DATA PROTECTION IMPACT ASSESSMENT (DPIA)

Als gemeente zijn we verplicht een DPIA uit te voeren als er sprake is van een verhoogd risico op een inbreuk op de persoonlijke levenssfeer van inwoners. De DPIA vul je aan het begin van je project in. Hierdoor kunnen maatregelen voor het beschermen van data en het voorkomen van risico's vooraf meegenomen worden in de afweging om tot de verwerking over te gaan of waar mogelijk meegenomen worden in het ontwerp op programma van eisen van de aan te schaffen applicatie.

De DPIA vul je samen met je I-adviseur in. De DPIA is onderdeel van de intakeprocedure van het afstemmingsoverleg.

Project: Opslag van analysedata in een datawarehouse
Afdeling: PIF
Verantwoordelijke: Tom Merkx
I-adviseur: Arjen Verhulst
Datum: 3-10-2022

1. Inleiding op het project

1.1 Wat zijn de algemene achtergronden van het project?

Als gemeente Nijmegen willen we steeds meer datagestuurd werken. Daarvoor is het nodig dat data eenvoudig beschikbaar is om te kunnen analyseren. Om dat mogelijk te maken, hebben we in Nijmegen een datawarehouse opgebouwd. Dat is een grootschalige verwerking van veel verschillende soorten (privacygevoelige) data. Om de verantwoordelijkheid en privacybescherming goed te waarborgen, toetsen we deze werkwijze door middel van deze DPIA.
Om grip te houden op de gegevenshuishouding en privacy is het belangrijk om over levering en gebruik van gegevens duidelijke afspraken te maken tussen (gegevens)leverancier en afnemer. De vastlegging van deze afspraken doen we in de vorm van een Gegevensleveringsovereenkomst (GLO). In de GLO wordt vastgelegd welke data overgehaald wordt, de reden waarom en hoe de uitwisseling technisch in elkaar zit.

In de praktijk komt het erop neer dat voor leveringen een set algemeen geldende kaders van toepassing is; deze kunnen de vorm van Algemene Voorwaarden aannemen (bijvoorbeeld: nooit doorleveren zonder toestemming van leverancier). Specifieke afspraken worden dan apart vastgelegd in de gegevensleveringsovereenkomst (GLO). Ten behoeve van het datawarehouse gaan we ook werken met GLO's, waarin doelbinding en rechtmatigheid wederkerig worden vastgelegd. Deze geven, in combinatie met de DPIA's van de afzonderlijke rapportages, een compleet overzicht van de data die in het datawarehouse verwerkt worden, en waarvoor deze verwerkt worden. Zie Bijlage 2 voor een visuele toelichting.

Daarnaast vergt de gegevensverzameling *an sich* in het hele datawarehouse een grondslag, o.b.v. artikel 6 van de AVG.

1.1 Welk probleem lost deze gegevensverwerking op?

We willen datagestuurd werken om onze werk- en beleidsprocessen beter, sneller en actueler te kunnen beoordelen en verbeteren.
Het organiseren van een centrale plek waar data verzameld en verwerkt worden, maakt het ook mogelijk een algemene 'data-waarheid' te creëren. Analyses worden niet langer gedaan op verschillende, van elkaar afwijkende bronnen. In plaats daarvan zorgt het datawarehouse voor één plek met kwalitatieve, gevalideerde data.

Als data niet – of slecht – voorhanden zijn, blijft het moeilijk de conclusies daaruit beschikbaar te stellen binnen de organisatie. Dat maakt het principe van datagestuurd werken – het gebruik van data in onze processen – nagenoeg onuitvoerbaar.

2. Doel en grondslag

2.1 Wat is het doel van de beoogde verwerking?

Het doel is een eenduidige gegevensverzameling op te bouwen, die ons in staat stelt onze beleids- en werkprocessen datagestuurd op te pakken.

2.2 Op welke grondslag is de verwerking gebaseerd?

Het DWH heeft geen eenduidige en duidelijke grondslag, want:
• de gemeente heeft geen toestemming van de betrokkenen gevraagd.
• er is geen overeenkomst naar privaatrecht met betrokkenen.
• een vitaal belang voor alle individuele inwoners is niet aanwezig.
• er is geen duidelijke wettelijke basis.
Er zijn echter wel aanknopingspunten bij de grondslag van "algemeen belang". Hoewel niet elke gemeente een datawarehouse heeft, vinden wij het goed verdedigbaar dat de grondslag voor het inrichten en gebruiken van een datawarehouse te vinden is in:

- Artikel 213a Gemeentewet
    - Het college verricht periodiek onderzoek naar de doelmatigheid en de doeltreffendheid van het door hem gevoerde bestuur.
- Artikel 6, lid 4 van de AVG, Jo artikel 23 van de AVG, mits het primaire wettelijke kader geen verbod heeft tot verdere verwerking:
  Wanneer de verwerking voor een ander doel dan dat waarvoor de persoonsgegevens zijn verzameld niet berust op toestemming van de betrokkene of op een Unierechtelijke bepaling of een lidstaatrechtelijke bepaling die in een democratische samenleving een noodzakelijke en evenredige maatregel vormt ter waarborging van de in artikel 23, lid 1, bedoelde doelstellingen houdt de verwerkingsverantwoordelijke bij de beoordeling van de vraag of de verwerking voor een ander doel verenigbaar is met het doel waarvoor de persoonsgegevens aanvankelijk zijn verzameld onder meer rekening met:
  ieder verband tussen de doeleinden waarvoor de persoonsgegevens zijn verzameld, en de doeleinden van de voorgenomen verdere verwerking;
  b) het kader waarin de persoonsgegevens zijn verzameld, met name wat de verhouding tussen de betrokkenen en de verwerkingsverantwoordelijke betreft;

c) de aard van de persoonsgegevens, met name of bijzondere categorieën van persoonsgegevens worden verwerkt, overeenkomstig artikel 9, en of persoonsgegevens over strafrechtelijke veroordelingen en strafbare feiten worden verwerkt, overeenkomstig artikel 10;
d) de mogelijke gevolgen van de voorgenomen verdere verwerking voor de betrokkenen;
e) het bestaan van passende waarborgen, waaronder eventueel versleuteling of pseudonimisering.

- Artikel 89 van de AVG
    - Waarborgen en afwijkingen in verband met verwerking met het oog op archivering in het algemeen belang, wetenschappelijk of historisch onderzoek of **statistische doeleinden**.
    - Gegevens in het datawarehouse worden verwerkt voor statistische doeleinden; het doen van onderzoek. Dat maakt het mogelijk wel op grote schaal persoonsgegevens te verwerken, mits de beveiliging en doelbinding van die gegevens goed op orde is.

3. Achtergrond

3.1 Welke besluitvorming heeft plaatsgevonden op dit onderwerp?

Het gebruik van een datawarehouse en de bijbehorende risico's worden voorgelegd aan het College van B&W, om het restrisico goed te beleggen.

De ambtelijke verantwoordelijkheid voor het datawarehouse ligt bij het Netwerk-MT Data. De voorzitter daarvan is vanuit het GMT gemandateerd voor het thema datagedreven werken, waarvan het datawarehouse een onderdeel is.

3.2 Zijn er eerder DPIA's uitgevoerd op dit onderwerp?

Er is in 2018 een DPIA opgesteld door externe auditor BKBO, zie Bijlage 3.
Daaruit zijn een aantal aanbevelingen voortgevloeid, waarvan in Bijlage 4 aangegeven is in hoeverre deze toegepast zijn of nog worden.

3.3 Is er sprake van stelselmatige monitoring?

Ja.

3.4 Is er sprake van matching of samenvoeging van datasets?

Ja.

3.5 Is er sprake van innovatief gebruik van data?

Ja. De verzameling van gegevens stelt ons in staat gebruik te maken van nieuwe technieken om data in te zetten, zoals voorspellende modellen en diepgaande analyses. Automatische besluitvorming is niet aan de orde, de data die uit deze analyses komen dienen als basis voor een menselijk besluit.

4. Gegevens

4.1 Gaat het om een eenmalige uitwisseling van persoonsgegevens of structureel?

Het betreft een structurele uitwisseling van (persoons)gegevens.

4.2 Welke persoonsgegevens worden verwerkt?

Er worden zowel 'persoonsgegevens' als 'bijzondere persoonsgegevens' verwerkt.
Daaronder vallen onder andere:

- NAW-gegevens van inwoners van Nijmegen
- Burgerservicenummers
- Cliëntgegevens over zorg- en inkomensvoorzieningen
    - Clientnummer
    - NAW
    - Soort voorziening verstrekt
    - Datums van- en tot-
- Belastinggegevens zoals bezwaren en aanslagen
    - Adres van aanslag/bezwaar
    - Bedrag
- Medewerker- en klantgegevens uit de backoffice van het KlantContactCentrum
    - Medewerkernaam
    - Onderwerp van contactmoment
- Gegevens over onroerende zaken uit de basisregistratie Kadaster en de basisregistratie WOZ
    - NAW gegevens eigenaar(s)
    - Waarde van het object
    - Soort object

Deze lijst is niet eindig, en wordt bij uitbreiding van het datawarehouse mogelijk langer. Een volledig overzicht van verwerkte gegevens is te vinden in de opgestelde Gegevensleveringsovereenkomsten en DPIA's.

4.3 Zijn dit bijzondere persoonsgegevens?

Ja. Het betreft (o.a.) gegevens over etnische afkomst (BRP) en gegevens over gezondheid (WMO).

4.4 Is de beoogde verwerking proportioneel?

De verwerking stelt de gemeente in staat haar processen beter, sneller en actueler te evalueren en aan te passen.

De gegevens worden op persoonsniveau overgehaald van het bronsysteem naar het datawarehouse.
Alleen gegevens die nodig zijn voor analyses worden overgehaald. Ook wordt rekening gehouden met wettelijke mogelijkheden die gelden voor de bron, wat verdere verspreiding en gebruik van de gegevens betreft.
Het is noodzakelijk de gegevens inhoudelijk en technisch te kunnen valideren, om te voorkomen dat gebruik wordt gemaakt van foutieve gegevens. Daardoor is het niet mogelijk gegevens meteen al te anonimiseren.
Daarom is er een splitsing aangebracht in het datawarehouse. De gegevens worden eerst overgehaald naar de zogeheten '*Load1*', waar slechts een beperkte groep beheerders toegang toe heeft.

Op basis van die gegevens wordt een '*Load2'* opgesteld, die slechts de gegevens bevat die nodig zijn voor een specifieke vraag. Deze *Load2* wordt ontdaan van herleidbare kenmerken, of gepseudonimiseerd wanneer dat niet mogelijk is. Daarmee bevat de Load2 geen herleidbare persoonsgegevens meer. Deze *Load2* bevat de datasets die uiteindelijk voor analyses gebruikt worden.

Dataleveringen van bron naar datawarehouse ('*Load1'*) worden vastgelegd in een gegevensleveringsovereenkomst (GLO). Deze GLO wordt opgebouwd uit de onderliggende datasets ('*Load2's'*) en hun (eventueel) bijbehorende DPIA's. Daarmee wordt getoetst of de verwerking van een dataset proportioneel is, en daarmee of de GLO de juiste doelbinding heeft.
Hier bestaat een risico dat een GLO en de onderliggende DPIA's verouderen en niet meer de actuele situatie beschrijven; zie risico 8.3.III.

4.5 Is de beoogde verwerking subsidiair?
Het is ook mogelijk voor elke rapportage of analyse de data afzonderlijk te vergaren, in plaats van te verzamelen in één datawarehouse. Dit heeft echter een aantal nadelen:

- Elk bronsysteem wordt meerdere keren bevraagd, waardoor dat mogelijk overbelast wordt. Dat heeft negatieve gevolgen voor de operationele processen.
- Eenzelfde bewerking wordt veelvuldig uitgevoerd: de verwerking van data richting het datawarehouse doorloopt een aantal stappen, die dan voor elke rapportage uitgevoerd moeten worden. Dit zorgt voor een grote extra belasting op de bron en het verwerkingsproces.
- Elke afzonderlijke verbinding tussen bron en rapportage moet dan ook beheerd worden, wat extra capaciteit kost.

Het op 1 plek organiseren van de technische koppeling tussen bron en analyse heeft daarmee een aantal voordelen op het gebied van capaciteit en belasting van systemen.

Daarbij fungeert datawarehouse dan ook als enige bron voor data. Analyses worden niet langer gedaan op verschillende, van elkaar afwijkende bronnen. In plaats daarvan zorgt het datawarehouse voor één plek met kwalitatieve, gevalideerde data.

Gemeente Nijmegen wil waar mogelijk aansluiten bij het landelijke Common Ground gedachtegoed, een set aan informatiearchitectuurprincipes. Een van deze principes is dat data zoveel mogelijk wordt opgehaald bij de bron, via services.
Veel applicaties en services zijn er (technisch) nog niet op ingericht om grote hoeveelheden data ook zo te kunnen ontsluiten. Dit principe is dus nog niet toepasbaar op het uitvoeren van analyses waarmee het inzetten van een datawarehouse een noodzakelijke, tijdelijke tussenstap is. In de toekomst zullen meer applicaties en processen hun data wel rechtstreeks kunnen ontsluiten, dat ze ook bruikbaar zijn voor analyses. Daar zullen we dan ook op overstappen.

4.4 Worden de persoonsgegevens buiten de EER gebruikt?
Nee.

5. Partijen

5.1 Welke partijen zijn bij de verwerking betrokken?

De gegevens worden uitgewisseld tussen verschillende softwareleveranciers van SaaS- of *on-premise* toepassingen, de gemeente Nijmegen en de iRvN (als beheerder van de database waarin het datawarehouse staat).

5.2 Wat zijn ieders rollen?

De gemeente Nijmegen is eigenaar van de gegevens, of deze nu uit een SaaS-oplossing komen of lokaal staan. Daarmee is gemeente Nijmegen bronhouder en verwerkingsverantwoordelijke.

De iRvN treedt in alle gevallen op als verwerker, in opdracht van gemeente Nijmegen. Er is een verwerkersovereenkomst opgesteld tussen de IRvN en gemeente Nijmegen.

5.3 Welke overeenkomsten worden aangegaan als gevolg van die rollen?

Voor elke levering van gegevens tussen bron en datawarehouse, wordt een GLO opgesteld tussen de bronhouder en het Datahuis. Voor elke dataset in Load2 en elke rapportage wordt de doelmatigheid getoetst middels een DPIA.

De verwerkersovereenkomst met de IRvN, zoals hierboven beschreven.

6. Zorgplicht

6.1 Op welke wijze is het project en haar analyses in begrijpelijke taal uit te leggen?

“Om slimmere en betere keuzes te maken over de stad, wil gemeente Nijmegen meer objectieve data gebruiken. Daarvoor is een technische tussenstap nodig; het opslaan van veel gevoelige gegevens in een speciale omgeving. Deze omgeving beveiligen we zo goed mogelijk[1]. Deze gegevens worden bij analyses anoniem en onherleidbaar gemaakt.”

6.2 Op welke wijze(n) worden betrokkenen geïnformeerd?

De manier waarop gemeente Nijmegen met haar data omgaat, wordt vastgelegd in de aankomende datastrategie. Deze zal eind 2022 verschijnen, en ingaan op de verschillende waarborgen die we nemen om verantwoord met de data van de inwoners van Nijmegen om te gaan.

Ook wordt het gebruik van de gegevens in afzonderlijke rapportages wordt vastgelegd in individuele DPIA’s, waarvan een overzicht in het openbare verwerkingsregister opgenomen wordt.

Het werken met een datawarehouse wordt ook opgenomen in het privacystatement van de gemeente Nijmegen, eind 2022.

6.3 Op welke manier zijn de rechten van betrokkenen geborgd in het proces?

Het datawarehouse is geen bronsysteem, maar wordt gevoed met data uit verschillende applicaties. In deze applicaties worden de rechten van betrokkenen, zoals het recht op inzage en recht op vergetelheid, geborgd.

Het datawarehouse volgt deze bronnen, en bevat alleen die gegevens die in de bronapplicaties staan.

[1] Het Informatiebeveiligingsbeleid van de gemeente Nijmegen: https://lokaleregelgeving.overheid.nl/CVDR628813

Het datawarehouse verhoogt het gemak waarmee burgers inzicht kunnen krijgen in welke gegevens van hen verwerkt worden en voor welke doeleinden, omdat deze informatie samenkomt in het datawarehouse. Zie ook risico 8.3.VIII.

7. Informatiebeveiliging en vernietiging

7.1 Hoe lang worden de gegevens bewaard?

Het datawarehouse volgt hierbij de bronapplicaties. De 'ruwe' data daaruit, blijven opgeslagen in de Load1, maar niet langer dan ze in de bronapplicatie opgeslagen worden. Daarmee blijven ze beschikbaar ter controle en validatie, en om datasets in Load2 te maken.
De datasets in Load2 bevatten alleen de gegevens die nodig zijn om een vraag te beantwoorden, en gaan dan ook niet verder terug in de tijd dan nodig.
Wanneer een analyse niet langer nodig is, of een vraag niet langer relevant, worden de bijbehorende gegevens ook weer uit het datawarehouse verwijderd.

Hierop geldt een uitzondering bij het overstappen naar een nieuwe bronapplicatie. Het is niet altijd mogelijk om de data uit de 'oude' bronapplicatie over te hevelen naar de vervanger. Het datawarehouse maakt het dan mogelijk systeemoverstijgende analyses te maken; in het datawarehouse komen dan data van zowel het oude als het nieuwe systeem samen. Hierbij volgen we wel de uiterlijke bewaartermijnen van het nieuwe systeem; we kijken niet verder terug.

7.2 Wie is verantwoordelijk voor de vernietiging van de gegevens?

De ambtelijke verantwoordelijkheid van de gegevensverzameling ligt bij de voorzitter van het Netwerk-MT Data, een concernmanager die vanuit het GMT gemandateerd is op het onderwerp Data. De praktische uitvoering van het beleid en het toezien daarop vallen onder verantwoordelijkheid van de CDO.

7.3 Is er een vastgesteld normenkader van toepassing op deze gegevensverwerking?

Nee.

7.4 Hoe worden de gegevens beveiligd? Hoe is het toezicht daarop georganiseerd?

De gegevens vallen onder beheer van de IRVN, in een on-premise database. De IRVN volgt de BIO-standaarden, waarmee deze ook van toepassing zijn op het datawarehouse.
De toegang tot de verzameling met gevoeligste gegevens is beperkt tot een zestal data-engineers en de datawarehouse-beheerder, die allen een VOG[2] (met screeningsprofiel 'Informatie') overlegd hebben. Zij kunnen alleen bij deze gegevens middels een beveiligde (Citrix-)verbinding. Hierop wordt toegezien door de beheerder van het datawarehouse.
Gebruikers maken op persoonlijke accounts gebruik van gegevens in het datawarehouse, waardoor de toegang tot gegevens goed te beperken is.
Wachtwoorden van die gebruikersaccounts veranderen driemaandelijks, waarmee onterechte toegang (door functiewijzing bijvoorbeeld) beperkt blijft.
Voor automatische processen worden generieke wachtwoorden gebruikt. Die processen komen echter pas tot stand na uitvoerige ontwikkeling en testen, wat via het persoonlijke account loopt.
Deze automatische processen resulteren in datasets die worden gebruikt voor rapportages, die op hun beurt weer strenge autorisatie kennen.

[2] Het toezicht op deze VOG's is binnen onze organisatie nog niet goed geregeld, zie risico VII.

Bij uitdiensttreding wordt de toegang tot het datawarehouse, samen met de toegang tot Citrix en andere accounts, automatisch opgezegd.

8. Risico's

8.1 Benoem hieronder de risico's voor betrokkenen door de gegevensverwerking

Iedere verwerking kan, ondanks genomen maatregelen, risico's met zich meebrengen. Het is belangrijk dat we als organisatie deze risico's in beeld hebben, zodat we besluiten kunnen nemen met inachtneming van eventuele risico's. Het gaat hier zowel om beveiligingsrisico's als andere privacyrisico's. Risico is geformuleerd in termen van de waarschijnlijkheid dat zich het risico voordoet (kans) afgezet tegen de hoeveelheid schade of gevolgen die het risico kan hebben (de impact). Kortweg: risico = kans x impact. In basis volgt dit het volgende schema:

	Risico impact		
Risico kans	Klein	Midden	Groot
Klein			
Midden			
Groot			

8.2 Benoem de maatregelen die nog noodzakelijk zijn.

Het gaat hier om maatregelen om de gevolgen te voorkomen of de effecten van die risico's te mitigeren, die nog niet genomen zijn. Beschrijf ook de impact van deze maatregelen, op het doel van de gegevensverwerking en de kosten van deze maatregelen.

8.3 Wat is het restrisico dat overblijft?

Dit is het risico dat overblijft na uitvoering van de geadviseerde maatregel(en).

I. Risico: Gebruik van gegevens zonder duidelijke aanleiding en/of gerechtvaardigd doel.

De grootschaligheid van het datawarehouse maakt het een uitdaging om alle verwerkingen die erin plaatsvinden bij te houden en te toetsen op rechtmatigheid.
Om dit toch goed en zorgvuldig te kunnen doen, beogen we enerzijds met een stelsel van gegevensleveringsovereenkomsten (GLO's) en DPIA's in kaart te brengen welke gegevens voor welk doeleinde gebruikt worden. Anderzijds moeten we constateren dat er geen dekkende grondslag is voor het verwerken van gegevens in een datawarehouse (zie risico II).
De data stromen van bron naar rapportage, de doelbinding loopt in tegengestelde richting. Het begint met een vraag, waarvoor gegevens nodig zijn ter beantwoording. Pas als die vraag is getoetst op doelbinding middels een DPIA, worden de benodigde gegevens klaargezet in het datawarehouse.
Er worden niet meer gegevens naar het datawarehouse gehaald, dan nodig is voor geïnventariseerde analyses. De data die – vastgelegd in een GLO – naar het datawarehouse worden gehaald, bestaat dus uit niet meer dan de som van alle DPIA's waarvoor deze gegevens benodigd zijn. In die DPIA's zijn de doelbinding en aanleiding voor de gebruikte data vastgelegd.
Zie Bijlage 2 voor een schematisch overzicht hiervan.

Het stelsel van GLO's en DPIA's wordt minimaal jaarlijks getoetst op volledigheid, actualiteit en juistheid (zie risico III).

II. Risico: Het verwerken van gegevens in een datawarehouse heeft geen wettelijke grondslag

Er is een gedeeltelijke grondslag te vinden in de verwerking van gegevens voor het 'algemeen belang'. Dat dekt niet de gehele lading, en het (rest)risico wat daardoor overblijft wordt middels een collegevoorstel bij het College van B&W belegd.

III. Risico: GLO's en DPIA's verouderen, en beschrijven niet langer de actuele situatie.

Om dit te voorkomen worden de bestaande GLO's en DPIA's regelmatig tegen het licht gehouden. De GLO's (van de Load1) worden minimaal jaarlijks getoetst op actualiteit en volledigheid door de data-architect. Deze draagt daarvoor de verantwoordelijkheid.
De DPIA's (van de Load2) vallen onder de verantwoordelijkheid van de *product-owner* van het ontwikkelteam. Deze is belast met het opstellen van een DPIA voor elke nieuwe dataset, en dus ook verantwoordelijk voor een jaarlijkse controle op actualiteit. Dit wordt gecombineerd met de controle op de GLO's.

IV. Risico: Gegevens worden onvoldoende geminimaliseerd

Een van de uitgangspunten bij de Algemene Verordening Gegevensbescherming, is dat alleen gegevens worden gebruikt die nodig zijn voor het uiteindelijke doel. Middels een GLO, de onderliggende DPIA's en de jaarlijkse toetsing daarvan, beperken we dit tot een minimum.
Onderdeel van een DPIA is ook dat gekeken wordt naar de toepassing van technieken om herleidbaarheid te voorkomen, zoals het anonimiseren en pseudonimiseren (hashen) van data.

V. Risico: Gegevens – en daarmee analyses – van onvoldoende kwaliteit leiden tot verkeerde conclusies

Bij het overhalen van gegevens naar het datawarehouse wordt nauw samengewerkt met de bronhouder. Waar mogelijk wordt gebruik gemaakt van definities en datamodellen zoals aangeleverd door de bronhouder, of die worden gezamenlijk met de bronhouder en eindgebruiker opgesteld.

Bij het bouwen van de 'datastroom' richting het datawarehouse, wordt uitvoerig getest of deze voldoet aan de kwaliteitseisen. Dit gebeurt zowel op technisch niveau als op inhoudelijk niveau. Dat laatste gebeurt ook weer in samenspraak met bronhouder en eindgebruiker.

Alle gegevens, en resultaten van analyses, worden voorgelegd aan de eindgebruiker ter validatie, vóór zaken geaccepteerd en uitgerold worden.

VI. Risico: Autorisatie voor het datawarehouse

In het datawarehouse staat een uiterst gevoelige gegevensverzameling. De toegang hiertoe beperken we dusdanig, dat slechts een handvol mensen toegang heeft tot de gevoeligste onderdelen van het datawarehouse.
De geprepareerde datasets voor analyses zijn door een groter aantal medewerkers te benaderen, maar zoveel mogelijk ontdaan van (persoons)gevoelige details.

VII. Risico: geen duidelijk proces rondom VOG's

Er is nu geen proces om VOG's en hun screeningsprofiel vast te leggen in personeelsdossiers, en daar ook periodiek op te toetsen.
De CDO onderzoekt dit, eind 2022 hebben we meer helderheid over de verantwoordelijkheid hiervoor.

VIII. Risico: Onrechtmatig gebruik van de gegevens in het datawarehouse

Logging stelt ons in staat terug te halen wanneer welke gegevens gebruikt zijn, waardoor rechtmatigheid en doelbinding getoetst kunnen worden. Het huidige beheersysteem van het datawarehouse *logt* geen gebruiksgegevens. Dat is een functionaliteit die wel gewenst is.
Richting de tweede helft van 2023 hebben we een verkenning gedaan naar datawarehouse-managementsystemen die ons van die functionaliteit kunnen voorzien.

IX. Rechten van betrokkenen zijn onvoldoende gewaarborgd

Betrokkenen hebben een aantal rechten:

1. Recht op inzage
2. Recht op vergetelheid
3. Recht op rectificatie en aanvulling
4. Recht op dataportabiliteit
5. Recht op beperking van de verwerking
6. Recht met betrekking tot geautomatiseerde besluitvorming en profilering
7. Recht op bezwaar

Wat de rechten op vergetelheid, rectificatie, beperking van verwerking en bezwaar betreft, is het datawarehouse volgend aan de bronapplicaties die de data leveren.
Deze verantwoordelijkheid voor de waarborging van deze rechten ligt dus bij de bron.
Het datawarehouse is slechts een kopie. Om te voorkomen dat de kopie afwijkt van de bron, legen we de bestaande tabellen en halen we de data opnieuw over vanuit de bron. Daarmee worden ook verwijderde of aangepaste gegevens op de goede manier overgehaald.
Voor veel datasets in het datawarehouse gebeurt dit dagelijks. Voor de meer complexe datasets, waarvoor veel rekencapaciteit benodigd is, doen we dit minimaal 1 keer per jaar wanneer deze persoonsgegevens bevatten.

Aan de rechten op inzage en dataportabiliteit is met het inrichten van een datawarehouse eenvoudig te voldoen.
De beheersystemen van het datawarehouse stellen ons in staat alle datasets (*Load1* en *Load2*) te doorzoeken en zo alle gegevens over een persoon of zaak te verzamelen. Daarna is het slechts een kleine stap om deze beschikbaar te stellen in een open bestandsformaat.

Het recht met betrekking tot geautomatiseerde besluitvorming en profilering blijft in stand. Gegevens uit het datawarehouse worden niet gebruikt voor automatische besluitvorming, maar dienen als onderligger voor een besluit dat door een professional genomen wordt[3].

[3] Zie ook het Digitaal Informatiebeleid: https://nijmegen.bestuurlijkeinformatie.nl/Document/View/679460f2-1d1f-4f84-aeb2-ee6d71a40e35

Bijlage 1 Ethiek

1. Data-alertheid

1.1 Wat zijn de ethische effecten van dit project?

Deze gegevensverwerking heeft veel effect op de privacy van inwoners. We doen geen analyses op persoonsniveau, maar de gegevens worden wel op dat niveau verwerkt. We moeten goed opletten dat de privacy van inwoners in de uiteindelijke analyses gewaarborgd blijft.

We maken geen gebruik van automatische besluitvorming, dus de veiligheid en autonomie van inwoners verandert niet.

Het datagestuurd werken – met name het openbaar beschikbaar stellen van data – zorgt ervoor dat de machtsverhouding tussen gemeente en inwoner een stukje gelijkwaardiger wordt.
Door het openbaar maken van de dataverwerkingen die we doen, middels het verwerkingsregister, geven we de inwoners van Nijmegen meer inzicht en controle over het gebruik van hun gegevens.

1.2 Welke inspanningen worden verricht om de mogelijk negatieve effecten van dit project te proberen voorkomen?

De data worden zo vroeg mogelijk in het analyseproces ontdaan van herleidbare gegevens. Daardoor zijn individuen niet herkenbaar in de analyses.
Dat neemt niet weg dat er alsnog veel persoonsgegevens verwerkt worden. Dit doen we zo zorgvuldig mogelijk, en proberen zoveel mogelijk te documenteren en transparant te maken (bijvoorbeeld middels DPIA's, GLO's en Open Data).

1.3 In hoeverre is er sprake van (semi-)automatische besluitvorming?

Er is geen sprake van automatische besluitvorming. De eindresultaten van analyses dienen ter ondersteuning van de bestaande besluitvorming, die gedaan wordt door een professional.
Middels verzoeken aan de gemeente kan een inwoner om verheldering van een besluit vragen, waarin de gemeente verplicht is te voorzien.
Omdat besluiten niet automatisch worden genomen, zijn ze altijd te herzien als daar aanleiding toe is.
Het principe van Open Data maakt ook de onderliggende data van besluiten inzichtelijk. Dit zal tot gesprekken leiden op een ander niveau, tussen inwoner en beleidsmaker.

2. Vooringenomenheid (bias)

2.1 Zijn alle verschillende groepen betrokkenen vertegenwoordigd in de dataset(s)?

Bij de meeste gegevensverzamelingen zijn niet alle inwoners van de gemeente betrokken. Analyses op die gegevens kennen dus per definitie een bias.
De medewerkers die deze analyses uitvoeren zijn zich hier goed van bewust. We blijven werken aan die bewustwording, zeker als meer en meer medewerkers met deze gegevens gaan werken.

2.2 Worden uitkomsten van de analyse weer gebruikt voor een vervolganalyse?

Er wordt bij analyses uitgegaan van de data uit het datawarehouse of rechtstreeks uit de bron. Machine Learning-modellen werken vaak iteratief en gebruiken dus de resultaten van een eerder model, maar ook hierbij dient de brondata altijd als basis.

2.3 Bestaat het gevaar dat bepaalde betrokkenen of groepen gediscrimineerd zouden kunnen worden door uw project?

Ja, gezien de meeste gegevensverzameling niet álle inwoners of álle kenmerken omvat, is dit een gevaar. Besluitvorming vindt echter niet plaats op alleen analyseresultaten. Een professional maakt deze besluiten op basis van data, maar ook zijn professionaliteit en intuïtie. Daarmee wordt het mogelijke gevaar van discriminatie aanzienlijk verkleind.

3. Transparantie

3.1 Bestaat het risico op (publieke) verontwaardiging?

Ja, de mogelijkheid bestaat dat betrokkenen het niet eens zijn met de verwerking van hun gegevens ten behoeve van analyse. Hierdoor is het belangrijk open te communiceren over het gebruik van hun gegevens, en vooral de meerwaarde daarvan voor henzelf en hun andere betrokkenen. Eind 2022 verschijnt de uitwerking van onze datastrategie, waarin deze communicatie beschreven wordt.

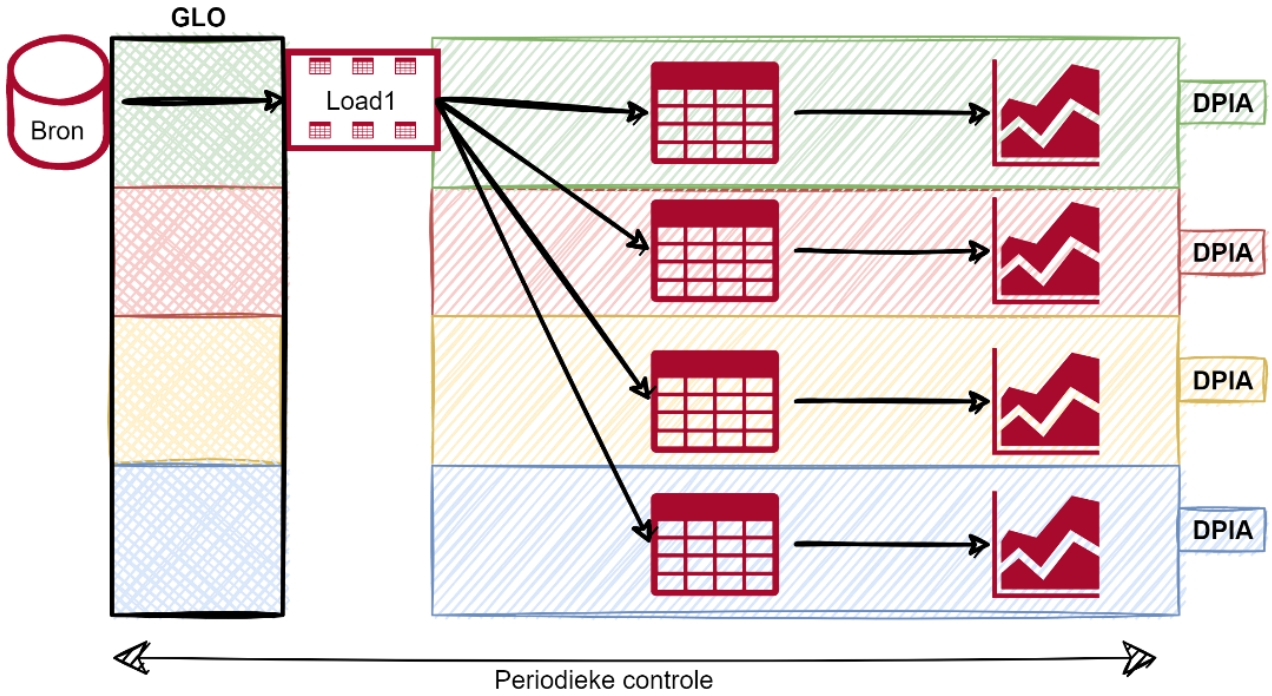
3.2 Kunnen betrokkenen bezwaar maken tegen uitkomsten van het project?

De uitkomsten van een project zijn gebaseerd op de conclusies getrokken door een professional, er is geen sprake van automatische besluitvorming.
Betrokkenen kunnen hun bezwaar kenbaar maken bij de betreffende professional, die hun bezwaren mee kan nemen in de uiteindelijke besluitvorming.

3.3 Is een opt-out mogelijk voor betrokkenen?

Betrokkenen kunnen contact opnemen met de gemeente als zij niet willen dat hun gegevens verwerkt worden. Dit wordt verwerkt in het bronsysteem; het datawarehouse is daarin volgend.

Bijlage 2



Bijlage 3

Met de link hieronder is de DPIA te openen, die in 2017 door BKBO uitgevoerd is.

DPIA Gemeente Nijmegen.datawareł

Bijlage 4

In de DPIA door BKBO (in de Bijlage 3) zijn een aantal aanbevelingen gedaan. Deze aanbevelingen worden hieronder nagelopen. Zo wordt er per aanbeveling aangegeven wat voor acties zijn ondernomen of nog genomen moeten worden.

Nummer	Aspect	Gedaan	Nog te doen
1	Rechten van inwoners		Eind 2022 hebben we de processen die de rechten van betrokkenen waarborgen, beschreven. Beoogd proces is al te vinden bij de risicomaatregelen hierboven (8.3.VIII).
2	Proceskennis	Met de benoeming van een Chief Data Officer en een data-architect, is de verantwoordelijkheid voor dit overzicht officieel belegd.	
3	CISO	De IRvN heeft de rol van CISO officieel belegd	
4	FG	De functie van FG is belegd bij Nijmegen	
5	VOG		Er is nu geen proces bij de organisatie om VOG's vast te leggen in personeelsdossiers, en daar ook periodiek op te toetsen. De CDO onderzoekt dit, eind 2022 hebben we hierover meer helderheid.
6	Beveiligings- en privacybewustzijn	Medewerkers zijn op de hoogte van de AVG, er is een procedure omtrent datalekken	
9	Werkplekken		Deze aanbeveling volgen we niet op.
10	Infrastructuur		In Q3 2022 start de IRVN met het programma Informatiebeveiliging, waarvan de ontwikkeling van een pentestprogramma een onderdeel is.
11	Mobiele devices	Dit is vastgelegd in het mobiele devices-beleid	

12	Bewerkersovereenkomst	Er is een verwerkingsovereenkomst afgesloten met de IRvN	
13	Audittrail		Loggen kan niet in het huidige beheersysteem, dus persoonlijke toegang tot gegevens kunnen we niet toetsen. Tweede helft 2023 hebben we zicht op een systeem dat dat wel kan bieden. Met het systeem van GLO's en DPIA's hebben wel een audittrail op welke gegevens verwerkt worden.
14	Gebruikte gegevensset	Er wordt middels deze DPIA een werkwijze opgezet, om met GLO's en DPIA's inzicht te krijgen in verwerkte gegevens	
15	Gegevensclassificatie		Classificatie van gegevenssets wordt opgepakt in het uitwerken van beleid rondom metadata
16	Melden van persoonsgegevensverwerkingen	De verantwoordelijkheid voor deze gegevensverwerkingen wordt middels een collegevoorstel bij het College belegd	
17	Versleutelde opslag en transport van gegevens		Versleuteling van gegevens wordt opgepakt in het uitwerken van beleid rondom cryptografie
18	Bewaartermijn van gegevens		Bewaartermijnen moeten worden vastgesteld
19	Vernietiging van gegevens		Er moet worden toegezien op handhaven van bewaartermijnen
20	Beheersing AO/IC		Beheersprocessen omtrent onze data wordt vastgesteld in beleidsdocumenten van het Datahuis
21	Eenduidig beleggen ambtelijke verantwoordelijkheid	Er is een Netwerk-MT Data, waarvan Tom Merkx voorzitter is. Hij is vanuit het GMT gemandateerd op het onderwerp Data.	

22	Doelbinding	Voor de datasets in Load2 die persoonsgegevens bevatten, zijn DPIA's gemaakt	Voor de datasets in Load1 worden GLO's opgesteld
23	Beveiligingsbeleid	Er is een informatiebeveiligingsbeleid opgesteld	
25	Inzage en correctierecht		Het datawarehouse is volgend aan de bronsystemen. Eind 2022 is beschreven hoe de processen zodanig ingericht zijn dat het datawarehouse niet te ver afwijkt van de bron.