Gemeente Amsterdam

Kunstmatige Intelligentie Impact Assessment

Algoritme:	Vakantieverhuur algoritme Alpha
Bestuurlijk verant -woordelijke:	Wethouder Laurens Ivens
Directie:	Wonen
Verwerkings-verantwoordelijke:	Henriëtte Rombouts
Bronhouder	Dignar van Onzenoord (Directie Basisinformatie)
Afdeling:	Woninggebruik en verbetering
Afdelingshoofd:	10.2.e
Werkproces:	Illegale vakantieverhuur
Proceseigenaar:	10.2.e
Opgesteld door:	10.2.e (Product Owner / Teamleider), 10.2.e (Handhavingsjurist), 10.2.e (Privacy Officer Cluster Ruimte & Economie), 10.2.e (Privacy Jurist CTO Innovatieteam), 10.2.e en 10.2.e (Data scientist CTO ontwikkelteam).
Datum:	1 september 2021
Versie	1.5

Bijlagen:

1) Overzicht rollen en verantwoordelijkheden
2) DPIA Toezicht en Handhaving
3) Excelbestand met overzichten van:
   - Verwerkte data
   - Opgestelde features
   - Business rules betreffende werkproces
   - Achteraf verwijderde features
4) BIO-toets
5) Weging betrokken (grond)rechten vakantieverhuurmodel
6) Architectuurnotitie
7) Rapport bias analyse
8) Technische documentatie

Met opmerkingen 10.2.e **]:** Klopt niet meer.

Inhoudsopgave

1. Inleiding p.
2. Bepalen noodzaak uitvoeren Kunstmatige Intelligentie Impact Assessment. p.
3. KI-toepassing, verantwoordelijkheden en werkproces p.
4. Beschrijf de baten van de toepassing p
5. Architectuur p.
6. Data en data governance
7. Wettige en ethische toepassing p.
8. Bias analyse p.
9. Technische documentatie p.
10. Risicomanagement p.
11. Informatiebeveiliging p
12. Beheer p
13. Communicatie p.
14. Verantwoording p.
15. Afweging en beoordeling p.

1. Inleiding

In november 2019 presenteerde de Rekenkamer Metropool Amsterdam het rapport "Handhaving Vakantieverhuur". In dit rapport worden aanbevelingen gedaan om de doeltreffendheid en doelmatigheid van de handhaving op vakantieverhuur te verbeteren. Dit kan door alleen urgente meldingen en signalen uit digitaal toezicht te onderzoeken. Handhaving Wonen heeft al langer de ambitie om data gestuurd te handhaven. In 2019 is gewerkt aan de ontwikkeling van een machine learning algoritme voor Toezicht en Handhaving van Wonen. Dit algoritme is een vorm van kunstmatige intelligentie en geeft een risico indicatie. De inzet ervan moet leiden tot slimmer handhaven. Met ondersteuning van het algoritme kunnen medewerkers efficiënter en effectiever handhaven. Er is geen sprake van automatische besluitvorming door het algoritme.

Het ontwikkelde algoritme verwerkt een beperkte set persoonsgegeven op grote schaal. Zie bijlage 3. Er zijn daarom een aantal assessments uitgevoerd om de risico's te identificeren en maatregelen te duiden om deze risico's te elimineren of te mitigeren. In deze KIIA worden een aantal onderdelen direct getoetst. Enkele andere onderdelen zullen een samenvatting bevatten van een analyse die in een ander document is gedaan. Deze KIIA is afgestemd op het Algemene Rekenkamer Rapport en het audit toetsingskader dat op 26 januari 2021 is gepubliceerd. De (concept) richtlijn van de Europese Commissie 'European Approach for AI' van 23 april 2021 is eveneens al op hoofdlijnen vertaald naar deze KIIA.

De eerste versie van deze assessment is op 23 januari 2020 besproken met de Commissie Persoonsgegevens Amsterdam (hierna: CPA). Op februari 2020 is een advies van de CPA ontvangen. Het oorspronkelijk plan was om middels een collegebesluit van 25 februari 2020 te starten met een Pilot in april 2020. Vanwege de coronapandemie is deze datum verschoven naar juni 2020, maar ook deze datum bleek vanwege de pandemie niet haalbaar te zijn. Sinds dien is de Pilot tot een nog onbekende tijd uitgesteld.

Gedurende het jaar zijn nieuwe inzichten verworven. Een aantal cruciale elementen dienen aan het ontwikkelproces te worden toegevoegd. Het uitstellen van de Pilot heeft daarom ook ruimte gecreëerd om dit te verkennen en toe te passen. Hierbij kan worden gedacht aan het uitvoeren van een 'onderzoek naar onbewuste vooringenomenheid over de onderliggende dataset en het algoritme. Dit onderzoek leidt ertoe het algoritme nader wordt onderzocht middels op (in)direct discriminerende elementen en dat de medewerkers van team Vakantieverhuur intensiever zijn betrokken bij rechtmatigheidsvraagstukken over de werking door het algoritme. De betrokken klassieke en sociale grondrechten zijn scherper in kaart gebracht en er is meer focus gekomen op het onderwerp risicomanagement en informatiebeveiliging.
Tot slot is door de CIO gewerkt aan een Stedelijk Kader waarin de verantwoordelijkheden bij het ontwikkelen en in gebruik nemen van hoge risico KI-toepassing zijn geduid. Hierin is bijvoorbeeld

vastgesteld dat bij een hoge risico KI-toepassing, voorafgaand aan een pilotfase, het algoritme moet worden geaudit door DMC. Er worden daarmee een aantal extra checks- en balances ingebouwd om tot een zo wettige, ethische en robuuste KI-toepassing te komen.

Voorgaande heeft geleid tot aanpassing van het algoritme, meer transparantie en een gestandaardiseerde aanpak voor de ontwikkeling van algoritmes binnen de gemeente Amsterdam. Deze KIIA bevat de belangrijkste informatie met betrekking tot het vakantieverhuur algoritme en zal waar nodig verwijzen naar verschillende bijlagen.

2. Bepalen noodzaak voor het doen van een kunstmatige intelligentie impact assessment

2.1 Wordt de kunstmatige intelligentie (hierna: KI) toegepast in een nieuw (maatschappelijk) domein?

Nee, de KI wordt toegepast binnen het domein van Toezicht & Handhaving in het kader van het bestrijden van woonfraude, illegale vakantieverhuur.

2.2 Wordt gebruik gemaakt van een nieuwe vorm van KI-technologie?

Nee, het gaat om een random forest regressie algoritme.

2.3 Heeft de KI een hoge mate van autonomie?

Nee, de KI is complex van aard en geschikt om risico indicaties te doen, het kunnen schatten van kansen op gebeurtenissen, het detecteren van verdachte gebeurtenissen of het waarnemen van veranderingen in een dataset. Het doet dit op basis van beschikbaar gestelde data. De KI heeft een lage mate van autonomie. Het geeft enkel een advies in de vorm van een risico indicatie voor illegale vakantieverhuur op een adres aan de verantwoordelijke medewerkers. Na dit advies volgt nog een uitgebreid onderzoek dat door verschillende medewerkers wordt uitgevoerd. Er is geen eigen "beslisruimte" voor de KI. Daarmee is dus geen sprake van geautomatiseerde besluitvorming

2.4 Wordt de KI toegepast in een complexe omgeving?

De KI wordt toegepast in een afgebakende omgeving die niet complex is. De input is vooraf bepaald en kan enkel wijzigen als er wordt gekozen voor (door)ontwikkeling van de KI. Bij het implementeren van de KI in de bedrijfsvoering zal de input van de KI ook niet wijzigen.

2.5 Wordt de KI toegepast binnen een 'gevoelig' werkproces?

Ja, er wordt onderzoek gedaan naar woonfraude. Dat betekent dat Amsterdammers in relatie worden gebracht met fraude door middelen van een (interne) melding of een signaal. Een woonadres wordt onder voorwaarden als onderzoekwaardig beschouwd. Het gevolg hiervan is dan een buitendienstonderzoek wordt uitgevoerd, wat betekent dat een tweetal toezichthouders bij een Amsterdammer aanbellen en mogelijk een woning binnentreden. De gemeente heeft toezicht en handhavingsbevoegdheden en de inzet hiervan kan hoge impact hebben op bewoners, zowel vanwege het binnentreden van een woning maar ook omdat een boetebesluit kan worden opgelegd als mogelijk resultaat van het werkproces. Om deze redenen dient dit werkproces als 'gevoelig' te worden beschouwd, en daarmee dus ook de inzet van de KI.

2.6 Heeft de KI eerder schade toegebracht aan fundamentele rechten van natuurlijke personen of groepen?

Nee, het gaat om een nieuw ontwikkeld algoritme. Schade aan fundamentele rechten en natuurlijke personen of groepen zijn nog niet bekend. Er kunnen wel risico's worden gesignaleerd waar mitigerende maatregelen voor zijn geduid.

2.7 Zijn er zorgen over de KI en de potentiële schade en impact?

Nee, waar risico's zijn gesignaleerd zijn een aantal mitigerende maatregelen geduid en uitgevoerd die voldoende zijn om de werking van het algoritme te testen gedurende een Pilot. Tijdens de pilot zullen de resterende risico's nader worden onderzocht.

2.8 Wordt gebruik gemaakt van gevoelige gegevens over personen?

Ja, er worden persoonsgegevens verwerkt in de KI. Het gaat om namen van geregistreerde bewoners van woonobjecten, geboortedatum, adres, postcode, woonplaats. Woonfraudezaken die de afgelopen 5 jaren zijn onderzocht worden verwerkt in de KI. Er worden geen bijzondere persoonsgegevens verwerkt zoals medische gegevens, etniciteit of seksuele voorkeuren.

2.9 Wordt een voorspelling gegenereerd op persoonlijke niveau op basis van persoonlijke kenmerken?

Er wordt geen voorspelling gegenereerd op persoonlijk niveau. Omdat er wel een voorspelling/ risico-indicatie wordt gegenereerd op adresniveau, moet dit worden beschouwd als een direct verband met de betreffende bewoner. Afgeleid van het adres, kan worden gesteld dat met de risico-indicatie een persoon al dan niet onderzoekwaardiger wordt. De risico-indicatie wordt **niet** gegenereerd op basis van geslacht, huidskleur of andere persoonlijke kenmerken.

2.10 Neemt de KI-beslissingen die natuurlijke of rechtspersonen in aanzienlijke mate treffen dan wel rechtsgevolgen voor hen hebben?

Er is geen sprake van besluitvorming of een beslissing maar dit kan (uiteindelijk) wel rechtsgevolgen hebben voor natuurlijke- of rechtspersonen. Zie vraag 8 voor toelichting.

2.11 In hoeverre is de output van de KI van invloed op het besluitvormingsproces binnen de organisatie?

Een te onderzoeken adres waar een vermoeden van woonfraude is komt voort uit een (interne)melding/ signaal. Deze melding wordt meestal ingediend door een burger (omwonende) of kan voortkomen uit scrapingdata. De KI ondersteunt de medewerker van Toezicht & Handhaving, de planner, om de meest risicovolle zaken uit de werkvoorraad te prioriteren en te selecteren voor een buitendienstonderzoek. Daarmee wordt de KI onderdeel van de specifieke afweging om eerder of later een buitendienstonderzoek te starten op een adres. Dit is van invloed op betrokkene, in dit geval de mogelijke overtreder van de huisvestingsverordening. Deze heeft namelijk een grote kans te worden geconfronteerd met een buitendienstonderzoek aan diens adres of in de woning.

Het geven van prioriteit aan te onderzoeken adressen is voor het werkproces bestrijden illegale vakantieverhuur relevant, gezien de vaak korte periodes waarin woonobjecten worden verhuurd. Om deze reden worden binnen het werkproces 'business rules' toegepast op basis waarvan momenteel handmatig meldingen en signalen worden geprioriteerd. De KI is mede geïnspireerd op deze business rules. Of er daadwerkelijk sprake is van illegale vakantieverhuur wordt vastgesteld door de verantwoordelijk toezichthouder en de projecthandhaver. Daarvoor wordt een voor- en buitendienstonderzoek uitgevoerd. Vervolgens wordt de woonfraudezaak intensief besproken in een debriefing met de medewerkers die tot besluitvorming kunnen overgaan.

De risico indicatie heeft een wezenlijke invloed op de planner, maar neemt geen autonome beslissingen. Dat blijft een maatwerkbeslissing per individueel geval. Een woonadres kan wel later of eerder worden onderzocht vanwege de werking van de KI-toepassing. Dat kan ertoe leiden dat een Amsterdammer eerder of later wordt geconfronteerd met een buitendienstonderzoek. Nu dit, indien woonfraude wordt geconstateerd, kan leiden tot een boetebesluit moet de invloed van KI-toepassing als belangrijk worden beschouwd.

2.12 Is er, mede bezien vanuit de geconstateerde en voorzienbare risico's, sprake van een hoge risico KI-systeem?

Ja, er is sprake van een hoge risico KI-toepassing.

2.13 Indien er sprake is van een (potentiële) hoge risico KI-toepassing, is de Autoriteit Persoonsgegevens geïnformeerd over de intentie om de KI-toepassing te ontwerpen en te ontwikkelen?

De Autoriteit Persoonsgegevens zal worden geïnformeerd nadat de audit is uitgevoerd, op basis van het rapport (indien van toepassing) maatregelen zijn getroffen en de startdatum voor de Pilot is gedefinieerd.

2.14 Conclusie (uitvoeren KI-assessment of niet?)

Het uitvoeren van de Kunstmatige Intelligentie Impact Assessment is noodzakelijk, gezien het feit dat persoonsgegevens worden verwerkt, er sprake is van een wezenlijk invloed op het besluitvormingsproces, hoewel niet geautomatiseerd, en de KI is geïnspireerd op basis van de business rules die momenteel worden toegepast. Deze assessment draagt bij om te kunnen bepalen of de werking en inzet van het algoritme rechtmatig is.

3. KI-toepassing, verantwoordelijkheden en betreffende werkproces

3.1 Beschrijf de toepassing en het doel van de toepassing.

Het primaire doel van het verwerken van persoonsgegevens "Toezicht en handhaven Wonen", welke noodzakelijk is voor de vervulling van een taak van algemeen belang of een taak voor het openbaar gezag, is:

- het houden van (risico gestuurd) toezicht op de naleving van de huisvestingsregelgeving en de leegstandswet, alsmede de Woningwet en de Wet algemene bepalingen omgevingsrecht (Wabo) aangaande het gebruik van woningen (bestrijden woonfraude);
- het nemen van effectieve handhavingsmaatregelen bij overtreding van de hieronder genoemde regelgeving.
- woningen beschikbaar maken voor de woningmarkt.

Het secundaire doel van het verwerken van persoonsgegevens met de KI is:

- reguliere meldingen effectiever en efficiënter te gaan onderzoeken door de onderzoekscapaciteit beter in te zetten middels een risico indicatie;
- patronen en verbanden vinden in historische data die in verband kunnen worden gebracht met illegale vakantieverhuur en dit toetsen tijdens een pilot.

Toezicht en Handhaving van directie Wonen werkt in beperkte mate informatie gestuurd. Op dit moment wordt hoofdzakelijk geacteerd na ontvangst van een (burger)melding en scrapedata. Er wordt een onderscheid gemaakt tussen twee type meldingen. Dit zijn issuemeldingen en reguliere meldingen. Issuemelding zijn meldingen waarbij is vermeld dat er op het moment van melden toeristen aanwezig zijn. Deze melding hebben verhoudingsgewijs een hoge "pakkans" (hit) en krijgen voorrang op regulier melding. Reguliere melding zijn alle overige meldingen. Het doel van de toepassing is om reguliere meldingen effectiever en efficiënter te gaan onderzoeken waar het algoritme bij kan ondersteunen.

De werkvoorraad met te onderzoeken zaken is normaliter groter dan de toezicht- en handhavingscapaciteit, omdat de kans op het aantreffen van illegale vakantieverhuur op een adres, na het ontvangen van een (burger)melding, na verloop van tijd afneemt. Doorgaans geldt hoe eerder een buitendienstonderzoek kan worden gestart na ontvangst van een melding, hoe groter de kans dat een onrechtmatige situatie (woninggebruik in strijd van met Huisvestingsverordening) kan worden hersteld

3.2 Beschrijf het werkproces (BPMn uitwerking en toelichting).

Toezichtfase

1) *In behandeling nemen woonfraude vermoeden en uitvoeren vooronderzoek.*

Nadat een melding is ontvangen van een burger, uit de interne organisatie van T&H (of via scraping) of van een ketenpartner als een te onderzoeken adres is aangedragen, wordt een "melding" in behandeling genomen.

De informatie in de melding wordt gecontroleerd. Hiervoor is de projectmedewerker verantwoordelijk. Er zijn meerdere bronnen gekoppeld aan het dossier. [redacted] Alle informatie is nodig voor de toezichthouder die het buitendienstonderzoek uitvoert.

2) Prioriteren & selecteren zaken

Uit de zaken in het zaaksysteem maakt de TOP-applicatie een lijst. Over het algoritme dat ten grondslag ligt aan het hitkanspercentage van woonfraude op een adres, zie het privacystatement Pilot Woonfraude Algoritme (https://www.amsterdam.nl/privacy/specifieke/privacyverklaringen-wonen/pilot-woonfraude-algoritme/). Daarnaast wordt een algoritme gebruikt om een lijst van adressen te genereren die de toezichthouders die dag gaat bezoeken. Dit algoritme borgt de kwaliteit van de te belopen lijst. Er worden verschillende variabelen tegenover elkaar afgewogen zoals de afstand van het ene adres ten opzichte van het andere adres, het hitkanspercentage, of er een actuele melding is, adressen uit een project en welk stadium in een onderzoek voorrang krijgt of juist wordt uitgesloten.

Het afwegen van de variabelen heeft tot doel het toezicht op en handhaving van de huisvestingsregelgeving efficiënter te kunnen uitvoeren. Per adres wordt ter plaatse onderzoek gedaan en dat onderzoek vormt de basis voor de besluitvorming van een eventueel op te leggen sanctie aan de hand van het aangetroffen en vastgestelde gebruik van de woning.

3) Constatering woonsituatie doen op locatie

Met de informatie uit de TOP-app gaan toezichthouders een buitendienstonderzoek uitvoeren aan de hand van de gegenereerde looplijst. Er wordt gebruik gemaakt van de andere module van de webapplicatie, TOP (Toezicht op Pad) waarop, naast de te onderzoeken adressen de belangrijkste zaakgegevens zichtbaar zijn. Er kunnen zich meerdere scenario's voordoen bij het uitvoeren van het buitendienstonderzoek:

- Het is mogelijk om het woonobject te betreden en een constatering te doen;
  - Er is sprake van een brandonveilige situatie;
  - Het vermoeden van woonfraude is bevestigd;
  - Er is geen vermoeden meer van woonfraude.

- Het is niet mogelijk om het woonobject te betreden en een constatering te doen.

Als het niet mogelijk is om een woonobject te betreden, dan wordt dit woonobject opnieuw ingepland voor het doen van een constatering op deze locatie (met evt. machtiging). Indien het wel mogelijk is om een constatering te doen en het vermoeden van woonfraude is bevestigd, dan keren de Toezichthouders terug naar kantoor om de bevindingen te bespreken met de projecthandhaver en eventueel de dag coördinator/teamleider en de handhavingsjurist. Dit wordt de debriefing genoemd. Als er geen vermoeden meer is van woonfraude of er kan geen constatering worden gedaan dan zal de volgende zaak worden bezichtigd. Van de bevindingen wordt altijd in een logboek een notitie gemaakt. Daarnaast is er een kladblok voor eigen notities. Als er geconstateerd wordt dat de BRP gegevens niet accuraat zijn, wordt dit teruggekoppeld aan de BRP.

Voorbereiding besluitvorming handhaving

4) Beslissing over woonfraude vermoeden nemen
Nadat er een constatering is gedaan op een adres worden de bevindingen en de resultaten uit het vooronderzoek besproken. In deze "debriefing" zal worden besloten of er al dan niet sprake is van woonfraude. De project handhaver, toezichthouders, dag coördinator/projectleider en de handhavingsjurist nemen deze beslissing gezamenlijk. De projecthandhaver is verantwoordelijk voor de uitvoering van deze beslissing indien er sprake is van woonfraude.

5) Verwerken informatie in rapport van bevindingen en het beeldverslag
Nadat een constatering is gedaan op locatie wordt er (o.a. bij woonfraude) een rapport van bevindingen en een beeldverslag opgesteld door de toezichthouders. [redacted] Bij geen woonfraude wordt een Buitendienst rapport opgesteld. Deze documenten worden toegevoegd aan de zaak en opgeslagen in Decos.

6) Opstellen voornemen te handhaven en richting kiezen handhavingsbesluit
Voordat een boete wordt verzonden, ontvangt de overtreder een voornemen tot handhaven. Het voornemen wordt opgesteld door de projecthandhaver. In dit voornemen wordt de overtreder geïnformeerd over het feit dat de gemeente voornemens is te handhaven en stelt de overtreder in de gelegenheid om een zienswijze in te dienen. Informatie uit het vooronderzoek en het buitendienstonderzoek wordt verwerkt in het voornemen. Het voornemen wordt toegevoegd aan de zaak en opgeslagen in Decos.

7) Behandelen zienswijze
Indien een zienswijze wordt ontvangen, dan moet deze inhoudelijk worden beoordeeld. Er zal een afweging moeten worden gemaakt of op basis van deze informatie zal worden afgezien van handhaven of niet. De projecthandhaver verwerkt de informatie uit het voornemen en bepaalt samen met de handhavingsjurist of de zienswijze wordt gehonoreerd.

3.3 *Welke processtap/ taak wordt ondersteund met de KI?*

De processtap 'Prioriteren & Selecteren zaken' wordt met de KI ondersteund.

3.4 *Worden bij deze processtap/taak beslisregels toegepast door medewerkers of systemen?*

Ja, zie hiervoor Excelbestand met overzichten van:

- Verwerkte data
- Opgestelde features
- Business rules betreffende werkproces
- Achteraf verwijderde features

In het tabblad 'business rules betreffende werkproces' zijn de beslisregels te zien voor deze processtap.

3.5 *Welke verantwoordelijkheden zijn er als het gaat om het verwerken van de gegevens?*

Illegale vakantieverhuur algoritme	
Verwerkersverantwoordelijke (persoons)gegevens en eigenaar van de KI	Gemandateerd Directeur Wonen: Henriëtte Rombouts
Gegevensverstrekker data uit zaaksysteem aan gegevensverwerker	Directeur Wonen: Henriëtte Rombouts
Gemeentelijke verstrekker Basisregistraties aan Wonen in zaaksysteem BWV (via Makelaarssuite)	Directeur Basisinformatie: Dignar van Onzenoort
Gegevensverwerker (ontwikkeling product)	Directeur CTO: Ger Baron

3.6 *Wie zijn betrokken bij de ontwikkeling van de KI?*

Zie bijlage 1 'overzicht rollen en verantwoordelijkheden'

3.7 *Welke verantwoordelijkheid heeft ieder rol?*

Zie bijlage 1 'overzicht rollen en verantwoordelijkheden'

3.8 *Beschrijf welke actoren een rol spelen in de toepassing.*

Zie bijlage 1 'overzicht rollen en verantwoordelijkheden'

3.9 Beschrijf welke KI-technologie wordt ingezet om het doel te bereiken.

De hiervoor beschreven informatie wordt verwerkt in een algoritme. Het algoritme is ontwikkeld om een antwoord op de vraag te kunnen formuleren of een risico indicatie kan worden gekoppeld aan een melding van illegale vakantieverhuur. Er is gekozen om een random forest regressie algoritme te ontwikkelen en deze is "voorspellend" van aard. Het algoritme valt in de categorie "gebied gebonden analyses" die een "risico indicatie" geeft op adresniveau[1].

Een random forest is een machine learning algoritme dat wordt toegepast op classificatie- en regressievraagstukken. Een random forest kan niet-lineaire relaties in data beschrijven en maakt daarvoor gebruik van een "tree based supervised learning" techniek. Daarmee is het algoritme redelijk complex van aard en geschikt om voorspellingen te doen, het kunnen schatten van kansen op gebeurtenissen, het detecteren van verdachte gebeurtenissen of veranderingen in een dataset te ontdekken.

Het ontwikkelde algoritme genereert op basis van de omstandigheden van het geval een risico-percentage voor woonfraude op een bepaald adres. Deze risico indicatie wordt berekend voor een melding die is ontvangen door een burger of als een medewerker een (interne) melding doet. Het algoritme creëert dus geen "nieuwe" te onderzoeken zaken, maar geeft een risico indicatie van illegale vakantieverhuur voor een ontvangen melding.

Het algoritme berekent eerst een "kans", wat een gemiddelde is van een zeer groot aantal berekeningen (middels kansbomen) dat het model heeft uitgevoerd. Het risico-percentage dat daar vervolgens statistisch mee wordt berekend, kan worden gebruikt voor het prioriteren van zaken. Het algoritme geeft ook uitleg over hoe die kans tot stand is gekomen. Dat wordt in dit document toegelicht.

Met opmerkingen [10.2.e]: @[10.2.e] : Zou jij dit kunnen aanpassen op de nieuwe situatie?

3.10 Beschrijf of er is nagedacht over manieren om het KI-systeem te ontwikkelen of de KI te trainen zonder of met zo weinig mogelijk gebruik van potentieel gevoelige gegevens of persoonsgegevens?

Er zijn adviezen ingewonnen bij de verantwoordelijk privacy officers. Uit die adviezen zijn concrete maatregelen naar voren gekomen. Er heeft dataminimalisatie plaatsgevonden omdat er negatief is geadviseerd over de burgerlijke staat (datum huwelijk, datum echtscheiding, A-nummer, A-nummer partner), nationaliteit, woonplaats, geboorteland en het BSN-nummer. Deze informatie is niet geleverd en daarmee buiten de KI gebleven. De geboortedatum is geaggregeerd van bv. 01-01-1970 naar het geboortejaar 1970. Overige persoonsgegevens zijn (waar mogelijk) gepseudonimiseerd. Zie hiervoor Excellijst bij WPD.

Na start van de Pilot is besloten om de variabelen die zijn gebaseerd op 'geslacht' eveneens te verwijderen. Dit waren de variabelen aantal_mannen en ratio_mannen. Zie hiervoor ook bijlage 3,

[1] Kamerstuk 26 643, nr. 641.

tabblad verwijderde variabelen. Gezien het gevoelige karakter van dit kenmerk en de lage meerwaarde ervan van de variabelen die hierop gebaseerd zijn, bestond er geen noodzaak om de variabelen in de KI-toepassing te behouden. Voorts zijn gedurende de Pilot ook ander features verwijderd gezien de geringe toegevoegde waarde voor het kunnen geven van een risico indicatie.

Tot slot worden tijdens het trainen automatisch de irrelevante features uit het model verwijderd. Dit gebeurt op basis van de zogenaamde 'feature importance'.

3.11 Wie zorgt ervoor dat de KI operationeel (beheer) blijft en wordt gemonitord?

Voorlopig zal de Werkplaats Datagestuurde Aanpak Ondermijning zorgen voor het beheer en monitoren van het vakantieverhuur model. Er zullen in de toekomst afspraken worden gemaakt met Directie Wonen om het beheer onder te brengen in een DevOps-team. Het ligt in lijn der verwachting dat ontwikkelaars van de Werkplaats enige tijd zullen assisteren bij deze werkzaamheden. Er wordt momenteel aan een beheerplan gewerkt.

4. Beschrijf de baten van de KI toepassing

4.1 Wat zijn de baten voor de organisatie?

- Om tot een effectievere en efficiëntere inzet van de beschikbare toezicht en handhavingscapaciteit te komen en om zo meer woningen beschikbaar te maken voor de woningmarkt.
- Het werkproces wordt transparanter gemaakt en er ontstaat meer aandacht voor de relevantie en de vereiste zorgvuldigheid bij het gebruik van data.

4.2 Wat zijn de baten voor het individu?

- Er worden woning vrij gemaakt waarin kan worden gewoond;
- Voor burgers en omwonende zal het pijnpunt overlast (toeristen) effectiever kunnen worden bestreden. Voor de (potentiële)overtreder is dat anders. Deze ondervindt enkel nadeel van de inzet van de KI. De kans dat deze persoon wordt "gepakt" is in theorie groter;
- Op middellange termijn neemt de koop- en huurprijs van woningen af doordat meer woningen vrijkomen voor de woningmarkt;
- Minder kansrijke zaken zullen niet worden behandeld door een toezichthouder. Dat betekent dat de kans dat een onschuldige burger geen impactvolle interactie heeft met een toezichthouder toeneemt.

4.3 Wat zijn de baten voor de maatschappij als geheel?

- Overlast door woonfraude wordt teruggedrongen;
- Door een grotere woningvoorraad zal de koop- en huurprijs mogelijk worden beïnvloed;
- Massatoerisme heeft vergaande gevolgen voor winkelaanbod, leefbaarheid en karakter van de stad. Door woonfraude effectiever te bestrijden zullen deze belangen worden gediend.

5. Architectuur

5.1 Hoe ziet de huidige architectuur eruit voor het gebruik van de KI?



5.2 Hoe zit de toekomstige architectuur eruit voor het gebruik van de KI? Geef hierbij een toelichting.



Zie voor aanvullende informatie bijlage 6.

6. Data en data governance

6.1 *Beschrijf welke brondata worden gebruikt binnen het werkproces en in het kader van de toepassing.*

Het algoritme is ontwikkeld en getraind met informatie uit meerdere bronnen. Het gaat om de oude applicatie (workflow systeem) BWV, de BAG en het nieuwe zaaksysteem AZA. Meldingsgegevens, procesmatige gegevens (starten volgende processtap) en persoonsgegevens (uit het BRP) worden verwerkt uit BWV en AZA. Gegevens uit de BAG worden bevraagd middels een API. Deze gegevens worden deels handmatig door medewerkers van directie Wonen, Toezicht & Handhaving ingevoerd in BWV en AZA. Ook wordt informatie "ingeladen" via andere applicaties. Het gaat dan om de gegevens uit het BRP en BAG die wordt geïmporteerd in BWV of AZA.

Met opmerkingen [10.2.e]: @[10.2.e] Risico: fouten in handmatig toegevoegde data.

Met opmerkingen [10.2.e]: [10.2.e] Risico: fouten in data bronsystemen.

6.2 *Beschrijf of de brondata voor de KI is aangevuld met afgeleide data?*

Nadat de brondata is geselecteerd, kunnen op basis van deze brondata zogenaamde 'features' (hierna: variabelen) worden opgesteld. Voorbeelden hiervan zijn een postcodegebied of de betreffende maand. Het kenmerk 'adres' is zo geaggregeerd naar het niveau van postcodegebied en 'datum wanneer zaak afgerond' vertaald naar 'maand' waaruit blijkt in welke maand de kans op vakantieverhuur het grootst is. In bijlage 3 is een overzicht gegeven van alle variabelen die worden gebruikt. Er wordt dus geen nieuwe brondata toegevoegd, maar hier kan wel afgeleide data uit worden opgemaakt.

Deze afgeleide data kan op twee manieren tot stand komen. Deze zijn:

- De data scientist maakt variabelen op basis van business rules;[10.2.d] [redacted]
- De data scientist bedenkt zelf variabelen;[10.2.d] [redacted]

6.3 Heeft de gemeentelijke organisatie volledige controle en beheersing (eigenaarschap) over de gebruikte data voor de KI?

Ja, de gemeente heeft de volledige controle en beheersing over de gebruikte data voor het model en het model zelf. Het model is door de gemeente zelf ontwikkeld. Externe medewerkers hebben ondersteund bij de ontwikkeling van de gemeente. Zij hebben een geheimhoudingsverklaring getekend. Interne medewerkers van de gemeente hebben de regie over de ontwikkeling van het model gehad.

6.4 Beschrijf de reikwijdte van de gegevens in uw gegevenssets die worden onderzocht (bijv. of ze persoonsgegevens bevatten)?

In het algoritme worden categorie gegevens gebruikt:

- Identificerende gegevens
- Huisvestingsgegevens
- Gegevens over gebouwen
- Gegevens uit woonfraudezaak

Zie bijlage 3 voor een volledig overzicht.

6.5 Wat is de bewaartermijn van de gegevens en voor verwerking ten behoeve van de KI-toepassing?

De gemeente bewaart persoonsgegevens voor zolang dit noodzakelijk is voor de uitvoering van haar taken, of zoals is bepaald in de Archiefwet. Dit betekent dat de persoonsgegevens na de in de selectielijst van de Archiefwet aangegeven periode worden verwijderd. Deze termijnen zijn bij enkel toezichthouden 5 jaar en bij handhaving 10 jaar.

Zolang de KI wordt gebruikt, moet de gebruikte dataset worden bewaard voor archiveringsdoeleinden. Dit is onder meer noodzakelijk om ook achteraf te kunnen onderzoeken of en zo ja, welke 'vooroordelen' (biassen) de dataset / de KI bevat.

De KI-toepassing heeft de gebruikte data niet meer nodig om zelfstandig te kunnen werken om een risico-indicatie te kunnen doen per nieuw ontvangen melding. Zolang de Pilot duurt zal de data worden bewaard en pas bij implementatie (indien dit het geval is) worden gearchiveerd.

6.6 Hoe is de data geleverd en verzameld?

De gebruikte data is afkomstig uit het werkproces (zie 7.1). Er is geen data apart verzameld ten behoeve van dit project.

Voor het trainen van de KI is gebruik gemaakt van een eenmalige data dump van het zaaksysteem BWV en AZA (en daarin opgenomen BRP-gegevens) naar de 10.2.g en daarnaast de API van de

BAG. Bij het gebruiken van de KI worden de gegevens direct opgehaald uit de database van BWV en AZA via de API van de BAG.

6.7 Welke afspraken zijn er gemaakt over de levering en verzameling van de data?

De ruwe data wordt eerst verwerkt door een data engineer, voordat de data scientist deze te zien krijgt. Dit zijn twee verschillende personen. De verwerking bestaat uit het hashen van gevoelige data. Alleen de gehashte data is beschikbaar gemaakt voor de data scientists.

6.8 Wat is de kwaliteit en kwantiteit van de beschikbare data?

De hoofddataset bestaat uit zo'n 100.000 woonfraudezaken op +- 70.000 verschillende adressen. Verder is er data over zo'n 2 miljoen personen beschikbaar, die alleen gebruikt is wanneer een persoon met een zaak te maken had. Tot slot is gebruik gemaakt van een dataset met de adreshistorie uit de BRP, die ongeveer 1,6 miljoen rijen omvat.

De datakwaliteit is over het algemeen in orde afgezien van een paar kleine problemen. Zo is er geen aanvullende data beschikbaar over oude adressen die nu niet meer bestaan, maar waar wel een zaak over in de dataset zit. Deze zaken zijn alsnog meegenomen zonder aanvullende adresinformatie. Een ander voorbeeld zijn een aantal zaken waarbij de begindatum na de einddatum lag. Deze zijn eruit gefilterd.

6.9 Is de data passend voor de KI-toepassing? Zo ja, waarom?

Ja, de data die is verwerkt is noodzakelijk om het werkproces uit te voeren. Er is voor gekozen om een deel van de beschikbare data (in BWV) niet te verwerken. De resterende data, in samenhang met de getroffen maatregelen, is als passend beschouwd om een model mee te ontwikkelen.

6.10 Heeft de data passende 'statistische' eigenschappen? Dit geldt zowel voor de individuele dataset als ook voor een combinatie van datasets.

Ja, want het model is in staat om frauduleuze zaken van niet-frauduleuze zaken te onderscheiden, dus de statistische eigenschappen zijn geschikt om dat doel te bereiken.

6.11 Zijn er relevante aannames, met betrekking tot de informatie die de gegevens geacht worden te meten en te duiden?

We nemen aan dat de gebruikte data correct verzameld is en representatief voor toekomstige woonfraudezaken in Amsterdam.

6.12 Zijn er 'data gaps' of tekortkoming in de data? Zo ja, kunnen er maatregelen worden troffen om dit te corrigeren? Welke maatregelen zijn dit?

Zie vraag 7.9 over de missende aanvullende adresinformatie. Dit zou gecorrigeerd kunnen worden door historische BAG-data op te vragen. Dit is niet gedaan, omdat het slechts een klein aantal zaken betreft.

6.13 Zijn passende maatregelen genomen om categorieën persoonlijke data te beschermen? Zo ja, welke maatregelen zijn die (bv: encryptie, pseudonimisatie, aggregatie, anonimisatie)

Ja, zie hiervoor bijlage 3, overzicht verwerkte data.

6.14 Welke datavoorbereiding- en verwerkingsmethodieken zijn toegepast? (annotatie, labelling, cleaning, aggregatie)

Het label van een zaak is positief (wel woonfraude) als de zaak als 'RAP naar HAN' ('rapportage naar handhaving') is opgevoerd in BWV. Het label van een zaak is negatief (geen woonfraude) als de zaak wel onderzocht is, maar zonder stadium 'RAP naar HAN' in BWV.

De data is opgeschoond o.a. door duplicaten te verwijderen, datums op te schonen, hoofdletters te verwijderen uit stringvelden en missende waardes op te vullen.

Daarnaast zijn alle features gecreëerd op zaakniveau. Hiervoor was het nodig om bijvoorbeeld persoonsdata te aggregeren naar het zaakniveau door het aantal personen te tellen, etc. De volledige lijst van features is te vinden in bijlage 3.

6.15 Zijn er afspraken gemaakt over de bewaartermijnen van de verwerkte data en de verwijdering/ archivering van de data als de termijn is verlopen of de data niet meer noodzakelijk is voor de ontwikkeling van de KI? Zo ja, wat zijn deze afspraken?

Data die uiteindelijk niet wordt gebruikt voor het model moet worden verwijderd. De data engineer en data scientist hebben hiervoor periodieke afstemming met de betrokken jurist en informatiebeveiligingsadviseur.

Voor bewaartermijn zie vraag 7.5.

6.16 Is de training-, validatie- en testdata relevant, representatief, foutloos en compleet?

Er is veel aandacht besteed aan de "datakwaliteit". De kwaliteit van de data bepaalt in grote mate hoe goed het algoritme werkt. De data die is onderzocht komt uit het zaaksysteem van Toezicht en Handhaving. Data-analisten hebben uitvoerig onderzoek gedaan naar de bruikbaarheid van de beschikbare data. Er is geconstateerd dat de informatie in het zaaksysteem even volledig is ingevuld. Het opvoeren van een foutieve datum, bv. 1-1-2108 in plaats van 1-1-2018 of het niet opvoeren van een stadium is hier een voorbeeld van. Daardoor is informatie in het zaaksysteem niet altijd volledig of correct geweest. Het zaaksysteem bevat geen controlemechanisme om dit te

corrigeren of onder de aandacht te brengen van de gebruikers. De data analisten hebben dit "gerepareerd".

Daarnaast zijn er ook technische mankementen geweest, wat in de loop der jaren heeft gezorgd voor een zekere "vervuiling" van data. Data analisten hebben deze vervuiling in de gebruikte dataset zoveel mogelijk "opgeschoond". Er is echter geconcludeerd dat een deel van de beschikbare data onbruikbaar is voor het algoritme. Deze informatie is dan ook voor een deel buiten het algoritme gelaten maar gedeeltelijk ook erin verwerkt. Na het uitvoeren van statistische analyses en controles blijkt desondanks dat de datakwaliteit voldoende is om een statistisch "betrouwbare" kansberekening te kunnen berekenen. Of een kansberekening juist is, moet worden getest gedurende een pilot aan de hand van nieuwe zaken.

Een belangrijk element bij het gebruik van data zijn definitiebepalingen. Alle definities en aannames zijn gecontroleerd met de medewerkers van de afdeling T&H Wonen. Dat betekent dat de gebruikte data op dezelfde manier wordt geïnterpreteerd. Tot slot is er tijdens de analyses altijd controle uitgeoefend door de data analisten en de medewerkers van de afdeling.

6.17 Zijn gedurende de training van de KI-toepassing de uitkomsten en relevante ontwikkelkeuzes geëvalueerd en gedocumenteerd?

Tijdens het ontwikkelen van het model is de gangbare manier van werken gebruikt dat na iedere verandering aan het model en bij iedere nieuwe feature is gekeken of dit het model verbetert. Veranderingen zijn uiteindelijk alleen doorgevoerd als dit het geval bleek. Het algoritme is ontwikkeld met 80% van de beschikbare data. Vervolgens is met 20% van de overige beschikbare data het algoritme gevalideerd.

6.18 Zijn er duidelijke criteria/parameters vastgesteld voor het correct functioneren van de KI?

Er zijn twee waardes gebruikt om het correct functioneren van de KI te evalueren: de 'AUC' (area under the receiver-operator curve) voor het trainen en daarnaast de 'lift' voor de pilotfase. Deze begrippen zijn uitgelegd in het pilotplan. Ook de criteria die hieraan worden gesteld staan daar beschreven.

Naast deze twee waardes wordt tijdens het trainen van het model ook berekend wat de bias is op diverse features. Het model wordt alleen in gebruik genomen als de bias (en uiteraard de AUC en lift) acceptabel is.

Met opmerkingen [10.2.e]: [10.2.e] Voor het correct functioneren zoals in deze vraag gevraagd wordt is natuurlijk ook de bias van belang en het liefst zouden we hier ook zeggen dat de bias na de pilot nogmaals gecheckt wordt (dus hoe het in de praktijk uit heeft gepakt). Maar we hebben de BRP-data natuurlijk alleen voor een eenmalige analyse vooraf geleverd gekregen. Voor nu hebben we die check na de pilot hier dus gewoon genegeerd. Maar puntje van aandacht

6.19 Welke methoden worden gebruikt om te verifiëren of de KI binnen de gestelde parameters blijft?

Op dit moment wordt niet continu gemonitord of de KI binnen de gestelde parameters blijft. Het is gepland dat dit geïmplementeerd gaat worden zodra het model daadwerkelijk in productie

genomen wordt. Dat zouden bijvoorbeeld automatische notificaties kunnen zijn wanneer een van de waardes uit vraag 6.18 zich sterk verandert.

6.20 Is de output van de KI consequent?

Er zijn twee maatregelen genomen om ervoor te zorgen dat het model consequent handelt. De eerste maatregel is het implementeren van een zogenaamde 'random seed'. Dit betekent dat iedereen die het model traint met dezelfde dataset het model ook op dezelfde manier traint en op dezelfde waarden uit komt. Verder wordt het getrainde algoritme opgeslagen in een zogenaamde 'pickle' file. Dit bestandstype zorgt ervoor dat het getrainde model ook in de praktijk dezelfde resultaten en handeling verricht. Daarnaast kan het model ook overgedragen worden naar andere systemen waar het ook consequent kan handelen.

6.21 Welke output data genereert de KI?

De KI geneert een risico-indicatie in de vorm van een percentage. Dit percentage wordt gekoppeld aan een (interne) melding van illegale vakantieverhuur. De risico-indicatie wordt getoond in de webapplicatie TOP naast het te onderzoeken adres. Hier wordt ook het icoon van een 'robot' bij getoond, zodat duidelijk is dat het percentage is gegenereerd door het model. Naast de risico-indicatie worden ook de features getoond die de risico-indicatie stuwen (hoger percentage) en laten dalen (lager percentage).

6.22 Is het mechanisme van peer review toegepast? Zo ja, welke afspraken zijn gemaakt?

In het team maken we altijd gebruik van het 4 ogen principe en daarnaast keur je nooit je eigen code goed

Met opmerkingen [10.2.e]: @[10.2.e] : zou jij hier kunnen toevoegen in hoeverre hier afspraken over zijn gemaakt in het team? Vraag ihkv de audit

6.23 Welke relevante designkeuzes voor de KI-toepassing zijn gemaakt?

Er wordt gebruik gemaakt van een model dat in de data science wereld erg gangbaar is en goed begrepen/onderzocht, namelijk een random forest. Dit model is gekozen, omdat het een goede mix biedt van transparantie/uitlegbaarheid en prestaties. Het model is geoptimaliseerd op basis van de data en verschillende parameterinstellingen.

6.24 Is er onderzoek gedaan naar onbewuste vooringenomenheid in de dataset en de KI-toepassing? Zo ja, wat zijn de uitkomsten hiervan en welke maatregelen zijn getroffen?

Ja zie hiervoor hoofdstuk 9.

7. Wettige en ethische toepassing

7.1 Wat is de wettelijke grondslag voor de beoogde verwerking van de gegevens ten behoeve van het ontwikkelen, trainen en inzetten van de KI?

Operationele toezicht- en handhavingstaak
- Artikel 6 lid 1 sub e AVG;
- Artikel 174a Gemeentewet;
- Artikel 14 Woningwet;
- Artikel 8, 21, 32 en 35 Huisvestingswet;
- Artikel 5:13 – art. 5:19 Algemene wet bestuursrecht;
- Artikel 3.1.2 lid 5, 4.2.1 en 4.2.2 van de Huisvestingsverordening;

Testen en valideren van algoritme
- Artikel 6 lid 1 sub e AVG;
- Artikel 174a Gemeentewet;
- Artikel 14 Woningwet;
- Artikel 8, 21, 32 en 35 Huisvestingswet;
- Artikel 5:13 – art. 5:19 Algemene wet bestuursrecht;
- Artikel 3.1.2 lid 5, 4.2.1 en 4.2.2 van de Huisvestingsverordening;

Onderzoek naar onbewuste vooringenomenheid)?
- Artikel 3.8 Wet BRP (verbijzondering AVG voor levering BRP-gegevens);
- Autorisatiebesluit (leveringsovereenkomst):

Onderliggende algemene beginselen behoorlijk bestuur en te eerbiedigen (grond)rechten:
- Artikel 3:2 - 3:4 lid 1 Algemene wet bestuursrecht (zorgvuldigheidsbeginsel);
- Artikel 2:4 Algemene wet bestuursrecht (verbod op vooringenomenheid);
- Artikel 3:47 Algemene wet bestuursrecht (motiverings- en rechtszekerheidsbeginsel).
- Artikel 1 van de Grondwet en artikel 14 EVRM (gelijkheidsbeginsel).
- Artikel 14 EVRM en art. 1 Gw discriminatie verbod
- Artikel 2 Algemene wet gelijke behandeling

Er is een grondslag en bevoegdheid om onderzoek te kunnen doen naar woonfraude, illegale vakantieverhuur, door verschillende bronnen te raadplegen om patronen en verbanden te vinden. Dit wordt uiteraard begrensd door te toetsen aan het evenredigheidsbeginsel, waarin proportionaliteit en subsidiariteit centraal staan (art. 5:13 Awb.

7.2 Welke waarden en belangen spelen een rol in de context van de KI toepassing?

Zie template 'weging betrokken (grond)rechten en waarden bij inzet KI

7.3 Hoe worden de waarden en belangen van betrokken actoren geborgd?

Zie template 'weging betrokken (grond)rechten en waarden bij inzet KI

7.4 Hoe wordt gedurende het ontwikkelproces getoetst aan deze waarden en belangen wat betreft de werking van de KI-toepassing?

Analyseren werkproces en informatiegebruik

Het werkproces van Toezicht en Handhaving is inzichtelijk gemaakt door het opstellen van aparte procesbeschrijvingen (BPMn) en door het uitvoeren van een DPIA. Dit zorgt voor een breder inzicht op het mogelijke effect bij de inzet van een algoritme en om strategische keuze te maken.

Analyseren en juridisch / ethisch toetsen gewerkte data en de business rules

In bijlage 3, tabbladen 'Overzicht verwerkte data' en 'Business Rules, is een overzicht gegeven van de toegepaste business rules. Los van de wettelijke verplichting om toezicht te houden als een burger daarom verzoekt, wordt gekeken hoe de onderzoekscapaciteit zo effectief mogelijk kan worden ingezet. De business rules zijn een weergave van de afwegingen die worden gemaakt om een melding al dan niet prioriteit te geven voor een onderzoek. Deze business rules zijn getoetst op juridische en ethische houdbaarheid. Een weging van deze business rules is ook in dit overzicht terug te vinden.

De conclusie hiervan is dat de manier waarop een onderscheid wordt gemaakt in de prioritering van meldingen en signalen gerechtvaardigd zijn. Deze business rules leiden niet tot het maken van een ongerechtvaardigd onderscheid tussen groepen of personen.

Selecteren brondata

Met het opstellen van de PIA en DPIA heeft een informatie analyse plaatsgevonden. Daaruit is gebleken welke informatie op welke momenten uit welke bronnen worden geraadpleegd of wordt ontvangen. Aanvullend wordt het WPD doorlopend, waarin in een separate PIA (onderdeel WPD) wordt beschreven de beoogde inzet van de KI, de benodigde data en de beveiligingsmaatregelen die moeten worden getroffen. Het doorlopen van dit proces zorgt voor het inwinnen van een akkoord bij de privacy en security officers. Aanvullend moet akkoord worden ingewonnen bij de verwerkersverantwoordelijke(n). Dit heeft tot dataminimalisatie geleidt, bijvoorbeeld door het kenmerken nationaliteit of verblijfsstatus niet te verwerken.

De geselecteerde brondata

Tijdens de ontwikkeling van het algoritme is kritisch gekeken naar de gebruikte dataset. Hiervoor is een privacy assessment opgesteld, wat heeft geleid tot dataminimalisatie van de gevraagde dataset. Slechts die informatie die relevant is voor het vaststellen van een overtreding van de Huisvestingswet, is opgenomen in de dataset waarop het algoritme is ontwikkeld. Informatie als geboorteplaats, nationaliteit, burgerlijke staat en geboorteland zijn daarmee niet verwerkt in het

algoritme. In het werkproces wordt ook geen onderscheid gemaakt in te onderzoeken adressen op basis van deze kenmerken. Dat zorgt ervoor dat er geen vertekend beeld ontstaat op persoonsniveau of voor specifieke groepen. Omgekeerd kunnen deze gegevens op een later moment worden gebruikt om te toetsten of het ontwikkelde algoritme niet onbedoeld alsnog discrimineert of ongelijke behandeling van burgers tot gevolg heeft. Dit wordt toegelicht in onderdeel 6, bias analyse.

De 'input features' voor het algoritme

Nadat de brondata is geselecteerd, kunnen op basis van deze brondata combinaties worden gemaakt, al dan niet geïnspireerd op basis van business rules. Een voorbeeld hiervan is het aantal m2 woonoppervlak per ingeschrevene op een adres. Waar de brondata dus 'aantal m2 oppervlak woonobject' en 'aantal ingeschreven personen' bevat, komt er een extra afgeleide tabel bij waarin het gemiddelde staat van beide. In bijlage 3 is een overzicht gegeven van alle tabellen die worden gebruikt. Er wordt dus geen nieuwe brondata toegevoegd, maar hier kan wel afgeleide data uit worden gefilterd. De features moeten passend zijn bij de processtap die moet worden ondersteund met het model.

Bepalen grenswaarde

Niet elke feature is even belangrijk voor de KI. Beredeneerd vanuit de beginselen van dataminimalisatie (noodzakelijkheidsbeginsel) wordt met de betrokken Product Owner een zogenaamd grenswaarde vastgesteld, waarmee het model in zekere zin wordt 'afgeslankt'. Dit levert de situatie op dat de KI naar minder features 'kijkt' en tegelijkertijd statistisch gezien een relevante risico indicatie blijft genereren. Hier moet wel worden opgemerkt dat, vanwege het machine learning karakter van het model, in de toekomst de relevantie van de features kunnen wijzigen en mogelijk andere features relevanter worden voor het model. Features die nu wel binnen de grenswaarde vallen kunnen in de toekomst dus buiten de grenswaarde kunnen vallen.

Valideren van de features (ontwikkeltraject)

Tijdens dit ontwikkeltraject gaat de data scientist onderzoeken welke features op zichzelf en in samenhang met elkaar het sterkst mogelijk model vormen en de meest betrouwbare risico indicatie geeft. Dit wordt gemeten aan de hand van de statistisch gezien meest 'betrouwbare' risico indicatie. Tijdens dit proces worden features toegevoegd of verwijderd. Het resultaat van dit ontwikkeltraject zijn de opgesomde features in bijlage 3, tabblad 'features'.

Per feature is 'statistisch' inzichtelijk gemaakt in hoeverre deze van invloed is op de voorspellende waarde van het model. Hieruit kan worden afgeleid of een bepaalde features, bijvoorbeeld gebaseerd op een business rule, statistisch gezien van meerwaarde is en dus of het specifiek kijken naar bepaalde zaken in een onderzoek iets toevoegt aan de uitkomst van een onderzoek. Dit kan worden meegenomen in de afweging om bepaalde features al dan niet te gebruiken.

Medewerkers worden betrokken bij de vraag in hoeverre een feature passend is bij de uitvoering van de processtap. Wordt deze feature al min of meer toegepast binnen het werkproces of is het

een nieuw 'inzicht' die statistisch gezien relevant is. Er wordt per feature genoteerd in hoeverre hier sprake van is.

Uitvoeren bias analyse
Zie hiervoor onderdeel 9.

Bespreken features, juridische / ethische toetsing en vaststellen welke features worden verwerkt.
Nadat het model is ontwikkeld, moet er een weging worden gemaakt met de medewerkers team Vakantieverhuur, de teamleider en het verantwoordelijk afdelingshoofd of deze features al dan niet moeten / kunnen worden gebruikt voor de uitvoering van de wettelijke taak. Voor een deel bestaan de input features namelijk een weergave van hoe het werkproces wordt uitgevoerd maar voor een deel bestaat dit ook uit afwegingen die een data scientist heeft gemaakt. Dit moet aandachtig worden bekeken en wordt telkens herhaald. Hierbij is onder meer getoetst aan artikel 26 van het Internationaal Verdrag betreffende de uitbanning van alle vormen van rassendiscriminatie, artikel 1 van Protocol nr. 12 bij het Europees Verdrag voor de Recht van de Mens jo. art. 8 EVRM. Deze uitwerking is weergeven in bijlage 2, tabblad 'features' en bevat de overwegingen die zijn gemaakt per feature, indien deze als 'gevoelig' is aangemerkt.

Tot slot wordt door de eindverantwoordelijk vastgesteld wat de definitief te gebruiken features zijn. Daarmee wordt ook de eindversie van het model bepaald.

Valideren tijdens Pilot
Het algoritme moet worden gevalideerd om te kunnen concluderen of de gegenereerde risico indicatie "correct" en daarmee bruikbaar zijn. Dat is al gedaan door een deel van de historische data te verwerken[2] in het algoritme en de uitkomsten te vergelijken. De resultaten hiervan zijn positief. Nu moet het algoritme ook worden getoetst in de praktijk. Dit wordt gedaan in de vorm van een pilot. Tijdens de pilot worden alleen ontvangen regulier meldingen, die zijn voorzien van een risico indicatie van illegale vakantieverhuur, onderzocht. De uitvoering van de pilot is beschreven in een validatieplan. Deze is gezamenlijk opgesteld met het data science team van het CTO. Vanwege de coronamaatregelen is het validatieplan herschreven nu er een periode lang geen buitendienstonderzoeken zijn uitgevoerd.

Conclusie
Op basis van bovenstaande analyses stelt het team vast dat binnen het werkproces en in het algoritme geen ongerechtvaardigd onderscheid wordt gemaakt tussen personen of groepen.

7.5 Is er, in de gebruikssituaties waarin er mogelijk sprake kan zijn van negatieve gevolgen voor de grondrechten, een effectbeoordeling op het gebied van grondrechten uitgevoerd?

Ja, zie bijlage 5 'weging betrokken (grond)rechten en waarden bij inzet KI'.

7.6 In hoeverre hebben betrokkene een kwetsbare positie ten opzichte van de gebruiker van de KI (bezien vanuit macht, kennis, economische en sociale omstandigheden of leeftijd).

De betrokkene heeft in het werkproces illegale vakantieverhuur een kwetsbare positie. Er kan buiten medeweten van betrokkene een melding of signaal op diens adres worden ontvangen. Zonder hiervan op de hoogte te zijn wordt een vooronderzoek naar het adres en daarmee ook de persoon gedaan. Er worden (persoons)gegevens van deze persoon verwerkt en er wordt een woonfraudezaak aangemaakt.

Het aanmaken van een woonfraudezaak zal leiden tot een buitendienstonderzoek op het betreffende adres. De betrokkene is niet verplicht om mee te werken aan dit onderzoek, hoewel dit doorgaans wel het geval is.

Het toepassen van de KI in het werkproces zorgt niet voor een verandering in de positie van deze betrokkene.

[2] Het algoritme is ontwikkeld met 80% van de beschikbare data. Vervolgens is met 20% van de overige beschikbare data het algoritme gevalideerd.

7.7 Wordt door de KI bepaalde personen of groepen schade toegebracht of ongunstig behandeld wat niet gerechtvaardigd is en in verhouding staat tot hun sociale gedrag of de ernst ervan?

De KI-toepassing kijkt niet naar het sociale gedrag van betrokkenen. Zie hiervoor het overzicht van de verwerkte data en de input features. Er is met zorgvuldigheid gekozen voor een beperkt gebruik van data en de input features. Het ligt niet in lijn der verwachting dat personen of geroepen schade wordt toegebracht of ongunstig behandeld. Momenteel is de werking van de KI-toepassing nog niet gevalideerd gedurende een Pilot fase. Pas na afronding van de Pilot kan deze vraag zuiverder worden beantwoord. Gezien de bekende en voorzienbare risico's en de getroffen maatregelen, wordt starten met het valideren van de KI-toepassing in een Pilotfase als gerechtvaardigd beschouwd.

7.8 Wordt door keuze architectuur of de gebruikersinterface van de KI menselijk gedrag, meningen of beslissingen beïnvloed?

In de Webapplicatie TOP kan de risico indicatie per adres met de toelichting van de (minst)belangrijke features worden getoond. De risico indicatie wordt uitgedrukt in een percentage. Het tonen van een hogere of lagere risico indicatie zal de gebruiker altijd tot op zekere hoogte beïnvloeden. Om het risico van protocollering van medewerkers te mitigeren, wordt een gebruikersinstructie opgesteld en worden workshops gegeven.

Gebruikers hebben te allen tijde de mogelijkheid om zelfstandig te beslissen welke zaken wel of niet worden geselecteerd voor een buitendienstonderzoek. Gebruikers worden door keuze architectuur of de gebruikersinterface hier niet in geremd of beperkt.

7.9 Versterkt of vergroot het KI-systeem menselijke capaciteiten?

De KI vergroot menselijke capaciteiten. De KI geeft een risico indicatie op basis van data-analyses. Deze data-analyse wordt vele malen sneller en nauwkeuriger uitgevoerd door een machine learning algoritme dan een medewerker. In korte tijd kunnen een veelvoud aan risico indicaties worden gemaakt ten opzichte van een medewerker.

7.10 Wie is de "human in control" en wat zijn de momenten of hulpmiddelen voor menselijke interventie?

Nadat de planner een adres heeft geprioriteerd en geselecteerd voor een buitendienstonderzoek, gaan de toezichthouders daadwerkelijk het buitendienstonderzoek uitvoeren. De toezichthouders doen een constatering van de woonsituatie op het adres. De bevindingen leggen de toezichthouders vast in een buitendienstrapport. Als bij de constatering het vermoeden van woonfraude wordt bevestigd, dan keren de toezichthouders terug naar kantoor en zal in een debriefing de zaak worden besproken met de projecthandhavers, de handhavingsjurist en mogelijk de dag coördinator.

In de debriefing wordt een beslissing genomen of er sprake is van woonfraude of niet. Als er sprake is van woonfraude, dan zal projecthandhaver een voornemen en later een beschikking opstellen. Iedere woonfraudezaak wordt door meerdere expertises bekeken. Daardoor is er maatwerk, voldoende menselijke interventie en geen sprake van een geautomatiseerde besluitvorming middels een KI.

Daarnaast wordt een menselijk toezichthouder aangewezen die toetst of de KI naar behoren presteert. Deze persoon wordt aangewezen als er kan worden gestart met de Pilot. In de gebruikersinstructies zullen de bijbehorende taken en verantwoordelijkheden worden beschreven.

7.11 In hoeverre is het mogelijk om een persoon 'uit' de KI te halen (opt-out?)

Een verzoek tot verwijdering van persoonsgegevens kan worden ingediend via de gemeentelijke website. Via het loket zal directie Wonen dit verzoek ontvangen. Er moet vervolgens worden onderzocht in hoeverre het verzoek moet worden gehonoreerd (bewaartermijnen) met als gevolg verwijdering van de persoonsgegevens uit de dataset. Voor de data engineer is het eenvoudig om de betreffende persoonsgegevens uit de onderliggende dataset te halen als daartoe de opdracht wordt gegeven. Daarmee is het mogelijk om de persoon uit de onderliggend dataset te halen en afleid uit de werking van de KI.

7.12 Zijn er voorzorgsmaatregelen genomen om overmatig vertrouwen in of op het KI-systeem bij werkprocessen te voorkomen?

Het model is niet gericht op gedragsbeïnvloeding van de burger/ melder of de overtreder van de huisvestigingsverordening. Punt van aandacht is wel de rol – met de professionele keuzes – van de planners en de toezichthouders. Het gevaar bestaat dat deze te veel "geprotocolleerd" worden, wat het maken van een zelfstandig, professioneel oordeel in de weg kan gaan staan.

Getroffen maatregelen:

- Gebruikersinstructies gegeven aan de planners en toezichthouders over de werking en risico's bij gebruikmaken van het algoritme;
- Teamleider betrekken bij de ontwikkeling en het monitoren van het algoritme;
- Per zaak de risico indicatie tonen met de bijhorende toelichting waarop de voorspelling is gebaseerd;
- De risico indicatie is één van die 4 factoren die meeweegt om een adres te selecteren voor een buitendienstonderzoek. Naast de risico indicatie kan bv. worden geselecteerd op planning (logische en efficiënte looplijst maken) en of er een hercontrole moet plaatsvinden;
- Het werkproces is ook uitvoerbaar zonder de risico indicatie.

7.13 In hoeverre kan de output van de KI worden teruggedraaid, gecorrigeerd of buiten beschouwing worden gelaten?

Er zijn een aantal mogelijkheden:

- Het werkproces kan worden uitgevoerd zonder de KI. De menselijk toezichthouder controleert doorlopend de werking van het algoritme en kan deze zelfstandig uitschakelen;
- De output van de KI kan worden genegeerd door de gebruiker;
- Een Amsterdammer die (eerder) betrokken is (geweest) bij een woonfraude zaak kan verzoeken om correctie of verwijdering van diens gegevens.

7.14 Is er getoetst in hoeverre er sprake is van profilering en in hoeverre dat is toegestaan?

De definitie van profilering is "elke vorm van geautomatiseerde verwerking van persoonsgegevens waarbij aan de hand van persoonsgegevens bepaalde persoonlijke aspecten van een natuurlijke persoon worden geëvalueerd, met name de bedoeling zijn beroepsprestaties, economische situatie, gezondheid, persoonlijke voorkeuren, interesses, betrouwbaarheid, gedrag, locatie of verplaatsingen te analyseren of te voorspellen".

De KI-toepassing valt onder de reikwijdte van deze definitie. Binnen het werkproces worden business rules ofwel beslisregels toegepast. Als aan die regels is voldaan krijgt een ontvangen melding een hogere of lagere prioriteit om te onderzoeken. Daarmee kan worden gesteld dat er sprake is van een profiel waar ontvangen meldingen 'naast' kunnen worden gelegd.

De output van de KI is dat bij een ontvangen melding een risico indicatie wordt gegeven voor illegale vakantieverhuur. De KI is geïnspireerd op de gehanteerde beslisregels van de business. Daarnaast kan het statistische relevante en nieuwe 'verbanden' bevatten die zijn gecreëerd door de data scientist. Een signaal of melding wordt dus vergelijken met deze verbanden, waarna een risico indicatie wordt opgesteld.

7.15 Worden kwetsbare groepen vanwege hun leeftijd, fysieke of mentale handicap schade toegebracht door gebruik te maken van de KI?

Hier is geen sprake van. Het werkproces is melding/ signaal gericht, waarbij geen onderscheid wordt gemaakt op basis van kwetsbare groepen vanwege bv. leeftijd, fysieke of mentaal handicap.

7.16 Wordt door gebruik te maken van de KI de geloofwaardigheid van natuurlijke personen gedurende een bepaalde periode op basis van hun sociale gedrag geëvalueerd of geclassificeerd?

Nee, er wordt niet gekeken naar sociaal gedrag van Amsterdammers. Er wordt enkel gekeken of [redacted] 10.2.d [redacted].

7.17 Worden door de KI voorspellingen gegenereerd op persoonsniveau op basis van persoonlijke kenmerken die leiden tot een 'sociale score'?

Nee, de output van de KI leidt niet tot een 'sociale score'. De output van de KI leidt tot een meer of minder een onderzoekwaardig adres.

7.18 Wordt door de KI een natuurlijke persoon schadelijk of ongunstig behandeld in een sociale context die niet gerelateerd is aan de context waarvoor de data is verzameld/ gegenereerd?

Nee, zie hiervoor de primaire en secundaire verwerkingsgrondslag onder vraag 4.1.

7.19 Is er een menselijke toezichthouder aangewezen om de werking van de KI te monitoren en om effectieve maatregelen als de KI niet naar behoren presteert?

Deze wordt aangewezen met bijbehorende taken en verantwoordelijkheden als de Pilot start.

7.20 Is de onderliggende logica in de KI-toepassing passend bij de wettelijke taak die moet worden uitgevoerd met de inzet hiervan?

Ja, zie hiervoor ingevulde template 'verwerkte data, business rules, features'.

7.21 Is de KI-toepassing voldoende transparant en controleerbaar? Zo ja, waarom?

Ja, zie hiervoor ingevulde template 'verwerkte data, business rules, features'. De relevante verwerking en beslispunten over de werking van de KI-toepassing zijn gewogen en beschreven, gevolgd door vaststelling. De gekozen technische oplossing (random forest regressie algoritme) is verhoudingsgewijs 'transparant' ten opzichte van andere KI-toepassingen zoals deep learning. Een combinatie van type technische oplossing en intensieve toetsing maakt dat de KI-toepassing voldoende transparant en controleerbaar is.

7.22 Is de ontwikkelde KI-toepassing wettig en ethisch verantwoord? Zo ja, waarom?

De ontwikkelde KI-toepassing is met voldoende checks and balances ontwikkeld om te kunnen concluderen dat de KI voldoet aan wettelijke kaders en ethische aspecten om een Pilotfase te kunnen starten. Uit de Pilotfase moet blijken of er, al dan niet met (technische) aanpassingen, ook kan worden overgegaan tot implementatie van de KI-toepassing.

8. Bias analyse

De bias analyse omvat de volgende onderdelen:

1. Te onderzoeken ("gevoelige") groepen bepalen
2. Hypotheses opstellen over features die tot indirecte bias leiden
3. Metrics selecteren passend bij het project
4. Directe bias analyseren
5. Indirecte bias analyseren
6. Bias op niet meetbare variabelen analyseren
7. Gevonden bias wegen en beoordelen met verantwoordelijke directie
8. Bias waar nodig mitigeren
9. Conclusies opstellen

Deze onderdelen hoeven niet exact in de weergegeven volgorde uitgevoerd te worden en hebben bovendien een iteratief karakter: uitkomsten van één onderdeel kunnen ertoe leiden dat een ander onderdeel nogmaals uitgevoerd of aangepast wordt. Dit herhaalt zich totdat alle onderdelen naar tevredenheid zijn afgerond. De onderdelen worden hieronder in detail toegelicht.

8.1 Te onderzoeken ("gevoelige") groepen bepalen

Het kernprincipe wat we toepassen om bias te achterhalen is om de uitkomsten van de KI te vergelijken voor verschillende groepen. De *punten* waarop wordt vergeleken (de metrics) worden beschreven in onderdeel 2. De *groepen* die worden vergeleken zijn gebaseerd op kenmerken waarop we geen ongerechtvaardigd onderscheid mogen en/of willen maken. We hebben de volgende kenmerken geïdentificeerd waarvoor dit het geval is:

- Nationaliteit*
- Geboorteland*
- Postcode*
- Huidskleur
- Etniciteit
- Geslacht*
- Leeftijd*
- Burgerlijke staat*
- Seksuele geaardheid
- Religie
- Politieke opvatting
- Verblijfsstatus*
- Zwangerschap
- Gezondheid
- Sociale klasse
- Genetica

De met een asterisk gemarkeerde kenmerken zijn bij de gemeente bekend over de mensen die, doordat zij wonen op een adres in de dataset, opgenomen zijn in de dataset.

8.2 Metrics selecteren passend bij het project

De KI kan op verschillende manieren beoordeeld worden, bijvoorbeeld:

- *Accuracy*: Percentage KI voorspellingen wat juist is.
- *Selection rate*: Percentage zaken wat door de KI aangewezen wordt als fraude.
- *False positive rate (FPR)*: Percentage van alle niet-fraudezaken dat (ten onrechte) door de KI als fraude wordt aangewezen. Zegt iets over de kans dat een zaak onnodig wordt onderzocht, daarom hoe lager, hoe beter.
- *True positive rate (TPR)*: Percentage van alle fraudezaken dat (terecht) door de KI als fraude wordt aangewezen. Zegt iets over de pakkans, daarom hoe hoger, hoe beter.
- *False negative rate (FNR)*: Percentage van alle fraudezaken dat (ten onrechte) door de KI niet als fraude wordt aangewezen. Complementair aan de TPR.
- *False discovery rate (FDR)*: Percentage van alle fraudevoorspellingen wat onjuist is.

Zulke cijfers noemen we *metrics*. Een metric kan worden uitgerekend voor de hele dataset of juist per groep, waarna we kunnen vergelijken of bepaalde groepen door de KI anders behandeld worden, waardoor zij mogelijk benadeeld worden. Het is doorgaans niet mogelijk om alle metrics tegelijkertijd te verbeteren, waardoor er met de specifieke projectcontext in gedachten keuzes gemaakt moeten worden. Zo zijn in medische toepassingen *false negatives* doorgaans ongewenst, omdat dat betekent dat er een ziekte over het hoofd wordt gezien, terwijl in andere gevallen *false positives* meer nadelen met zich meebrengen.

Om een beeld te geven van de hoeveelheid bias in de KI zijn de volgende metrics gekozen:

- *Disparate impact*: Verhouding tussen de selection rate van twee groepen.
- *Average odds difference*: Gemiddelde van het verschil tussen twee groepen in FPR en het verschil in TPR.
- *False discovery rate difference*: Verschil in FDR tussen twee groepen.
- *False positive rate difference*: Verschil in FPR tussen twee groepen.
- *Generalized entropy index*: Oorspronkelijk afkomstig uit onderzoek naar inkomensongelijkheid; geeft uitdrukking aan de ongelijkheid tussen groepen én binnen een groep.

8.3 Hypotheses opstellen over features die tot indirecte bias leiden

Dat een bepaald kenmerk niet als feature in de KI zit, betekent niet dat de KI hier geen bias tegen kan hebben. Een bekend voorbeeld is dat KI zonder nationaliteit als feature alsnog een bias op nationaliteit kan hebben doordat postcode *wel* in de KI zit en sterk samenhangt met nationaliteit. Voor alle features in de KI zijn daarom hypotheses opgesteld over mogelijke relaties met de gevoelige kenmerken. Als zo'n samenhang leidt tot een bias op het gevoelige kenmerk noemen we dat indirecte bias.

8.4 Directe bias en controlevariabelen analyseren

Met directe bias wordt bedoeld een bias veroorzaakt door het direct meenemen van een gevoelig kenmerk als feature in de KI. De gevoelige kenmerken die zijn overwogen voor de KI zijn geslacht en leeftijd. De exacte features die hieruit zijn afgeleid, zijn te vinden in VERWIJZING NAAR TEMPLATE

Gevoelige kenmerken die niet worden ingezet als feature kunnen worden gebruikt als controlevariabele. Dit betekent dat het gevoelige kenmerk *niet* in de KI wordt meegenomen, maar dat we er *wel* groepen mee maken om te zien of de metrics verschillen tussen deze groepen.

Het analyseren gebeurt als volgt:

1. Deel de dataset op in twee groepen op basis van een gevoelig kenmerk.
2. Bereken de bovengenoemde metrics voor beide groepen.
3. Bekijk of de metrics verschillen tussen de groepen.

Als de metrics niet (substantieel) van elkaar verschillen tussen groepen op basis van kenmerk X, dan kunnen we concluderen dat de KI geen (substantieel) verschil maakt op basis van kenmerk X. Is dit het geval, dan kunnen ook de hypotheses over indirecte bias betreffende dit kenmerk van de lijst worden gehaald.

Als de metrics wel (substantieel) van elkaar verschillen, dan is er mogelijk sprake van een ongerechtvaardigde bias. Deze resultaten worden verzameld en meegenomen naar onderdeel 7.

8.5 Indirecte bias analyseren

Met indirecte bias wordt bedoeld een bias veroorzaakt door het meenemen van een op het eerste gezicht onschadelijke feature die samenhangt met een gevoelig kenmerk.

Onderdeel 3 heeft geleid tot hypotheses over de features die indirecte bias zouden kunnen veroorzaken. Onderdeel 4 heeft deze lijst mogelijk verkleind, als al geconcludeerd kon worden dat de KI geen verschil maakt op basis van een gevoelig kenmerk.

Het analyseren gebeurt op de manier zoals in onderdeel 4 beschreven, met als enige aanpassing dat de groepen nu niet gebaseerd zijn op een gevoelig kenmerk, maar op een van de features die voorkomt op de lijst met indirecte-bias-hypotheses.

Als de metrics niet (substantieel) van elkaar verschillen tussen groepen op basis van feature X, dan kunnen we alle indirecte-bias-hypotheses voor feature X ontkrachten. Als de metrics wel (substantieel) van elkaar verschillen, dan is de hypothese mogelijk waar en is er mogelijk sprake van een ongerechtvaardigde bias. Deze resultaten worden verzameld en meegenomen naar onderdeel 7.

8.6 Bias op niet meetbare variabelen analyseren

Lang niet alle gevoelige kenmerken op de lijst in onderdeel 1 worden geregistreerd door de gemeente of zelfs de overheid in het algemeen.

8.7 Gevonden bias wegen en beoordelen met verantwoordelijke directie

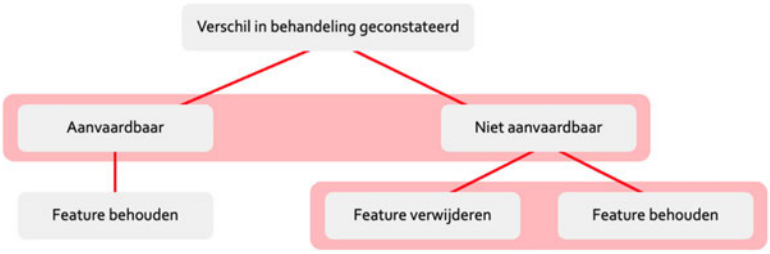
Het uitvoeren van de voorgaande stappen resulteert in een lijst van features met een indicatie van de mate waarin deze features bias veroorzaken tegen bepaalde groepen, uitgedrukt in de metrics. Vervolgens wordt iedere feature voorgelegd aan de verantwoordelijke directie.

Wanneer een verschil in behandeling is geconstateerd (oftewel: de metrics verschillen substantieel tussen twee groepen), is de eerste vraag of dit verschil aanvaardbaar is of onaanvaardbaar. Een verschil zou bijvoorbeeld aanvaardbaar kunnen zijn als er een achterliggend verband is tussen een derde kenmerk en de uitkomsten van de KI, dat het verschil verklaart, en dit achterliggende verband is niet problematisch. Bij het nemen van deze beslissingen zijn meerdere invalshoeken relevant, hieronder weergegeven. Het is daarom van belang dat er mensen vanuit verschillende rollen en achtergronden betrokken zijn bij het wegen van de uitkomsten van de bias analyse.



Wanneer een verschil aanvaardbaar wordt geacht, kan de feature behouden blijven. Wanneer een verschil niet aanvaardbaar wordt geacht, kan ervoor gekozen worden om de feature volledig uit de KI te verwijderen, of om te proberen de feature te behouden, waarover meer in onderdeel 8.

De overwegingen die gemaakt moeten worden zijn hieronder schematisch weergegeven.



8.8 Bias waar nodig mitigeren

De biases die niet aanvaardbaar worden geacht, worden in dit onderdeel aangepakt. De ene aanpak is het weglaten van features. De andere is om de bias te proberen te verminderen door het doen van aanpassingen in de KI (bv. de loss function of postprocessing) of de data (bv. herwegen voor een gebalanceerdere dataset).

Na dit onderdeel worden de onderdelen 4 t/m 7 altijd herhaald, omdat aanpassingen aan de KI invloed kunnen hebben op de uitkomsten van de bias analyse voor andere features. Bovendien is noch het weglaten van features noch het aanpassen van data of de KI een waterdichte oplossing (respectievelijk door indirecte bias en imperfecties in de methodes), dus ook de bias van de feature/gevoelige groep die is aangepast moet opnieuw bekeken worden.

8.9 Conclusies opstellen

Nadat alle onderdelen zijn doorlopen en in onderdeel 7 uiteindelijk alle overgebleven verschillen als aanvaardbaar beoordeeld worden, worden de conclusies van de bias analyse vastgelegd in en Rapport Bias Analyse. Dit omvat de oorspronkelijk gevonden en uiteindelijke biases, eventuele maatregelen die genomen zijn, en de overwegingen uit onderdeel 7.

8.10 Samenvatting rapport (in Engels)

Het rapport in het kader van de bias analyse is te vinden in bijlage 7.

Direct and indirect bias have been analyzed, obtaining useful insights regarding the model and the data used by it.

The first important outcome is that investigating only cases created after a melding has been received, leads to a high bias in most of the categories checked. The processes used in the past to generate new cases, not relying only on the melding, were important to reduce the bias of the investigations.

Therefore, the creation of cases using other methodologies will resume, merging them with the cases opened after a melding is received, since this leads to no bias in the data.

The current model has been trained on data containing both types of cases, therefore it is not carrying bias. If in the future, the creation of cases through other methodologies other than

the melding would not resume, the dataset will be biased, and the model will probably learn the bias from it.

The second important outcome is that men were discriminated against when the model was trained using sex as a feature. The error committed by the model over cases where men were living in the apartment was bigger than the error when also women were living in the apartment. The outcome of this analysis combined with the low feature importance led to the removal of the feature sex.

The purpose of this bias analysis was to identify and ensure that model was not biased, focusing only on detecting bias in the algorithm and not in the underlying data. It would be interesting in the future to extend it by analyzing the cases opened by citizens' signals to un understand better the discriminations that the citizen signals carry.

8.11 Resultaat van delen bevindingen met directie Wonen

Dit zal op 25 oktober worden besproken met Wonen.

9. Technische documentatie[3]

9.1 Wat zijn de interacties tussen de KI-toepassing met hard- en software dat geen onderdeel uitmaakt van de KI-toepassing zelf?

De TOP-applicatie stuurt een zaaknummer naar het API-endpoint van de KI-toepassing, dat de output terug levert zoals beschreven in vraag 7.21. De KI-toepassing communiceert ook met de [redacted] voor logging. Deze verbinding wordt automatisch gemaakt doordat de KI-toepassing draait bij Datapunt. Tot slot interacteert de KI-toepassing ook met de databases waaruit de input data wordt opgehaald, zoals beschreven in hoofdstuk 6.

Zie bijlage 8 voor extra technische informatie.

9.2 Wat zijn de revelante soft- of firmware versies?

Er wordt een recente versie van python gebruikt (op moment van schrijven, 06/07/2021, python 3.7) en de nieuwste versies van gangbare machine learning packages. Een volledige lijst met de specificaties is te vinden in het pyproject.toml bestand in de Gitlab repository.

9.3 Op welke manier is het KI-systeem in gebruik genomen en door wie?

De KI is verpakt in een Docker-container en op die manier gedeployed op de infrastructuur van [redacted]. Dit is gedaan door de werkplaats Datagestuurde Aanpak Ondermijning.

9.4 Op welke hardware functioneert of zou de KI-toepassing moeten functioneren?

De KI-toepassing zou moeten functioneren op iedere actuele hardware met minstens een dual-core processor en 8GB memory.

9.5 Is er een gebruikersinstructie en (indien nodig) installatie instructie opgesteld?

Er is een gebruikershandleiding opgesteld.

9.6 Wat zijn, indien van toepassing, de vooraf bepaalde (toekomstige) wijzigingen in de KI-toepassing en diens prestaties?

- Regelmatige monitoring van prestaties en bias

[3] Zie Richtlijn European Approach for AI voor alle documentatievereisten.

- Verbeteringen aan het algoritme qua prestaties/bias
- Toevoegen van nieuwe data

9.7 Hoe wordt omgegaan met de onvoorspelbaarheid van het handelen van de KI?

De KI ondersteunt de business om de meest kansrijke zaken uit de werkvoorraad te prioriteren en te selecteren voor een buitendienstonderzoek. Daarmee wordt de KI onderdeel van de specifieke afweging van de business om al dan niet een buitendienstonderzoek te starten op een adres. Een medewerker kan zien waarop de statistische kans op illegaal vakantieverhuur is gebaseerd. De redenering achter deze voorspelling wordt getoond aan de medewerker. Daardoor is inzichtelijk waarop de KI de risico-inschatting van illegale vakantieverhuur heeft gebaseerd. Dit is wel van invloed op de betrokkene, in dit geval de mogelijke overtreder van de huisvestingsverordening, omdat de ene melding meer prioriteit kan krijgen dan een andere melding. Of er daadwerkelijk sprake is van illegaal vakantieverhuur wordt vastgesteld door de verantwoordelijke toezichthouders in samenwerking met de projecthandhaver. Daarvoor wordt een administratief- en buitendienstonderzoek uitgevoerd. Vervolgens wordt het dossier intensief besproken in een debriefing om tot een besluitvorming te kunnen overgaan. De KI neemt dus geen zelfstandige beslissingen of er al dan niet sprake is van illegale vakantieverhuur.

9.8 Welke maatregelen zijn genomen om de verklaarbaarheid te borgen?

- Er wordt gebruik gemaakt van een algoritme wat goed te begrijpen is.
- Er wordt gebruik gemaakt van zogenaamde 'SHAP values' om voor individuele voorspellingen uit te leggen welke features er een positief dan wel negatief effect op hebben gehad en hoe groot dit effect was. Dit wordt aan de medewerkers getoond binnen de TOP-applicatie.

9.9 Is er een overzicht van alle veranderingen die aan de KI-toepassing zijn aangebracht gedurende de levenscyclus?

Alle veranderingen worden gelogd in een zogenaamde 'change log' in de Gitlab repository.

9.10 Is er een beheerplan opgesteld?

Zie vraag. 12.1

10. Risicomanagement

10.1 Welke risicomanagement systeem wordt toegepast en welke instrumenten worden daarbij ingezet?

- *DPIA*
- *BIO-toets (wordt gedaan over het hoofdproces waarvan het algoritme een onderdeel is)*
- *Weging betrokken (grond)rechten en maatregelen (op basis van IAMA)*
- *Beoordeling features*
- *Bias analyse*
- *KIIA*
- *Document 'Taken en verantwoordelijkheden bij het gebruik van algoritmen en rekenmodellen door de gemeente Amsterdam'.*

10.2 Is er een risico analyse uitgevoerd over de toepassing van de KI?

Er is een risicoanalyse uitgevoerd op het (hoofd)proces waarin het algoritme wordt gebruikt: Illegale Vakantieverhuur. De risicoanalyse is niet specifiek op het algoritme zelf gedaan.

10.3 Is er een overzicht van de geïdentificeerde risico's?

Ja, zie uitkomst DPIA, BIO-toets en overzicht 'betrokken (grond)rechten toetsing. Er mist een lijst met specifieke risico's van het algoritme op het gebied van informatiebeveiliging.

10.4 Kunnen deze risico's worden geëlimineerd of worden gereduceerd?

Zie uitkomst BIO-toets en overzicht 'betrokken (grond)rechten toetsing. Hier is aangegeven in hoeverre de risico's kunnen worden geëlimineerd of gereduceerd. De geïdentificeerde risico's gaan veelal over het proces in het algemeen.

10.5 Is er een overzicht opgesteld van de getroffen of nog te treffen maatregelen?

Ja, zie overzicht 'betrokken (grond)rechten toetsing'. Specifiek voor informatiebeveiliging is er nog geen overzicht van getroffen of nog te treffen maatregelen.

10.6 Zijn potentiële veiligheidsrisico's vastgesteld van (andere) voorzienbare toepassingen van de technologie, met inbegrip van onbedoeld of kwaadwillig misbruik ervan? Is er een plan om deze risico's te beperken of te beheren?

Nee, deze zijn niet beschreven.

10.7 Is het acceptabel, gezien de risico's en de beoogde maatregelen, om de KI-toepassing in gebruik te nemen?

Gezien de getroffen maatregelen en de werking van de KI-toepassing is het acceptabel om de KI-toepassing in gebruik te nemen gedurende de Pilot.

Met opmerkingen [10.2.e]: @[10.2.e] Wie heeft dit besloten en op basis van wat?

10.8 ***Wordt het risicomanagementsysteem gedurende de levenscyclus van de KI-toepassing toegepast?***

De gehanteerde instrumenten zijn beschreven in het beheerplan en zullen periodiek moeten worden geactualiseerd en worden herzien door de business en het verantwoordelijke beheerteam.

10.9 ***Wordt het risicomanagementsysteem periodiek geëvalueerd?***

De instrumenten worden zelfstandig periodiek geëvalueerd. Dit wordt gedaan door de Chief Informatie Office. Deze instrumenten worden tevens voorgelegd ter vaststelling.

Met opmerkingen [10.2.e]: @[10.2.e] Weet jij over welk risicomanagementsysteem het hier gaat?

11. Informatiebeveiliging

11.1 Wordt loginformatie over de werking van het algoritme bewaard en toegankelijk gemaakt?

Analyseomgeving

Binnen de[10.2.g], waar het algoritme ontwikkeld wordt, vindt enige logging plaats over gebruikers en uitgevoerde acties. De logs zijn opvraagbaar bij de product owner van de[10.2.g].

Er zijn geen afspraken over monitoring en bewaartermijnen.

Productieomgeving

In de productieomgeving wordt er op verschillende plekken gelogd:

- *In[10.2.g] wordt gelogd of het model werkt, of het data geleverd krijgt uit de verschillende databases en of het in staat is een voorspelling te berekenen. Ook wordt gelogd wanneer de API bevraagd wordt om voorspellingen te genereren (meestal 1 á 2 keer per dag automatisch).*
- *[10.2.g] wordt gebruikt om een nieuwe versie van het model te deployen naar productie dan wel acceptatie. Hierin wordt gelogd welke versie wanneer gedeployd is.*
- *Versiegeschiedenis (eerdere broncodes van het model) staan in Gitlab.*
- *Het is onduidelijk op welke manier de toegang tot het model zelf in de productieomgeving wordt gelogd.*

Er zijn geen afspraken gemaakt over monitoring en bewaartermijnen.

11.2 Wordt gecontroleerd of toegangsrechten up-to-date zijn met betrekking tot de omgeving waarin het algoritme functioneert?

Toegangsrechten worden beschreven in het document Aanvraagprocedure toegang tot het algoritme *en* Autorisaties toegang tot en ontwikkeling en beheer van het algoritme.

Let op: De scope van het document is beperkt.

11.3 Worden toegangsrechten aangepast zodra er een uitdiensttreding of functiewijziging van een werknemer plaatsvindt?

Toegangsrechten worden beschreven in het document Aanvraagprocedure toegang tot het algoritme *en* Autorisaties toegang tot en ontwikkeling en beheer van het algoritme.

Let op: De scope van het document is beperkt.

11.4 Worden toegangsrechten uitgegeven door daarvoor bevoegde personen?

Toegangsrechten worden beschreven in het document Aanvraagprocedure toegang tot het algoritme *en* Autorisaties toegang tot en ontwikkeling en beheer van het algoritme.

Let op: De scope van het document is beperkt.

11.5 Wordt functievermenging voorkomen bij de toegang van gebruikers tot het algoritme?

Functies en taken worden omschreven in het document Autorisaties toegang tot en ontwikkeling en beheer van het algoritme.

Let op: De scope van het document is beperkt.

11.6 Wordt er gebruikgemaakt van generieke beheeraccounts? Staat het aantal beheeraccounts in logische verhouding tot de beheerders?

Er worden geen generieke beheeraccounts gebruikt.

Let op: Het gaat hier alleen om de gebruikersaccounts van de leden van de werkplaats Ondermijning. Administrators van de infrastructuur bij OIS hebben meer rechten.

11.7 Wordt er bij het inrichten van toegangsrechten van verschillende gebruikersgroepen/rollen gebruikgemaakt van naamgevingsconventies en systematiek?

De werkplaats maakt gebruik van de diensten/infrastructuur van directie OIS. OIS maakt geen gebruik van naamgevingsconventies en systematiek.

11.8 Worden er naamgevingsconventies gebruikt voor gebruikers en beheerders, zodat zij geïdentificeerd kunnen worden?

De werkplaats maakt gebruik van de diensten/infrastructuur van directie OIS. OIS maakt geen gebruik van naamgevingsconventies en systematiek.

11.9 Voeren beheerders werkzaamheden als beheerder en werkzaamheden als gewone gebruiker uit onder 2 verschillende gebruikersnamen?

Ja, er bestaan verschillende accounts met verschillende rechten. Deze worden omschreven in het document Autorisaties toegang tot en ontwikkeling en beheer van het algoritme. *Voor beheerstaken wordt een ander account gebruikt dan voor normale werkzaamheden.*

11.10 Hebben gebruikersaccounts (geen) directe toegang tot onderliggende componenten?

Nee.

Let op: Het gaat hier alleen om de gebruikersaccounts van de leden van de werkplaats Ondermijning. Administrators van de infrastructuur bij OIS hebben meer rechten.

11.11 Bestaat er een functiescheiding tussen aanvragen, autoriseren en verwerken van wijzigingen in gebruikersaccounts en toegangsrechten?

Ja, toegangsrechten worden beschreven in het document Aanvraagprocedure toegang tot het algoritme *en* Autorisaties toegang tot en ontwikkeling en beheer van het algoritme.

Let op: De scope van het document is beperkt.

11.12 Is het wachtwoordbeheer interactief en zijn de wachtwoorden van geschikte kwaliteit?

Onduidelijk.

11.13 Worden wijzigingen in de code van het algoritme op een gecontroleerde wijze uitgevoerd? Denk aan het testen en accorderen/autoriseren van wijzigingen.

Ja, maar een omschrijving van het proces ontbreekt.

11.14 Is het algoritme beveiligd, zodat er geen risico is op ongeautoriseerde toegang, wijziging, beschadiging en/of dataverlies?

Het model wordt ontwikkeld in de 10.2.g [redacted]*, waar een beperkt aantal personen toegang tot heeft (zie autorisatiematrix). Het getrainde model wordt vervolgens in productie gezet.*

Een omschrijving van het model in en de veiligheid van de productieomgeving mist.

11.15 Worden er back-ups van het algoritme gemaakt en kunnen het algoritme en de data hersteld worden?

Oudere versies van het model zijn te vinden in Gitlab. De gebruikte data staan in de zaaksystemen.

Er is geen proces voor back-up en herstel beschreven.

11.16 Is er sprake van security by design?

Voor een deel is er sprake van security by design binnen de werkplaats, maar ze is tevens afhankelijk van informatiebeveiligingsmaatregelen die zijn getroffen door ketenpartners. Denk aan directie OIS voor de infrastructuur en omgevingen waarin wordt gewerkt, directie Basisinformatie voor de levering van gegevens, directie Wonen voor de levering van gegevens en de gebruikers van het algoritme.

Er is lange tijd geen overzicht geweest van deze getroffen maatregelen waardoor de status op het gebied van informatiebeveiling nog niet volledig in kaart en beschreven is. Hier wordt nog aan gewerkt.

Er zijn binnen de gemeente nog maar weinig kaders op het gebied van de beveiliging van algoritmes, waardoor op basis van eigen inzichten en externe informatie al lerende maatregelen worden bedacht en getroffen. Deze maatregelen zijn niet op een centrale plek gedocumenteerd.

12. Beheer

12.1 Is er een beheerplan opgesteld?

Er wordt momenteel gewerkt aan een beheerplan. Hierover is contact tussen de Werkplaats en de IV-manager van Wonen. Het oorspronkelijke beheerplan wordt komende maanden afgestemd op het product machine learning algoritmes, zodat duidelijk wordt welke werkzaamheden en expertises noodzakelijk zijn in het kader van beheer. Het beoogde DevOps team van OIS is ook geïnformeerd over de ontwikkelingen en zal hier naar verwachting een rol in gaan spelen.

12.2 Is er voldoende capaciteit om de KI-toepassing operationeel te houden en wijzigingen door te voeren?

Gedurende de Pilot wel. Na de Pilot zal een DevOps team met voldoende capaciteit moeten worden uitgerust. Dit moet nog nader worden besproken.

12.3 Is er voldoende kennis aanwezig om de KI-toepassing operationeel te houden en wijzigingen door te voeren?

Gedurende de Pilot wel. Na de Pilot zal een DevOps team van voldoende expertise moeten worden voorzien. Dit moet nog nader worden besproken maar er zijn voldoende mogelijkheden.

12.4 Worden de werking van de KI-toepassing periodiek geëvalueerd?

Ja, dit wordt gedaan door de Product Owner, menselijk toezichthouder, gebruikers en het toekomstig DevOps-team.

13. Communicatie

13.1 Is er een gebruikersinstructie opgesteld?

Ja, er is een eerste versie opgesteld. Deze wordt geactualiseerd en in lijn gebracht met concept wet- en regelgeving van de Europese Commissie.

13.2 Zijn mechanismen ingesteld om gebruikers in kennis te stellen van de redenen en criteria achter de resultaten van het KI-systeem?

Er worden workshops gegeven en werkinstructies opgesteld om medewerker bewust te maken van de kansen en risico's van het gebruik van KI. Daarnaast wordt toegelicht hoe de uitkomst van het algoritme moet worden geïnterpreteerd. Hierbij moet worden opgemerkt dat per uitkomst (risico-indicatie van woonfraude) de belangrijkste features (variabelen) worden getoond die bepalend zijn geweest voor de risico indicatie. Eindgebruikers zijn betrokken bij evaluaties en doorontwikkeling van het algoritme.

13.3 Is duidelijk gemaakt richting gebruikers en betrokkenen wat het doel van het KI-systeem is en wie of wat er profijt kan hebben van het product of de dienst?

Ja, er worden workshops gegeven en werkinstructies opgesteld om dit toe te lichten.
Daarnaast wordt aan betrokkenen (burgers die te maken krijgen met toezicht en handhaving Wonen) in de privacystatement te kennen gegeven dat er gebruik wordt gemaakt van een KI, bij het prioriteren van de meldingen Woonfraude. Deze is op de website van de gemeente Amsterdam worden geplaatst. De KI is eveneens opgenomen in het verwerkingsregister en in het algoritmeregister. Er heeft tevens een interview plaatsgevonden met de Volkskrant. Er is een artikel gepubliceerd op 26 februari 2020.

13.4 Is inzichtelijk gemaakt welke kenmerken, beperkingen en potentiële tekortkomingen van het KI-systeem duidelijk kenbaar gemaakt.

Ja, er worden workshops gegeven en werkinstructies opgesteld om dit inzichtelijk te maken. Daarnaast zijn met de Product Owner en gebruikers de opgestelde features besproken. Daardoor is er kennis aanwezig bij de gebruikers over de kenmerken, beperkingen en potentiële tekortkomingen van de KI-toepassing.

14. Verantwoording

14.1 ***Hebt u mechanismen ingesteld die de controleerbaarheid van het systeem door interne en/of externe onafhankelijke actoren mogelijk maken, bijvoorbeeld in de vorm van waarborging van de traceerbaarheid en het bijhouden van de processen en resultaten van het KI-systeem?***

Ja, dit ontwikkeling van het algoritme is gedocumenteerd in Gitlab. Deze informatie is niet openbaar. Daarnaast zijn diverse (risico) analyses uitgevoerd die kunnen worden gedeeld met interne en/of externe onafhankelijke actoren. De bevindingen uit de Pilot worden vastgelegd in het zogenaamde 'validatieplan'.

14.2 ***Is er, naast de interne initiatieven of kaders voor het toezicht op ethiek en verantwoording, ook sprake van een vorm van externe begeleiding of zijn er ook controleprocessen ingesteld?***

Er zijn externe adviseurs ingehuurd om de KI te optimaliseren en voor adviesvragen. Er wordt ook een juridisch advies de Functionaris Gegevensbescherming en de Commissie Persoonsgegevens Amsterdam ingewonnen.

14.3 ***Hebt u een mechanisme ingesteld om relevante belangen en waarden die door het KI-systeem in het gedrang komen, alsook potentiële afwegingen daartussen, vast te stellen?***

Ja, er zijn meerdere (risico)analyses uitgevoerd om deze belangen en waarden in kaart te brengen. Om deze belangen en waarden te eerbiedigen is onderzocht welke maatregelen kunnen worden getroffen. Zie hiervoor het overzicht 'weging betrokken (grond)rechten vakantieverhuur modell'.

15. Afwegingen en conclusie

15.1 ***Is de toepassing proportioneel?***
De toepassing van de KI valt binnen de bestuursrechtelijke publieke taak: efficiënte uitoefening van een toezichthoudende taak in een branche / werkterrein waarin momenteel het aantal meldingen van overtreding het aantal 'beloopbare zaken' ver overstijgt. Hierbij is rekenschap gegeven van rechten als eerbiediging van de persoonlijke levenssfeer, recht op informatie over overheidsoptreden, transparantie omtrent verwerking van persoonsgegevens en dataminimalisatie. Tevens zijn de noodzakelijke technische en organisatorische maatregelen genomen ter beveiliging van de verwerkte persoonsgegevens bij het ontwikkelen van het algoritme. De impact op het individu en de maatschappij wordt met voldoende technische en organisatorische waarborgen beschermd en de beperking van enkele grondrechten valt binnen de (in de rechtspraak) ontwikkelde grenzen als vrijheid van informatie, en recht op een eerlijk proces.

15.2 ***Is met minder ingrijpende middelen hetzelfde doel te verwezenlijken?***
Het doel om een risico indicatie van woonfraude op een adres te genereren is niet op een andere manier mogelijk. Dit type algoritme is passend voor dit doel. Dat betekent dat complexere algoritmes bewust niet zijn toegepast.

15.3 **Gaat het om positive sum in plaats van zero sum?**
Het betreft een positive sum. Bij een juist werkend algoritme zal de capaciteit van de beoordelaars aanvragen effectiever worden benut. Een positief effect voor de burger (een mogelijk overtreder) is dat deze minder bezoek van een toezichthouder heeft te verwachten, indien de berekend is dat de kans op een overtreding relatief klein is. Dat zal als positief ervaren worden, aangezien bezoek van een toezichthouder Wonen aan een woning in de regel als belastend ervaren wordt.

16.4 ***Zijn er rest-risico's?***
Het algoritme richt zich op kenmerken van de woning en niet van diens bewoner of eigenaar. Daarmee kan worden gesteld dat er geen reële rest-risico's zijn.

Er zijn wel risico's van politieke aard. Er bestaat een reëel risico dat de Amsterdamse bevolking grote bezwaren heeft met de inzet van een algoritme wat woonfraude kan voorspellen. Dit bezwaar zal zich kunnen manifesteren via media-aandacht, de politiek of bezwaar- en beroep procedures. Het is daarom van belang om transparant te zijn richting betrokkenen, de wethouder en het college van B&W volledig te informeren over de materie. Er is een collegebesluit gegeven over toepassing van het algoritme in de bedrijfsvoering.

16.5 ***Hoe wordt verantwoordelijkheid genomen voor verder gebruik?***
Er ligt een politiek bestuurlijke verantwoordelijkheid voor het gebruik van het algoritme. Hoe dit concreet vorm gaat krijgen, zal nog moeten worden uitgekristalliseerd

Gemeente Amsterdam

Bijlage 1: Overzicht rollen en verantwoordelijkheden

Directie Wonen	
Rol	***Verantwoordelijkheden***
Wethouder Wonen (bestuurlijk-politiek verantwoordelijk)	- De betrouwbaarheid van het algoritme die worden toegepast door de directie; - Het ter besluitvorming aan het college voorleggen van algoritmen met een hoog risico.
Directeur Wonen (ambtelijk verantwoordelijk	- Verantwoordelijk voor de betrouwbaarheid van algoritmen die ontwikkeld - Eindverantwoordelijk en eigenaar van KI-toepassing; - Bepalen al dan niet in gebruik nemen van de KI-toepassing (MT-besluit); - Informeren verantwoordelijk Wethouder/ College; - Geeft akkoord voor het verwerken van de benodigde data; - Een akkoord geven op de door het ontwikkel-team of opdrachtnemer opgestelde acceptatiecriteria (ofwel 'definition of done') volgens de vastgestelde standaard voor de start van het ontwikkelproces (zie: 'Ontwikkel-team').
Afdelingshoofd Woonruimteverdeling en Gebruik	- (Mede)verantwoordelijk/ eigenaar KI vanuit mandaatbesluit; - Akkoord geven voor werking KI-toepassing; - Vaststellen of afkeuren assessments en templates; - Geeft akkoord voor het verwerken van de benodigde data; - Een akkoord geven op de door het ontwikkel-team of opdrachtnemer opgestelde acceptatiecriteria (ofwel 'definition of done') volgens de vastgestelde standaard voor de start van het ontwikkelproces (zie: 'Ontwikkel-team');
Product Owner/ proceseigenaar	- Voert taken en verantwoordelijkheden uit die zijn opgedragen vanuit verantwoordelijk directeur en/of afdelingshoofd; - Vertalen van de bedrijfsbehoefte naar requirements voor KI; - Duiden van de gebruikersbehoefte bij gebruik KI; - Aansturen betrokken medewerkers; - Aanleveren vereiste documentatie;

	- Inwinnen adviezen.
Beleidsadviseur	- Maakt de vertaling tussen werking KI en (op te stellen) beleid (indien nodig); - Adviseert over de werking van de KI-toepassing; - Neemt deel aan weging en beoordeling werking KI.
Handhavingsjurist	- Adviseert of de werking van de KI binnen de bevoegdheden van de wettelijke taak blijft; - Ondersteund bij het opstellen van de vereiste documentatie; - Analyseert de betrokken (grond)rechten.
Senior Medewerker (planner/ toezichthouder)	- Adviseert of de werking van de KI passend is bij het ondersteunen van de wettelijk taak; - Geeft de gebruikersbehoefte weer; - Toekomstig gebruiker van de KI-toepassing.
Privacy jurist	- Toetst in hoeverre wordt voldaan aan de bepalingen uit de AVG; - Adviseert over welke en hoe data moet worden verwerkt; - Ondersteund bij het opstellen van de vereiste documentatie; - Zorgt voor het inwinnen van advies bij de Privacy Officers, FG en CPA (indien nodig).
Security officer	- Zorgt van de vertaling van gesignaleerde risico's naar te treffen (technische) maatregelen, zodat er sprake is van een veilig en robuust product; - Adviseert over een veilige levering van de benodigde data; - Ondersteund bij het opstellen van de vereiste documentatie. - Zorgt voor het inwinnen van advies bij de CISO (indien nodig)
Data Analist/ Engineer	- Zorgt voor de datalevering en pseudonimiseert/anonimiseert/versleuteld waar mogelijk; - Geeft inzicht over de datakwaliteit; - Geeft inzicht over de betekenis van de beschikbare data;
Architect	- Geeft advies aan het ontwikkelteam van de KI-toepassing hoe de KI-toepassing in lijn met Stedelijke Architectuurbeginselen moet worden ontwikkeld en beschikbaar gesteld;

	- Geeft advies aan het ontwikkelteam van de KI-toepassing hoe de KI-toepassing binnen het applicatielandschap van Directie Wonen moet worden ingepast.
Communicatieadviseur	- Adviseert over de communicatie naar betrokkenen, de media, de wethouder en het College B&W.

Werkplaats DAO	
Rol	***Taken en verantwoordelijkheden***
Product Owner	- Verantwoordelijk voor het maximaliseren van de waarde het het product; - Stelt de acceptatiecriteria en prioriteiten vast voor het ontwikkelteam op basis van de gebruikers- en bedrijfswensen van de opdrachtgever; - Aanspreekpunt voor de Product Owner opdrachtgevende directie; - Zorgt ervoor dat voldoende resources beschikbaar zijn; - Eindverantwoordelijk voor alle door het ontwikkelteam opgestelde en aangeleverde documentatie ten behoeve van de ontwikkeling van het product. - Verlenen en intrekken van (toegangs)rechten voor Data Engineer en Data Scientist voor werken in 10.2.g in samenwerking met Datapunt. - Controleert periodiek de rechten
Business Analist	- Ondersteunt de opdrachtgever en product owner bij het inzichtelijk maken van de gebruikers- en bedrijfswensen; - Ondersteunt bij het opstellen van de vereiste documentatie; - Ondersteunen en begeleiden van de verschillende fases in de ontwikkeling van producten - Beschrijven en analyseren van de relevante processen, systemen en data in de huidige en gewenste situatie. - Het ophalen, analyseren en vertalen van de business doelstellingen, uitdagingen en behoeften naar potentiële oplossingen, door bijvoorbeeld idee generatie workshops; - Ophalen van feedback op bij de gebruikers en vertaalt dit naar concrete mogelijke vervolgstappen voor het team;
Data Engineer	- Ontvangt de geleverd data en zorgt ervoor dat deze data wordt 'geprepareerd' (o.a. gepseudonimiseerd), zodat ermee kan worden gewerkt; - Is bevoegd om autorisaties te verlenen/ rechtenstructuur aan te passen van de database;

	- Verantwoordelijk voor model deployment.
Data Scientist	- Ontwikkelt, traint en test de KI; - Voert het onderzoek op onbewuste vooringenomenheid uit; - Ondersteunt bij het opstellen van de vereiste (technische) documentatie; - Tevens beheerder van de ontwikkelde producten. Zorg ervoor dat deze operationeel blijven totdat is besloten om het product al dan niet te implementeren in de bedrijfsvoering.
Security Officer	- Adviseert het ontwikkelteam op het gebied van informatiebeveiliging ten behoeve van het ontwikkelen van de KI; - Zorgt voor de vertaling van gesignaleerde risico's op het gebied van informatiebeveiliging naar te treffen (technische) maatregelen, - Ondersteunt bij het opstellen van de vereiste documentatie; - Ziet erop toe dat de maatregelen uit de BIO worden nageleefd en geïmplementeerd.
Jurist	- Voorziet het ontwikkelteam van juridisch advies ten behoeve van het ontwikkelen van de KI; - Zorgt voor de vertaling van gesignaleerde (juridische) risico's naar te treffen (technische) maatregelen; - Adviseert over de manier van het verwerken van data; - Ondersteunt bij het opstellen van de vereiste documentatie; - Ziet erop toe dat de beschreven maatregelen worden nageleefd; - Controleert periodiek de verwerkte en beschikbare data.

Directie Basisinformatie	
Rol	***Verantwoordelijkheden en bevoegdheden***
Directeur	- Formeel gegevensverstrekker van basisregistraties
Adviseur gegevensmanager	- Het maken van stelselafspraken, het toezien op de naleving van deze stelselafspraken - Verifieert of de afnemer bevoegd is de gevraagde gegevens te ontvangen volgens een autorisatieproces
Data Engineer	- Zorgt voor de datalevering en pseudonimiseert/anonimiseert/versleuteld waar mogelijk; - Geeft inzicht over de datakwaliteit;

	- Geeft inzicht over de betekenis van de beschikbare data;
Privacy jurist	- Toetst in hoeverre wordt voldaan aan de bepalingen uit de AVG; - Adviseert de gegevensverstrekker over de data die wordt verwerkt;
Security officer	- Adviseert de gegevensverstrekker over een veilige levering en gebruik van de benodigde data;

Directie Data en Informatie (OIS)	
Rol	***Verantwoordelijkheden en bevoegdheden***
Product Owner	- Leveren van een [redacted]
Beheerder	- Operationeel houden van de [redacted] 10.2.g en treft beheermaatregelen