Stap 10	Management en bestuurlijke goedkeuring gegevens voor het algoritme
Doel	Goedkeuring data die benodigd zijn voor het ontwikkelen van algoritme voor heronderzoeken
Omschrijving	De product owner stelt een voorstel op m.b.t. de benodigde data voor het ontwikkelen van het algoritme. De data scientist levert input hiervoor aan de product owner. Het voorstel wordt ter goedkeuring voorgelegd aan de proceseigenaar, directeur Inkomen en wethouder.
Stap 13	Onderzoeken verbanden tussen variabelen en bijzondere persoonsgegevens
Doel	Zorgen voor een ethisch verantwoorde dataset
Omschrijving	Het doel is om variabelen die een sterk verband hebben met variabelen die wijzen op vooringenomenheid, niet mee te nemen in het algoritme. Om deze reden nemen we bijzondere persoonsgegevens niet op. In deze stap wordt voorkomen dat zgn proxy variabelen worden meegenomen in het ontwikkelen van het algoritme. Na deze stap is duidelijk welke data gebruikt gaat worden om het algoritme te ontwikkelen.
Stap 14	Opstellen voorstel trainingsdata voor proceseigenaar
Doel	Het motiveren van de keuze voor de benodigde dataset voor het ontwikkelen van het algoritme.
Omschrijving	Na het onderzoek op het ethisch toepassen van de data voor het algoritme, moeten de resultaten en conclusies gemotiveerd worden. De motivatie wordt samen met het advies ter goedkeuring voorgelegd aan de proceseigenaar. Zonder formele goedkeuring van de dataset wordt niet gestart met het ontwikkelen van het algoritme. De onderbouwing voor de trainingsdata en advies wordt vastgelegd in een notitie voorstel en voorgelegd ter goedkeuring aan de proceseigenaar. In het voorstel moet ook gemotiveerd worden waarom bepaalde (relevante/nuttige) variabelen en kenmerken NIET worden meegenomen om het algoritme te ontwikkelen.
Stap 15	Goedkeuren voorstel trainingsdata
Doel	Het beschikken over de goedkeuring van de proceseigenaar om de voorgestelde dataset te gebruiken voor het ontwikkelen van het algoritme.
Omschrijving	De proceseigenaar neemt op basis van het voorstel m.b.t. de gewenste dataset voor het ontwikkelen van het algoritme, een besluit.



	De data scientist, Adviseur HH en product owner hebben vanwege hun kennis een belangrijke adviseerende rol richting de proceseigenaar, zodat deze de impact en consequenties van het besluit goed kan doorgronden.
Wat zijn de risico's die overblijven?	Risico's zijn altijd aanwezig bij ontwikkeling van een algoritme. Door de te treffen processtappen zoals hierboven weergegeven worden risico's gemitigeerd. Het blijft wel van belang dat alle processtappen gevolgd (en gedocumenteerd) worden. Onder processtappen wordt ook de processtappen in de IAMA bedoeld.
Hoe groot is de impact voor de betrokkene?	Hoog
Hoe groot is de kans dat het risico zich voordoet?	Midden
Aanvullende maatregelen?	Voer de processtappen in de Procesbeschrijving 'Ontwikkelen algoritme voor heronderzoeken' en de processtappen conform de 'IAMA' uit (en documenteer) om continu risico's te identificeren en te kunnen mitigeren.

Norm 10	Het gebruik van bijzondere persoonsgegevens en/of gebruik van gevoelige gegevens dient zo veel mogelijk vermeden te worden.
---------	---



Hoe wordt voldaan aan de norm?	De brondata uit Socrates en RMW/Raak wordt gebruikt. Beide systemen worden gebruikt voor het onder andere vaststellen van het recht op een uitkering en voor de uitbetaling daarvan. De bijzondere persoonsgegevens die uit deze systemen wordt opgehaald, worden gebruikt voor het correlatieonderzoek. Op deze manier wordt geborgd dat er in het daadwerkelijke model geen discriminatoire variabelen zitten. In het uiteindelijke model dat gebruikt wordt zullen geen bijzondere persoonsgegevens worden gebruikt. Nationaliteit is niet opgenomen in het gebruikte model. Daarnaast is in juni 2021 een aantal variabelen aangemerkt die mogelijk indirect verband houden met nationaliteit (oftewel proxy-variabelen). Deze worden niet meer gebruikt bij de ontwikkeling van het model voor 2022. Het onderzoek naar de vraag of er nog variabelen worden gebruikt die sterke correlatie vertonen met bijzondere persoonsgegevens (en dus kunnen worden aangemerkt als proxy-variabelen) is als stap 17 opgenomen in het 'proces Ontwikkelen algoritme voor heronderzoeken'. Zie tekst onder norm 4.
Wat zijn de risico's die overblijven?	Zie tekst onder norm 4.
Hoe groot is de impact voor de betrokkene?	Zie tekst onder norm 4.
Hoe groot is de kans dat het risico zich voordoet?	Zie tekst onder norm 4.
Aanvullende maatregelen?	Zie tekst onder norm 4.

Norm 11	Vrije tekstvelden in applicaties dienen te worden vermeden.
Hoe wordt voldaan aan de norm?	Ingevoerde teksten in vrije tekstvelden in met name RMW/Raak, betreffen veelal een subjectieve omschrijving van persoonskenmerken van uitkeringsgerechtigden. Aangezien hier sprake kan zijn van een bepaalde mate van vooringenomenheid heeft de directeur Inkomen (proceseigenaar) besloten om alle variabelen/features met vrije invoer tekstvelden uit de verdere ontwikkeling van het algoritme te houden.
Wat zijn de risico's die overblijven?	Het gebruik van (resterende) variabelen die bestaan uit vrij invulbare velden, vergroten de kans op subjectiviteit en mogelijke bias. Zie verder norm 1.
Hoe groot is de impact voor de betrokkene?	Hoog Het uiteindelijke heronderzoek wordt door een mens uitgevoerd. Dit geldt tevens voor het nemen van een zorgvuldig besluit betreffende het uitgevoerde heronderzoek. Wel kan een burger bij bias in het algoritme de kans lopen om eerder/ vaker te worden uitgenodigd voor een heronderzoek op basis van subjectieve oordelen.
Hoe groot is de kans dat het risico zich voordoet?	Midden In de procesbeschrijving 'Ontwikkelen algoritme voor heronderzoeken' zijn expliciet stappen opgenomen waarin kritisch wordt gekeken naar de gebruikte variabelen. Er



	zijn waarborgen ingebouwd om vooringenomenheid binnen het algoritme te voorkomen. Er is sprake van een voorspellend algoritme en geen voorschrijvend algoritme. Het uiteindelijke onderzoek wordt door een mens uitgevoerd en er zijn voldoende waarborgen in het primaire proces aanwezig voor het nemen van een zorgvuldig besluit.
Aanvullende maatregelen?	1. Onderzoek periodiek (tijdens ontwerp en uitvoering) het risico-inschattingsmodel en de toepassing daarvan op discriminatie en bias, zodat deze binnen de grenzen van het discriminatieverbod blijven en voorkom daarmee dat onwenselijke systematische afwijkingen voor specifieke werkzoekenden of groepen ontstaan. 2. Onderzoek, onderbouw en leg vast (periodiek) de noodzakelijkheid van de persoonsgegevens (historische data) die als input dienen voor het risico-inschattingsmodel. 3. Laat een onafhankelijk algoritme toetsingscomité het risico-inschattingsmodel te toetsen.

4.5 Bewaartermijn

Norm 12	Persoonsgegevens worden niet langer worden bewaard dan noodzakelijk voor het doel van de verwerking en niet langer dan wettelijk is voorgeschreven.
Hoe wordt voldaan aan de norm?	Het algoritme, alle documentatie en persoonsgegevens betreffende de ontwikkeling van het algoritme betreft het proces "data science projecten", nummer P3693 in het informatiebeheerplan. De bewaartermijn is in principe "10 jaar na bestaansrecht van het procesobject". Vrij vertaald 10 jaar na de oplevering van het algoritme. Voor het betreffende proces geldt een formele doorlopende machtiging van vernietiging (d.d. 8-6-2021). Aangezien het een algoritme is met een ingrijpend maatschappelijk effect, is op advies van Records Management een permanente een permanente bewaartermijn op grond van categorie 2.1.2. uit de gemeentelijke selectielijst vastgesteld. De documentatie en data betreffende het algoritme wordt conform de kaders van Informatiebeheer gearchiveerd in het systeem DIVA. Na het verstrijken van de bewaartermijn moet de documentatie en data daadwerkelijk vernietigd te worden of in betreffende dit algoritme overgedragen worden aan het Stadsarchief voor permanente archivering. Binnen afdeling Informatiebeheer en OBI van BCO is aandacht voor het op lange termijn uitvoering geven hieraan. In dit kader wordt dit jaar i.s.m. privacy, security en informatiebeheer (SPIDER) gestart met een uitgebreide analyse op alle datasets. Dit geldt dus niet alleen voor deze dataset. Mochten er bevindingen naar boven komen die niet compliant zijn, dan zal dit worden rechtgezet. Tevens lopen er acties t.b.v. continu monitoring hierop (meta datamanagement).
Wat zijn de risico's die overblijven?	Geen of niet tijdige vernietiging met als gevolg dat de data in handen komt van onbevoegden.

Artikel 5.2 lid 1 Woo



Hoe groot is de impact voor de betrokkene?	Laag
Hoe groot is de kans dat het risico zich voor- doet?	Laag
Aanvullende maatregelen?	Controleer of daadwerkelijk na afloop van de bewaartermijn de betreffende data is vernietigd.

4.6 Verwijdering

Norm 13	Indien persoonsgegevens niet langer nodig zijn moeten deze, zowel fysiek als digitaal, worden verwijderd. De persoonsgegevens moeten op een dusdanige manier verwijderd dat de gegevens niet meer te benaderen en/of te gebruiken zijn
Hoe wordt voldaan aan de norm?	Voor het betreffende proces geldt een permanente bewaartermijn. Zie orm 12). Dit betekent dat de persoonsgegevens volgens de concern brede kaders en procedures van Informatiebeheer voor permanente archivering worden overgedragen aan het Stadsarchief.
Wat zijn de risico's die overblijven?	Geen overdracht aan het Stadsarchief voor permante archivering met als gevolg dat de data in handen komt van onbevoegden.
Hoe groot is de impact voor de betrokkene?	Laag
Hoe groot is de kans dat het risico zich voor- doet?	Laag
Aanvullende maatregelen?	Controleer of daadwerkelijk voor permanente archivering is overgedragen aan het Stadsarchief.

4.7 Locatie

Norm 14	Verwerking van persoonsgegevens buiten de EU is alleen toegestaan, indien het derde land over een passend beschermingsniveau beschikt
Hoe wordt voldaan aan de norm?	Niet van toepassing. Er vindt geen gegevensuitwisseling plaats buiten de EU. De gegevens die worden gebruikt zijn beschikbaar gesteld uit de systemen BRP, RMW/Raak en Socrates. De gegevens worden slechts binnen de EU gebruikt. Dit geldt tevens voor de DAP'R omgeving, waarvan de Datawarehouse en Azure onderdeel van uitmaken. De omgeving voldoet aan de Baseline Informatiebeveiliging Overheid (BIO) – eisen.
Wat zijn de risico's die overblijven?	Niet van toepassing
Hoe groot is de impact voor de betrokkene?	Niet van toepassing



Hoe groot is de kans dat het risico zich voor- doet?	Niet van toepassing
Aanvullende maatregelen?	Niet van toepassing

4.8 Anonimiseren en/of pseudonimiseren

Norm 15	De mogelijkheid van anonimiseren of pseudonimiseren worden optimaal benut.
Hoe wordt voldaan aan de norm?	Voordat de persoonsgegevens uit Socrates, RMW/Raak in het Concern Datawarehouse worden geladen, worden de privacygevoelige gegevens gepseudonimiseerd. Zie tevens uiteenzetting van norm 5. Omdat geboortemaand en buurt van belang zijn voor het algoritme, wordt bij het doorladen van Concern Datawarehouse naar Azure Datalake dataminimalisatie toegepast op de persoonsgegevens geboortedag en postcode. Bij de geboortedatum wordt de geboortedag verwijderd. Bij de postcodegegevens wordt de letteraanduiding verwijderd. De BRP-data voor het vaststellen van een mogelijke correlatie met afkomst/ras worden direct na de afronding van de correlatiecheck vernietigd conform de kaders van de Archiefwet en het concern informatiebeheerbeleid. De BRP-data belanden niet in de (trainings)data voor het algoritme. De BRP-data worden enkel gepseudonimiseerd vrijgegeven aan de data scientists van team Advanced Analytics van OBI die bij de ontwikkeling van het algoritme betrokken zijn. Deze data scientists zijn in vaste dienst bij de gemeente Rotterdam. De BRP -data worden tweemaal aan de data scientists ter beschikking gesteld en door hen gebruikt, te weten: 1. Voor het uitvoeren van het correlatieonderzoek; en 2. Vervolgens nogmaals voor een controletest. Het beschikbaar stellen en het gebruik van de BRP-data worden verantwoord in een DPIA en maakt onderdeel uit van een ethische toetsing (IAMA). De definitieve DPIA wordt aan het einde van het ontwikkeltraject en voorafgaande van de ingebruikname van het algoritme afgerond. Indien nodig wordt een afzonderlijke gebruikersovereenkomst gesloten tussen de data-eigenaar binnen cluster Dienstverlening en de gebruiker Advanced Analytics van cluster BCO. Mochten er aanvullende beheermaatregelen nodig zijn, dan worden deze in samenspraak met betrokkenen uit de clusters DV en W&I en BCO (OBI) genomen. De adviezen van de Privacy Offices W&I en DV worden hierin meegenomen.



	Mochten ondanks de ingebouwde waarborgen risico's zijn betreffende het gebruik van de BRP-data, dan wordt direct geëscaleerd naar opdrachtgever (directeur uitvoering W&I), de data-eigenaar (DV) en teammanager team Advanced Analytics (BCO). Zo nodig wordt het correlatieonderzoek direct stopgezet.
Wat zijn de risico's die overblijven?	Niet van toepassing
Hoe groot is de impact voor de betrokkene?	Niet van toepassing
Hoe groot is de kans dat het risico zich voor doet?	Niet van toepassing
Aanvullende maatregelen?	Niet van toepassing

4.9 Informatieplicht

Norm 16	Betrokkenen dienen op het moment dat gegevens verzameld worden volledig te worden geïnformeerd over de verwerking van die gegevens.
Hoe wordt voldaan aan de norm?	Bij het indienen van een aanvraag voor een uitkering op basis van de Participatiewet wordt betrokkene erop gewezen dat zijn gegevens worden verzameld voor de uitvoering van de Participatiewet. Er is een privacyverklaring aanwezig op rotterdam.nl: Uw gegevens bij Werk en inkomen Rotterdam.nl. Hierin staat uitgelegd welke gegevens de gemeente verwerkt bij een uitkering, waarom de gemeente die gegevens nodig heeft en met welke instanties de gegevens eventueel gedeeld worden. Controleren en vaststellen van het recht op uitkering wordt benoemd. Ook wordt uitgelegd welke bronnen worden geraadpleegd en hoe de betrokkene zijn/haar rechten kan uitoefenen. Het selecteren op basis van een risico-inschattingsmodel wordt hier niet genoemd. De communicatie over dat uitkeringsgerechtigden mogelijk voor een heronderzoek worden uitgenodigd op basis van de uitkomsten van het algoritme, verloopt via het algoritmeregister op de website van de gemeente Rotterdam: Algoritmeregister Rotterdam.nl Onderhavige verwerking wordt opgenomen in het register van verwerkingen. Een extract daarvan wordt gepubliceerd op rotterdam.nl. Daarnaast wordt een samenvatting van onderhavige DPIA gepubliceerd op rotterdam.nl: Verwerkingsregister gemeente Rotterdam Rotterdam.nl
Wat zijn de risico's die overblijven?	Betrokkenen worden niet, te laat of onvolledig geïnformeerd over de verwerking van gegevens.
Hoe groot is de impact voor de betrokkene?	Laag



Hoe groot is de kans dat het risico zich voor- doet?	Laag
Aanvullende maatregelen?	1. Informeer de door het risico-inschattingmodel geselecteerde werkzoekende. 2. Vul de huidige privacyverklaring aan met info over het risico-inschattingmodel. 3. Zorg dat - vooruitlopend op de Europese Verordening AI/ WET OP DE ARTIFICIËLE INTELLIGENTIE - voldaan wordt aan artikel 13.

Norm 17	Indien de gegevens in de verwerking via een externe derde worden verkregen worden de betrokkenen uiterlijk een maand na het verkrijgen van de informatie op de hoogte gesteld. Indien niet voldaan wordt aan een van de uitzonderingen. De informatie die de betrokkene ontvangt voldoet aan Art. 14 van de AVG.
Hoe wordt voldaan aan de norm?	Niet van toepassing. De gegevens in de verwerking worden niet via een externe derde verkregen.
Wat zijn de risico's die overblijven?	Niet van toepassing
Hoe groot is de impact voor de betrokkene?	Niet van toepassing
Hoe groot is de kans dat het risico zich voor- doet?	Niet van toepassing
Aanvullende maatregelen?	Niet van toepassing

4.10 Rechten van betrokkenen

Norm 18	Indien een betrokkene een inzageverzoek doet, kan eenvoudig worden aan- gegeven welke gegevens van betrokkene worden verwerkt.
Hoe wordt voldaan aan de norm?	Betrokkenen kunnen hun AVG-rechten, via een vastgesteld concern brede procedure, bij de gemeente uitoefenen. Dit verloopt via het concern proces https://www.rotterdam.nl/loket/aanvraag-avg/ Het is mogelijk om inzage te geven in de gegevens die verwerkt worden door een extract van het register van verwerkingen te raadplegen/ overleggen. Daarnaast kunnen bestanden door daarvoor geautoriseerde en gemandateerde medewerkers worden doorgezocht op pseudo-ID en vervolgens kan inzage worden gegeven in: - De vraag of de burger is uitgenodigd voor een heronderzoek op basis van de uitkomsten van het algoritme of op basis van een andere selectiemethode; - Hoe hoog het risico-inschattinggetal voor de individuele burger is.



	- De in het model gebruikte variabelen en features met bijbehorende relatieve belangrijkheid.
Wat zijn de risico's die overblijven?	Geen
Hoe groot is de impact voor de betrokkene?	Niet van toepassing
Hoe groot is de kans dat het risico zich voor- doet?	Niet van toepassing
Aanvullende maatregelen?	Niet van toepassing

Norm 19	Geautomatiseerde besluitvorming, zonder menselijke controle is niet toegestaan indien dit wordt toegepast om het gedrag, de aanwezigheid of de prestaties van mensen, persoonlijke voorkeuren, interesses, locatie etc in kaart te brengen en/of te beoordelen en/of te voorspellen.
Hoe wordt voldaan aan de norm?	Er vindt geen geautomatiseerde besluitvorming plaats. De risicoscores zijn een handvat om te bepalen wie voor een heronderzoek wordt uitgenodigd. Het onderzoek vindt plaats door een mens en de wijze waarop het onderzoek plaatsvindt is onafhankelijk van de wijze waarop de burger voor een heronderzoek is geselecteerd. Het besluit over de rechtmatigheid van de verstrekte bijstand en eventuele aanpassingen in de hoogte van de uitkering wordt altijd door een medewerker voorbereid of genomen. Er is wel sprake van 'profilering'. Het gebruik van het risico-inschattingmodel 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 risico-inschattingmodel is namelijk een algoritme dat de kans op rechtmatige of onrechtmatig verstrekking van bijstand inschat voor de huidige uitkeringsgerechtigden. Ten tweede heeft de verwerking betrekking op persoonsgegevens. De in het model gebruikte indicatoren zijn persoonsgegevens uit de uitkeringssystemen Socrates en RMW/RAAK. De informatie die gebruikt wordt voor de indicatoren in het risico-inschattingmodel zijn persoonsgegevens als bedoeld in artikel 4, aanhef en onder 1, van de AVG. Ten derde heeft de verwerking met het risico-inschattingmodel ten doel om individuele kenmerken te evalueren en categoriseren. Aan de hand van persoonlijke aspecten van de uitkeringsgerechtigde wordt geëvalueerd en op basis daarvan ingeschat hoe groot het risico is dat de uitkeringsgerechtigde een (on)rechtmatige bijstandsuitkering geniet en wordt een (uitnodiging voor) een heronderzoek daarop afgestemd. Het algoritme is alleen bepalend voor de vraag wie voor een heronderzoek zou moeten worden uitgenodigd. Het algoritme neemt zelfstandig geen geautomatiseerde besluiten.



Wat zijn de risico's die overblijven?	Geen
Hoe groot is de impact voor de betrokkene?	Niet van toepassing.
Hoe groot is de kans dat het risico zich voor- doet?	Niet van toepassing
Aanvullende maatregelen?	Niet van toepassing

Norm 20	Indien een betrokkene hier een verzoek toe doet, kunnen gegevens eenvoudig worden verwijderd (mits er geen bewaarplicht geldt).
Hoe wordt voldaan aan de norm?	Dit verloopt via het concern proces https://www.rotterdam.nl/loket/aanvraag-avg/ De bestanden kunnen handmatig worden doorgezocht op pseudo-ID en op verzoek en indien wettelijk toegestaan worden verwijderd.
Wat zijn de risico's die overblijven?	Niet van toepassing
Hoe groot is de impact voor de betrokkene?	Niet van toepassing
Hoe groot is de kans dat het risico zich voor- doet?	Niet van toepassing
Aanvullende maatregelen?	Niet van toepassing

Norm 21	Indien een betrokkene hier een verzoek toe doet, kunnen gegevens eenvoudig worden aangepast (mits er geen juridische bezwaren zijn).
Hoe wordt voldaan aan de norm?	Dit verloopt via het concern proces https://www.rotterdam.nl/loket/aanvraag-avg/ De bestanden kunnen handmatig worden doorgezocht op pseudo-ID en op verzoek kunnen gegevens worden aangepast mits er geen juridische bezwaren zijn.
Wat zijn de risico's die overblijven?	Niet van toepassing
Hoe groot is de impact voor de betrokkene?	Niet van toepassing
Hoe groot is de kans dat het risico zich voor- doet?	Niet van toepassing
Aanvullende maatregelen?	Niet van toepassing

4.11 Training en instructie van medewerkers



Norm 22	Medewerkers zijn op de hoogte hoe om te gaan met persoonsgegevens binnen de verwerking.
Hoe wordt voldaan aan de norm?	Medewerkers zijn zich bewust van het feit dat ze zorgvuldig om moeten gaan met persoonsgegevens. Er is periodiek aandacht voor privacy en beveiliging binnen de organisatie. Vertrouwelijkheid van gegevens is een onderdeel van de ambtseed. Medewerkers van de gemeente volgen een verplichte e-learning over privacy, informatiebeveiliging, informatiebeheer en integriteit. N.a.v. datalekken worden evaluatiegesprekken gevoerd in werkoverleggen. De verplichte e-learning is door ongeveer 77% van de medewerkers daadwerkelijk gevolgd. Er is blijvende aandacht voor periodieke trainingen en bewustwording van medewerkers met betrekking tot privacy-, informatiebeveiliging, informatiebeheer en integriteit. Specifiek voor de afdeling OBI geldt dat medewerkers worden geïnformeerd en getraind over de wijze hoe zij om moeten gaan met persoonsgegevens tijdens uitvoering van hun werkzaamheden. Dit omvat ook de werkinstructies. Verzoeken vanuit clusters aan de afdeling OBI, worden na akkoord van de cluster business consultant, voor raadpleging voorgelegd aan de Privacy-Officer van het betreffende cluster.
Wat zijn de risico's die overblijven?	Geen
Hoe groot is de impact voor de betrokkene?	Niet van toepassing
Hoe groot is de kans dat het risico zich voordoet?	Niet van toepassing
Aanvullende maatregelen?	Niet van toepassing



4 Risico Analyse beveiligingsmaatregelen

Controleer voor deze paragraaf bij de DISO of er een Risico Impact Assessment (RIA) is uitgevoerd op processen waaraan de verwerking is gerelateerd of applicaties die bij de verwerking worden gebruikt. Onderstaande vragen kunnen aan de hand van de uitgevoerde RIA worden beantwoord. Is er geen RIA uitgevoerd, betrek dan de DISO bij de beantwoording van deze vragen.

Het algoritme wordt ontwikkeld op het Azure DAP'R platform. Hierop is in 2021 en in mei 2022 een DPIA ('Consolideren en bewaren van gegevens in het Datalake') uitgevoerd. Ook op concern Datawarehouse is in 18 juni 2021 een DPIA ('Centrale datawarehouse') uitgevoerd. Voor de analyse op hoofdlijnen van de risico's en maatregelen op het gebied van informatiebeveiliging wordt verwezen naar de DPIA van Datalake en Concern Data Warehouse. De vastgestelde beheersmaatregelen worden vanuit het project SPIDER (Cluster BCO) geïmplementeerd.

Voor details over Datalake wordt verwezen naar de BIO-self assessment die in het kader van de migratie van het Datalake naar Microsoft Azure (eind 2021) is uitgevoerd. Zie voor de resultaten en voor de status van de implementatie van maatregelen <https://wiki.rotterdam.nl/display/DAPR/BIO+uitwerking+DAP%27R+op+Azure>.

De bronapplicaties Socrates en RMW/ RAAK vallen buiten de scope van onderhavige DPIA.

4.1 Fysieke beveiliging

Norm 23	Persoonsgegevens moeten adequaat worden beschermd. Dit geldt zowel voor de elektronische beveiliging (logische toegangsbeveiliging, encryptie) als fysieke beveiliging (diefstal informatie, niet afgesloten dossierkasten). IT-voorzieningen en apparatuur zijn fysiek beschermd tegen toegang door onbevoegden en tegen schade en storingen. De geboden bescherming is in overeenstemming met de vastgestelde risico's.
Hoe wordt voldaan aan de norm?	Aan de hand van policies, in dit kader technische regels, wordt er afgedwongen dat er alleen Azure resources gebruikt mogen worden die staan in datacenters in West-Europa. De Azure DAP'R omgeving is ingericht zonder public endpoints. Er is ook een directe koppeling met het Rotterdamse netwerk gelegd, waardoor de omgeving volledig is afgeschermd van het publieke internet. Toegang tot DAP'R is alleen mogelijk indien de medewerker voorkomt in de Active Directory van Rotterdam en waarbij beheer expliciet autorisatie heeft toegekend. Zie norm 24. Microsoft heeft ook geen toegang tot onze data. Zie https://docs.microsoft.com/en-us/azure/security/fundamentals/protection-customer-data.



	Fysieke beveiligingszone Voor elke locatie van de gemeente wordt een risicoanalyse en beveiligingsplan gemaakt, het zogeheten pandprofiel. Een pandprofiel kan ook voor een groep locaties worden gemaakt, bijv. als ze dicht bij elkaar staan of op dezelfde manier worden gebruikt. In het pandprofiel wordt de locatie verdeeld in één of meerdere zones. Het is niet noodzakelijk dat een locatie altijd alle zones heeft. Kantoren, ruimten en faciliteiten beveiligen Maatregelen voor beveiliging van ruimtes zijn genomen naar aanleiding van de zonering zoals die in het pandprofiel is vastgelegd. Daaronder valt ook het omgaan met informatie op papier en met waardevolle goederen. In het fysieke beveiligingsbeleid is aangegeven hoe wordt omgegaan met ontvangst van bezoekers. Iedere medewerker heeft een toegangspas, die nodig is om toegang te krijgen tot de gebouwen/ lokalen van de gemeente. Laptops en computers zijn beveiligd met key-locker en inloggen, en zijn op afstand te wissen bij verlies of diefstal. De Citrix-omgeving heeft 2FA. Iedere te ontsluiten bron heeft een eigen voorportaal, met eigen inlog. De databases zijn gefragmenteerd, dus toegang tot een database geeft niet automatisch toegang tot andere databases.
Wat zijn de risico's die overblijven?	Door de genomen maatregelen is de kans op de genoemde risico's geminimaliseerd.
Hoe groot is de impact voor de betrokkene?	N.v.t
Hoe groot is de kans dat het risico zich voordoet?	N.v.t.
Aanvullende maatregelen?	N.v.t.

4.2 Logische toegangsbeveiliging

Norm 24	Alleen bevoegde gebruikers mogen toegang hebben tot de informatiesystemen en -diensten die ze voor de uitvoering van hun taken nodig hebben en om onbevoegde toegang tot informatiesystemen te voorkomen.
Hoe wordt voldaan aan de norm?	Beveiligingsmaatregelen Authenticatie Voor de toegangscontrole tot de Rotterdamse werkplek en thuiswerkvoorziening wordt vertrouwd op de standaardbeveiliging die de gemeente aanbiedt. Op dit moment wordt veel thuisgewerkt, via de reguliere Rotterdamse manier van thuiswerken. Beide betreffen generieke voorzieningen en zijn daarmee geen onderdeel van deze DPIA.



	Toegang tot DAP'R is alleen mogelijk vanaf een workstation dat deel uitmaakt van het Rotterdamse kantoor netwerk, of via de Citrixomgeving. Om daadwerkelijke toegang te krijgen tot de data vindt een tweede authenticatie plaats. Beveiligingsmaatregelen Autorisatie Toegang tot DAP'R Azure is ten eerste alleen mogelijk als de account bekend is in de Rotterdamse Active Directory. Er wordt dan ook niet gewerkt met groepsaccounts. De basis van autorisatie op Azure wordt gevormd door beveiligingsgroepen en aangelopen vanuit twee richtingen: • Enerzijds vindt er autorisatie plaats op het gebied van toegang tot Azure applicaties (tools) en workspaces (onderdelen) en • Anderzijds autorisatie voor toegang tot het datalake (data). Iemand kan bijv. toegang hebben tot een Azure applicatie, maar zonder expliciet toegang tot het Datalake kan er geen data worden ingezien. Het uitgangspunt met betrekking tot autorisaties op het Datalake is dat niemand automatisch toegang heeft tot data. Toegangsverlening tot data is altijd een bewuste handeling van Functioneel Beheer. Er wordt binnen DAP'R Azure onderscheid gemaakt tussen "Standard" en "Elevated" rechten, waarbij bij "Elevated" er sprake is van autorisatie tot privacy en vertrouwelijke data. Er is een speciale procedure voor het toekennen van "Elevated" rechten, aangezien beperkte medewerkers met expliciete reden hier toegang toe mogen hebben. De verleende toegang, welke wordt vastgelegd in de autorisatiematrix, wordt elk kwartaal geëvalueerd. Het plan is om 2023 te gaan werken met een soort vergunningsmechanisme. Je krijgt een vergunning voor toegang tot data n.a.v. een gesteld doel (GLO/DUB-advies). Nadat de vergunning is afgelopen, worden de rechten automatisch ingetrokken.
Wat zijn de risico's die overblijven?	Onrechtmatige autorisatie: er is een handmatige procedure die gevolgd wordt bij functiewijziging of uitdiensttreding. Het bijhouden van autorisaties vereist daarom alertheid van de functioneel beheerders van het Datalake.
Hoe groot is de impact voor de betrokkene?	Hoog



Hoe groot is de kans dat het risico zich voordoet?	Middel
Aanvullende maatregelen?	Toets periodiek de toegang tot applicatie (users én de aan hen toegekende toegangsrechten).

4.3 Logging

Norm 25	Activiteiten die gebruikers en beheerders uitvoeren met persoonsgegevens worden vastgelegd in logbestanden⁵. Hetzelfde geldt voor andere relevante gebeurtenissen, zoals pogingen om ongeautoriseerd toegang te krijgen tot persoonsgegevens en verstoringen die kunnen leiden tot verminking of verlies van persoonsgegevens. De logbestanden worden periodiek gecontroleerd op indicaties van onrechtmatige toegang of onrechtmatig gebruik van de persoonsgegevens en waar nodig wordt actie ondernomen.
Hoe wordt voldaan aan de norm?	Er vindt bewaking plaats op ongeautoriseerde toegang. Op informatiebeveiligingsincidenten wordt gereageerd conform protocol 'meldplicht datalekken' (zie ook norm 27). Indien het om een echt noodgeval gaat, een grote actieve aanval, dan wordt er direct geschakeld met de SOC-afdeling. Er vindt gebruikerslogging plaats op alle relevante acties op data in Azure. Omdat ook uitgevoerde query's worden gelogd, is naderhand te reconstrueren wie wanneer toegang heeft gehad tot welke gegevens. De logging is gekoppeld aan persoonsgegevens (naam medewerker). Bewaartermijnen op logbestanden is daarom hier ook van toepassing.
Wat zijn de risico's die overblijven?	Geen
Hoe groot is de impact voor de betrokkene?	Niet van toepassing
Hoe groot is de kans dat het risico zich voordoet?	Niet van toepassing
Aanvullende maatregelen?	Niet van toepassing

4.4 Beheer

Artikel 5.2 lid 1 Woo



Norm 26	Software, zoals browsers, virusscanners en operating systems, wordt up-to-date gehouden. Ook installeert de verantwoordelijke tijdig oplossingen die de leverancier uitbrengt voor beveiligingslekken in deze software. Meer in het algemeen verkrijgt de verantwoordelijke tijdig informatie over technische kwetsbaarheden van de gebruikte informatiesystemen. De mate waarin de organisatie blootstaat aan dergelijke kwetsbaarheden wordt geëvalueerd en de verantwoordelijke treft geschikte maatregelen voor de behandeling van de risico's die daarmee samenhangen.
Hoe wordt voldaan aan de norm?	In het geval van Azure DAP'R signaleert KPN Inspark, cloud partner van de gemeente Rotterdam, of er updates van het platform beschikbaar dan wel nodig zijn. Functioneel Beheer geeft vervolgens de opdracht om de betreffende update(s) te installeren op de DAP'R omgeving. In de techniek zijn vele maatregelen getroffen om uitbuiting van kwetsbaarheden te voorkomen, o.a. is er een SOC. Beheer is een regulier proces. De mate waarin de organisatie en DAP'R blootstaat aan kwetsbaarheden wordt conform beleid getoetst door het uitvoeren van een pen test. De eerste pen test voor Azure staat voor dit jaar gepland. De bevindingen hiervan worden geëvalueerd en verantwoordelijke treft geschikte maatregelen om risico's te reduceren of weg te nemen.
Wat zijn de risico's die overblijven?	Verlies van data
Hoe groot is de impact voor de betrokkene?	Hoog: ingeval de sleutel gesloten wordt en de gepseudonimiseerde gegevens daarmee ontsleuteld kunnen worden.
Hoe groot is de kans dat het risico zich voordoet?	Laag
Aanvullende maatregelen?	1. Periodiek pen testen overwegen en uitvoeren. 2. RIA uitvoeren op Azure DAP'R en overige betrokken ict-systemen die ingezet worden voor het ontwikkelen van het algoritme.

4.5 Incident management

Norm 27	Er zijn procedures voor het tijdig en doeltreffend behandelen van informatie-beveiligingsincidenten en zwakke plekken in de beveiliging, zodra ze zijn gerapporteerd. Het beoordelen van de risico's voor de betrokkenen en het effectief informeren van de betrokkenen en, waar van toepassing, de toezichhouder is in deze procedures opgenomen. De lessen getrokken uit de afgehandelde incidenten worden gebruikt om de beveiliging waar mogelijk structureel te verbeteren. Als een vervolgprocedure na een informatiebeveiligingsincident juridische maatregelen omvat (civiel- of strafrechtelijk), wordt het bewijsmateriaal verzameld, bewaard en gepresenteerd overeenkomstig de voorschriften voor bewijs die voor het relevante rechtsgebied zijn vastgelegd.
---------	--



Hoe wordt voldaan aan de norm?	Meerdere partijen spelen hierbij een rol. Functioneel Beheer DAP'R vervult de centrale rol en draagt zorg voor het effectief informeren van betrokkenen en het opvolgen van processen. Er bestaat binnen de gemeente een standaardproces voor de melding en opvolging van incidenten en beveiligingslekken. (Protocol melding en afhandeling datalekken 23 mei 2018.pdf) Er vindt gebruikerslogging plaats op alle relevante activiteiten op Azure. Het is daarmee te achterhalen wie wat heeft bevraagd t.b.v. bewijsvoering. (Zie norm 25)
Wat zijn de risico's die overblijven?	• Het niet herkennen van een (mogelijk) datalek; • het niet (correct) volgen van de procedures door medewerkers.
Hoe groot is de impact voor de betrokkene?	Hoog
Hoe groot is de kans dat het risico zich voordoet?	Laag
Aanvullende maatregelen?	Doorlopend aandacht schenken aan kennis onder medewerkers ten behoeve van het herkennen van datalekken en het in voorkomende gevallen adequaat handelen.

4.6 Continuïteit

Norm 28	Door natuurrampen, ongevallen, uitval van apparatuur of opzettelijk handelen kunnen persoonsgegevens verloren gaan. Door in de organisatie continuïteitsbeheer in te richten worden de gevolgen tot een aanvaardbaar niveau beperkt, waarbij gebruik wordt gemaakt van een combinatie van preventieve maatregelen en herstelmaatregelen.
Hoe wordt voldaan aan de norm?	Azure heeft een compleet overzicht beschikbaar met SLA's per resource uitgeschreven: https://azure.microsoft.com/nl-nl/support/legal/sla/summary/. Het datalake voldoet aan LRS (Lokaal Redundant Opslag) voor de development en test omgeving. Dit houdt in dat we 3 kopieën hebben die bij upload/verwijdering/wijziging synchroon worden doorgevoerd. Dit gebeurt op 1 fysieke locatie. Voor acceptatie en productie, wordt er gebruik gemaakt van ZRS (Zone Redundant Opslag); opslag op meerdere fysieke locaties. Er wordt voldaan aan de eisen: maximaal toegestane dataverlies van 28 uur en maximale hersteltijd na een incident van 16 werkuren.
Wat zijn de risico's die overblijven?	Geen



Hoe groot is de impact voor de betrokkene?	Niet van toepassing
Hoe groot is de kans dat het risico zich voordoet?	Niet van toepassing
Aanvullende maatregelen?	Niet van toepassing

4.7 Verwerker

Meer informatie over verwerkers vind je op: <https://rio.rotterdam.nl/umbraco/organisatie/privacy/>

Norm 29	Indien (een deel van) de verwerking bij een andere partij plaatsvindt dient er een verwerkersovereenkomst te worden afgesloten en dient er toezicht te zijn op de verwerking van gegevens.
Hoe wordt voldaan aan de norm?	1. Inspark Managed Oxygen is verwerker. Een verwerkersovereenkomst is aanwezig. 2. Microsoft Azure: er is geen 'aparte' verwerkersovereenkomst afgesloten met Microsoft voor Azure. Onduidelijk is of de afspraken Microsoft gecheckt en afdoende zijn.
Wat zijn de risico's die overblijven?	Verlies van data en verlies van PKI-sleutels
Hoe groot is de impact voor de betrokkene?	Midden
Hoe groot is de kans dat het risico zich voordoet?	Midden
Aanvullende maatregelen?	Zoek uit of de afspraken met Microsoft i.v.m. Azure voor het ontwikkelen van het algoritme passend zijn.



5 Overige risico's

Norm. 30 Zijn er andere hoge risico's met betrekking tot de verwerking van persoonsgegevens die nog niet genoemd zijn?	Risico op onzorgvuldige/ onbehoorlijke/ onrechtmatige/ onverantwoorde verwerking: In deze tijd (toeslagenaffaire, Syri en institutionele racisme in zijn algemeenheid) ligt het gebruik van profielen in combinatie met algoritmes zeer gevoelig en is men zich meer bewust dan voorheen van de aanwezigheid van 'risico op onzorgvuldige/ onbehoorlijke/ onrechtmatige/ onverantwoorde verwerking' ondanks de getroffen en de te treffen maatregelen. Er zijn onzekerheden aanwezig, zoals het ontbreken van specifieke AI-wetgeving, het ontbreken van een Rotterdams algoritmebeleid/ proces en de implementatie van de algoritme governance is nog gaande.
Hoe groot is de impact voor de betrokkene?	Hoog
Hoe groot is de kans dat het risico zich voordoet?	Midden
Aanvullende maatregelen?	1. Ga na of het ontwikkelen van het nieuwe model nu wel het juiste moment is. 2. Laat het risico-inschattingsmodel toetsen door een onafhankelijk algoritme toetsingscomité. 3. Laat de PO en FG gezamenlijk beoordelen of onderhavige verwerking voorgesteld moet worden aan de Autoriteit Persoonsgegevens voor een ' voorafgaande raadpleging '. Zo ja, start het aanvraagproces.

Norm. 31 Zijn er andere hoge risico's met betrekking tot de verwerking van persoonsgegevens die nog niet genoemd zijn?	Betrokkenen: mogelijk draagvlak ontbreekt Betrokkenen zijn niet direct betrokken bij de ontwikkeling van het algoritme, wel indirect via de Cliëntenraad en gemeenteraad. Beide gremia worden proactief betrokken bij de ontwikkeling van het algoritme. Overige externe stakeholders worden geconsulteerd.
Hoe groot is de impact voor de betrokkene?	Laag
Hoe groot is de kans dat het risico zich voordoet?	Midden
Aanvullende maatregelen?	Blijf betrokkenen bij de realisatie de ontwikkeling van het algoritme-inschattingsmodel betrekken bi externe consultatiestrategie en het proactief informeren van de cliëntenraad.



6 Overzicht maatregelen op basis van de risico's

Hieronder worden de beoogde maatregelen, om de restrisico's te mitigeren, geschetst. Deze zijn in volgorde van de besproken beginselen en niet op volgorde van zwaarte.

Nr.	Risico omschrijving	Kans	Impact	Risico =K x I	Risico ac- cep- teren	Maatregel(en)
1.	Discriminatie van een werkzoekende. (norm 1)	1	3	3	Nee	1. Voer de IAMA op het nieuwe algoritmemodel z.s.m. uit en de daarin voorgesteld processtappen (en documenteer). 2. Zorg dat het proces 'Ontwikkelen algoritme voor heronderzoeken' aansluit met nieuwe ontwikkelingen rondom AI-wetgeving, zoals de AI-verordening. 3 a. Voer alle processtappen conform het proces 'Ontwikkelen algoritme voor heronderzoeken' uit (en documenteer). 3 b. Controleer periodiek of alle processtappen conform het proces 'Ontwikkelen algoritme voor heronderzoeken' zijn gevolgd. 4. Laat het risico-inschattingmodel toetsen door een onafhankelijk algoritme toetsingscomité.
2.	Financieel verlies werkzoekende (terugwerkende kracht + toekomstige uitkeringen) (norm 1)	1	3	3	Nee	Zie maatregelen bij nummer 1.
3.	Reputatieschade werkzoekende. (norm 1)	1	3	3	Nee	Zie maatregelen bij nummer 1.
4.	Onrechtmatige verwerking als ondanks de getroffen waarborgen het risico-inschattingmodel onbewust discriminerend werkt, waardoor bepaalde groepen/ werkzoekenden sneller uitgenodigd worden voor een heronderzoek dan andere. (norm 1)	1	3	3	Nee	Zie maatregelen bij nummer 1.



5.	Reputatieschade, verlies van vertrouwen, sancties toezichthouders, klachten en gerechtelijke procedures, financiële schade. (norm 1)	1	3	3	Nee	Zie maatregelen bij nummer 1.
6.	Verlies controle over eigen gegevens bij onrechtmatige verwerkingen (normen 4, 9 en 10)	1	3	3	Nee	1. Onderbouw waarom en welke bijzondere persoonsgegevens noodzakelijk zijn voor het bereiken van het doel en verwerk geen bijzondere persoonsgegevens in het algoritme als deze niet noodzakelijk zijn voor het bereiken van het doel. 2. Voer de processtappen in de Procesbeschrijving 'Ontwikkelen algoritme voor heronderzoeken' en de processtappen conform de 'IAMA' uit (en documenteer) om continu risico's te mitigeren.
7.	Bias (onwenselijke systematische afwijkingen van de uitkomsten van het algoritme voor de gehele groep of juist voor specifieke personen of deelgroepen) (norm 8)	1	3	3	Nee	Voer periodiek controles uit (tijdens ontwerp en uitvoering) om de juiste werking van het algoritme te garanderen.
8.	Het gebruik van (resterende) variabelen die bestaan uit vrij invulbare velden, vergroten de kans op subjectiviteit en mogelijke bias. Zie verder norm 1. (norm 11)	1	2	2	Nee	1. Onderzoek periodiek (tijdens ontwerp en uitvoering) het risico-inschattingmodel en de toepassing daarvan op discriminatie en bias, zodat deze binnen de grenzen van het discriminatieverbod blijven en voorkom daarmee dat onwenselijke systematische afwijkingen voor specifieke werkzoekenden of groepen ontstaan. 2. Onderzoek, onderbouw en leg vast (periodiek) de noodzakelijkheid van de persoonsgegevens (historische data) die als input dienen voor het risico-inschattingmodel. 3. Laat een onafhankelijk algoritme toetsingscomité het risico-inschattingmodel toetsen.
9.	Geen of niet tijdige vernietiging met als gevolg dat de data in handen komt van onbevoegden. (normen 12 en 13)	1	1	1	Nee	Controleer of daadwerkelijk na afloop van de bewaartermijn de betreffende data is vernietigd.



10.	Betrokkenen worden niet, te laat of onvolledig geïnformeerd over de verwerking van gegevens. (norm 16)	1	1	1	Nee	1. Informeer de op basis van het risico-inschattingmodel geselecteerde werkzoekende. We kunnen algemene informatie geven aan alle uitkeringsgerechtigden die voor een heronderzoek worden uitgenodigd. Niet voor alleen uitkeringsgerechtigden die specifiek worden uitgenodigd o.b.v. het algoritme, maar breder naar de hele doelgroep. 2. Vul de huidige privacyverklaring aan met info over het risico-inschattingmodel. 3. Zorg dat - vooruitlopend op de Europese Verordening AI/ WET OP DE ARTIFICIËLE INTELLIGENTIE - voldaan wordt aan artikel 13.
11.	Onrechtmatige autorisatie: er is een handmatige procedure die gevolgd wordt bij functiewijziging of uitdiensttreding. Het bijhouden van autorisaties vereist daarom alertheid van de functioneel beheerders van het Datalake. (norm 24)	1	2	2	Nee	Toets periodiek de toegang tot applicatie (users én de aan hen toegekende toegangsrechten).
12.	Verlies van data (norm 26)	1	2	2	Ja	1. Periodiek pen testen overwegen en uitvoeren. 2. RIA uitvoeren op Azure DAP'R en overige betrokken ict-systemen die ingezet worden voor het ontwikkelen van het algoritme.
13.	Het niet herkennen van een (mogelijk) datalek. (norm 27).	1	2	2	Nee	Doorlopend aandacht schenken aan kennis onder medewerkers ten behoeve van het herkennen van datalekken en het in voorkomende gevallen adequaat handelen.
14.	Het niet (correct) volgen van de procedures door medewerkers. (norm 27)	1	2	2	Nee	Doorlopend aandacht schenken aan kennis onder medewerkers ten behoeve van het herkennen van datalekken en het in voorkomende gevallen adequaat handelen.
15.	Verlies van data en verlies van PKI-sleutels (norm 29)	1	2	2	Nee	Zoek uit of de afspraken met Microsoft i.v.m. Azure voor het ontwikkelen van het algoritme passend zijn.
16.	In deze tijd (toeslagenaffaire, Syri en institutionele racisme in zijn algemeenheid) ligt het gebruik van profielen in combinatie met algoritmes zeer gevoelig en is	1	2	2	Nee	1. Ga na of het ontwikkelen van het nieuwe model nu wel het juiste moment is. 2. Laat het risico-inschattingmodel toetsen door een onafhankelijk algoritme toetsingscomité. 3. Laat de PO en FG gezamenlijk beoordelen of onderhavige verwerking voorgelegd moet worden aan



	men zich meer bewust dan voorheen van de aanwezigheid van 'risico op onzorgvuldige/ onbehoorlijke/ onrechtmatige/ onverantwoorde verwerking' ondanks de getroffen en de te treffen maatregelen. Er zijn onzekerheden aanwezig, zoals het ontbreken van specifieke AI-wetgeving, het ontbreken van een Rotterdams algoritmebeleid/ proces en de implementatie van de algoritme governance is nog gaande. (norm 30)					de Autoriteit Persoonsgegevens voor een 'voorafgaande raadpleging'. Zo ja, start het aanvraagproces.
17.	Betrokkenen zijn niet direct betrokken bij de ontwikkeling van het algoritme, wel indirect via de Cliëntenraad en gemeenteraad. Beide gremia worden proactief betrokken bij de ontwikkeling van het algoritme. Overige externe stakeholders worden geconsulteerd. (norm 31)	1	2	2	Nee	Blijf betrokkenen bij de realisatie de ontwikkeling van het algoritme-inschattingmodel betrekken bij externe consultatiestrategie en het proactief informeren van de cliëntenraad.



7 Planning maatregelen

Risico nr.	Proces-verantwoordelijke	Uitvoering door	Datum gereed
1	Directeur Uitvoering portefeuille Inkomen Annemarie de Rotte	Projectleider [REDACTED]	Q1 2023
2	Annemarie de Rotte	[REDACTED]	Q1 2023
3	Annemarie de Rotte	[REDACTED]	Q1 2023
4	Annemarie de Rotte	[REDACTED]	Q1 2023
5	Annemarie de Rotte	[REDACTED]	Q1 2023
6	Annemarie de Rotte	[REDACTED]	Q1 2023
7	Teamleider BCO OBI Advanced Analytics Projectleider Spider alytics	Data-scientist [REDACTED]	Q1 2023
8	Annemarie de Rotte	[REDACTED]	Q1 2023
9	Annemarie de Rotte	[REDACTED]	2033, 10 jaar na in gebruikname algoritme
10	Annemarie de Rotte	[REDACTED]	Q1 2023
11	DAP'R Consultant / Technical Lead [REDACTED]	DISO BCO	Q1 2023
12	[REDACTED]	DISO BCO	Q1 2023
13	[REDACTED]	Projectleider Spider	Q1 2023
14	[REDACTED]	Projectleider Spider	Q1 2023
15	[REDACTED]	Projectleider Spider	Q1 2023
16	Annemarie de Rotte	[REDACTED]	Q1 2023
17	Annemarie de Rotte	[REDACTED]	Q1 2023



8 Akkoord

De proceseigenaar is ervoor verantwoordelijk dat voor verwerkingen met een hoog risico privacy impact assessments worden uitgevoerd.

Om ervoor te zorgen dat eigenaarschap in de lijn wordt belegd wordt de inhoud van de DPIA, de bevindingen en de mogelijke maatregelen afgestemd met betrokkenen uit de lijn. Daarnaast wordt aan de Proceseigenaar gevraagd akkoord te geven op het rapport en de verbetervoorstellen.

De volgorde van akkoord is als volgt:

1. Deelnemers
2. Privacy Officer
3. Decentraal Information Security Officer
4. Proceseigenaar
5. Functionaris Gegevensbescherming

Rol	Naam	Datum akkoord	Akkoord via
Deelnemer workshop/ input	[REDACTED]		E-mail
Deelnemer workshop/ input	[REDACTED]		E-mail
Deelnemer workshop/ input	[REDACTED]		E-mail
Privacy Officer	[REDACTED]		E-mail
Decentrale Information Security Officer (DISO)	[REDACTED]		E-mail
Proceseigenaar W&I	Annemarie de Rotte		E-mail
Teamleider Data Science van afdeling OBI, cluster BCO: verantwoordelijke voor de ontwikkeling en training van het model.	Vacant		E-mail
Teamleider Data Warehouse van afdeling OBI, cluster BCO: verantwoordelijk voor de verkrijging, opslag en bewerking van data binnen Concern Data Warehouse en Datalake Azure	[REDACTED]		E-mail
Functionaris Gegevensbescherming	[REDACTED]		E-mail

Akkoord kan gegeven worden per e-mail. Indien deze mail opmerkingen bevat worden deze overgenomen en wordt het document daarmee als akkoord verondersteld.

Stakeholders geven akkoord voor



1. Deelnemer workshop – geeft akkoord voor inhoud en verslaglegging
2. Privacy Officer – geeft akkoord op diepgang en acties
3. DISO – geeft akkoord op diepgang en acties
4. Proceseigenaar – geeft akkoord voor verslag en accepteert acties
5. Functionaris Gegevensbescherming – geeft akkoord op de inhoud



9 Bijlage - Bronnen

A.1. Geïnterviewden

Datum	Geïnterviewde	Onderwerp

A.2. Documentatie

Nr	Documentnaam
1	Bijlage Proces 'Ontwikkel algoritme voor heronderzoeken'
2	Bijlage JD-advies BRP
3	Bijlage Deelproces 1 'vaststellen aantal heronderzoeken en aanpak'
4	Bijlage 'Second opinion adviesaanvraag JD'
5	Bijlage 'Correlatie covar en features' d.d. 12.12.2022
6	Bijlage Beantwoording van de schriftelijke vragen van het raadslid A.J. van Zevenbergen (SP) over 'Gemeentelijke algoritmes onder de loep' (d.d. 14-12-2021)
7	Bijlage Rapportage Handhaving cluster Werk en inkomen 2021