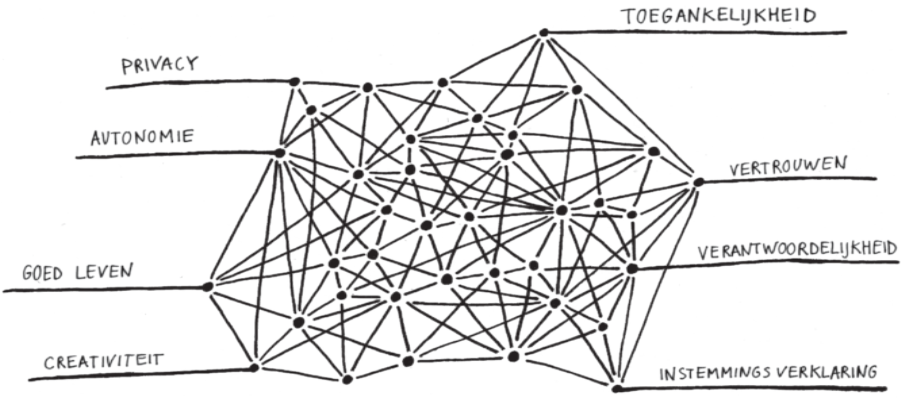
DRINGENDE DATAVERHALEN

Bewustwording van knelpunten in dataprojecten



- ter ondersteuning van de DEDA-handleiding -

INHOUDSOPGAVE

INLEIDING

Big Data en nieuwe analysepraktijken beloven grote voordelen voor (commerciële) bedrijven en publieke instellingen. De mogelijkheden en de voordelen van dataprojecten brengen echter ook moeilijkheden met zich mee. Deze moeilijkheden zijn gemakkelijk te negeren, maar kunnen er op de lange termijn voor zorgen dat goede bedoelingen leiden tot slechte resultaten.

Dataprojecten kunnen als gevolg hebben dat verschillende waarden van mensen, zoals privacy en autonomie, in het geding komen. In reactie hierop zijn er vanuit de overheid een aantal praktijken gereguleerd en zijn er wetten aangepast. De verzwaarde boetes voor het schenden van privacy zijn een voorbeeld van de pogingen van de EU om een verantwoord gebruik van persoonlijke informatie af te dwingen. Naast privacy zijn er ook andere problemen die voort kunnen komen uit dataprojecten. Zo kunnen datasets van schimmige herkomst zijn of uit hun context worden gelicht. Er kan sprake zijn van een vooringenomenheid in de datasets, modellen en algoritmen. Ook kunnen er vragen zijn rond belangenverstrengelingen van commerciële bedrijven en publieke instellingen. Men kan ook denken aan vragen omtrent de sociale impact van datagedreven beleid en de kritische evaluatie hiervan. Dit zijn slechts een klein aantal voorbeelden van de gebieden waarin de wet niet altijd opgaat of een duidelijke richtlijn biedt. Zulke grijze gebieden kunnen verhelderd worden door gebruik te maken van richtlijnen voor ethische besluitvorming.

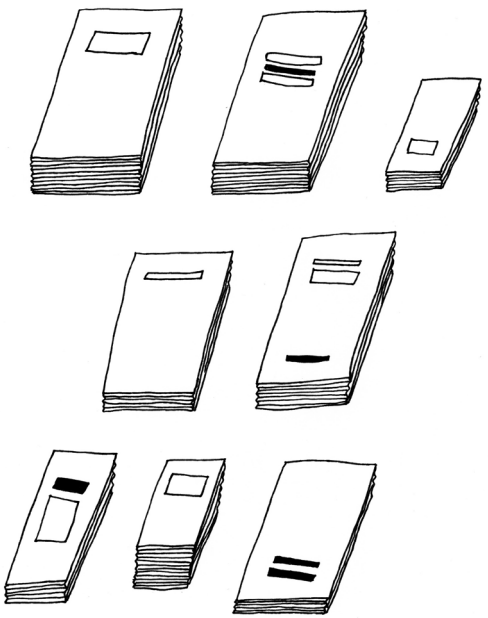
In dit boekje worden een aantal cases uiteengezet waarin ethische problemen in dataprojecten worden uitgelicht. Hiermee willen wij bijdragen aan het ontwikkelen van een gevoeligheid voor waarden die geschonden kunnen worden tijdens dataprojecten. Ook illustreren deze verhalen ethische knelpunten met betrekking tot data, modellen en algoritmen. De cases in dit boekje zijn op vier niveaus ingedeeld: ten eerste kan de datamanagement van dataprojecten niet op orde zijn, ten

tweede kan er te weinig transparantie over het project zijn, ten derde kan er te makkelijk besluitvorming op basis van data worden gemaakt en ten vierde kan er sprake zijn van belangenverstrengeling zijn in dataprojecten.

Bij iedere case staan er een aantal onderwerpen vermeld welke belangrijk zijn in de besproken case. Informatie over deze onderwerpen kan worden gevonden in de DEDA-handleiding. De paginanummers corresponderen met de paginanummers in de DEDA-handleiding.

Voor meer informatie over ethische problemen in dataprojecten, datamanagement en databeleid kunt u terecht op dataschool.nl/deda/. Ook kan u contact met ons opnemen via ***info@dataschool.nl***.

DATAMANAGEMENT



Kort samengevat houdt datamanagement het organiseren, categoriseren, opbergen, ophalen en onderhouden van data in. Binnen dataprojecten is een goed datamanagement essentieel om goed en verantwoord te werken met data. Wanneer het datamanagement niet in orde is, kan (gevoelige) data bijvoorbeeld te makkelijk bereikbaar zijn voor derde partijen (die misschien minder goede intenties hebben). Goede bedoelingen kunnen kort gezegd tot slechte resultaten en wantrouwen leiden wanneer het datamanagement niet in orde is. De volgende vier cases illustreren het belang van goed datamanagement.

Ashley Madison - Gestolen data
Februari 2017

14 anonimiseren
20 verantwoordelijkheid
21 communicatiestrategieën

- De case -

In juli 2015 heeft de hackersgroep The Impact Team alle gebruikersdata van de webservice Ashley Madison gestolen en deze op 18 augustus 2015 online geplaatst. Ashley Madison was een Canadese datingsite met de expliciete intentie om getrouwde mensen te helpen om een affaire met elkaar te beginnen. De slogan van de website luidde: “Life is short. Have an affair.” Bij inschrijving bij de datingssite werd de gebruiker beloofd dat inschrijving 100% discreet en anoniem zou zijn.

Op 15 juli 2015 werd de site gehackt door de hackersgroep The Impact Team. In een manifest gepubliceerd door The Impact Team verklaarden zij dat zij alle gebruikersdata hadden gestolen en deze online zouden plaatsen mits Avid Life Media (ALM, het moederbedrijf van Ashley Madison) Ashley Madison volledig offline zou halen. In het manifest verklaarde The Impact Team dat deze actie een antwoord was op leugens van ALM. Het bedrijf beloofde haar klanten namelijk dat zij voor $20 hun profiel bij Ashley Madison volledig konden laten verwijderen. Volgens the Impact Team waren echter alle klantgegevens nog steeds aanwezig in het klantenbestand. ALM zou met de full delete functie $1,7 miljoen hebben verdiend in 2014. De hackersgroep verklaarde dat zij als antwoord op deze leugens niet alleen het gehele klantenbestand, maar ook documenten van het bedrijf als mails en werknemersinformatie, online zouden plaatsen. Op 18 augustus werd, zoals beloofd, alle data online geplaatst. Het klantenbestand bevatte onder andere namen, adressen, telefoonnummers, mailadressen, zoekhistorie, creditcardgegevens,

fysieke beschrijvingen en seksuele voorkeuren van iedereen die zich ooit had ingeschreven bij de website.

Wereldwijd werd het datalek door velen als rechtvaardig gezien, aangezien de service van Ashley Madison was gebaseerd op leugens en ontrouw. De gebruikers, aan de andere kant, waren bang om geïdentificeerd en geassocieerd te worden met de dating service. Na het datalek volgde een wereldwijde moral shaming, zo was er een Australische radio DJ die in de uitzending een vrouw opbelde om te vertellen de naam van haar man op de lijst voorkwam. Ook was er een krant in Georgië die een lijst met alle namen van de mensen die in het klantenbestand waren opgenomen uit Georgië publiceerde in de krant. Een deel van de gebruikers kwam bovendien uit landen waar vreemdgaan en homoseksualiteit strafbaar is. Een gebruiker op Reddit[1] postte kort na het datalek dat hij een homoseksuele man uit Sauodi-Arabië was, die een account had aangemaakt om andere mannen te ontmoeten. Hij schreef “Ik kom uit een land waar homoseksualiteit bestraft wordt met de dood. Ik heb in Amerika gestudeerd en toen het account bij Ashley Madison aangemaakt. Ik ben single, maar ik gebruikte de site omdat ik homo ben; seks wordt bestraft met de dood in mijn land dus ik wilde de hookups discreet houden. Ik vraag jullie allen om dit bericht te delen.” Nadat de data online werd geplaatst moest de man vluchten uit zijn eigen land. Inmiddels heeft de man op Reddit laten weten dat hij naar Amerika is gevlucht. Dit zijn slechts enkele voorbeelden van wat het datalek wereldwijd teweeg heeft gebracht.

- Ethische knelpunten -

Ashley Madison beloofde haar klanten bij inschrijving discretie en anonimiteit. Na het datalek werd echter duidelijk dat Ashley Madison de persoonlijke data niet voldoende had beveiligd. Er waren volgens The Impact Team veel kwetsbaarheden in Ashley Madison’s source code[2] waardoor de hackers relatief makkelijk bij het klantenbestand konden. Bovendien was de data niet geanonimiseerd in het klantenbestand, waardoor klanten na het datalek makkelijk identificeerbaar waren.

[1] Reddit is een Amerikaanse sociale nieuwswebsite en discussieforum

[2] Source code is de broncode van een programma in een bepaalde programmeertaal

Daarnaast waren de gegevens van diegenen die hadden betaald om het profiel te laten verwijderen nog steeds aanwezig in het klantenbestand. Doordat alle informatie zoals namen, telefoonnummers, adressen aan elkaar was gekoppeld kon iedere gebruiker geïdentificeerd worden. Daarnaast stopte ALM na het datalek met het reageren op haar klanten. Er werd niets gedaan om het datalek uit te leggen, om vragen te beantwoorden of om de klanten te helpen.

Ashley Madison heeft hiermee het vertrouwen van hun klanten geschonden, door zowel de belofte van discretie als de belofte na betaling het profiel te verwijderen te breken. Bovendien heeft het bedrijf de privacy van de klanten geschonden en daarmee persoonlijke schade toegediend aan de betrokken personen.

- Aandachtspunten -

Deze case illustreert het belang van een zorgvuldige datamanagement wanneer er met gevoelige data wordt gewerkt. Wanneer het datamanagement niet op orde is kan dit negatieve gevolgen hebben waarbij persoonlijke levens worden beïnvloed. Vragen die gesteld kunnen worden wanneer een organisatie beschikt over gevoelige data:

- Hoe is het datamanagement geregeld?
- Zijn er gevoelige gegevens opgeslagen?
- Is het nodig om de dataset(s) te anonimiseren of om de data te pseudonimiseren?
- Is het nodig of verantwoord alle gegevens te bewaren of kunnen gegevens ook verwijderd worden?
- Wie is er verantwoordelijk als iets mis gaat?
- Zijn er communicatiestrategieën voor het geval dat er iets mis gaat?
- Hoe kunt u de problemen communiceren met het publiek (en/of de media)?

De besproken knelpunten hoeven niet de enige te zijn in deze case. Wat zijn, volgens u, nog andere ethische knelpunten of moeilijkheden in deze case?

Belastingdienst - De Broedkamer
Februari 2017


23 privacy
20 verantwoordelijkheid
12 algoritmen

- De case -

Op 1 februari 2017 zond het Vara-programma Zembla de uitzending Prutsen en pielen zonder pottenkijkers uit. Zembla spreekt in deze documentaire het vermoeden uit dat de Belastingdienst in Nederland de financiële en persoonlijke gegevens van elf miljoen belastingbetalers en twee miljoen bedrijven in de periode van 2013 tot 2016 onvoldoende heeft beveiligd.

In 2013 werd een nieuw onderdeel van de Belastingdienst, de Broedkamer (later Data & Analytics), verantwoordelijk om alle gegevens van belastingbetalers en bedrijven te koppelen en te analyseren om sneller en goedkoper te kunnen werken. De data die gekoppeld werd bestond uit tientallen zaken, zoals financiële gegevens, adressen, telefoonnummers en reisgedrag. Al deze data was volgens Zembla onvoldoende beveiligd, zo hadden de data-analisten van de Broedkamer onder andere te makkelijk toegang tot de data. Zembla spreekt het vermoeden uit dat binnen de Broedkamer geen gebruik gemaakt werd van autorisatieprofielen en dat er niet door middel van logging werd bijgehouden wie er toegang had tot de data en wat er met de data gebeurde. Zo was de data niet alleen onvoldoende beveiligd, maar kon er ook niet achterhaald worden wie er in de periode van 2013 tot 2016 de data heeft ingezien en wat er met de data is gebeurd.

Een ander probleem dat werd aangekaart in de uitzending is dat de algoritmen die de Broedkamer bedacht en gebruikte niet extern

gecontroleerd werden. De Broedkamer zou algoritmen hebben bedacht dat kon voorspellen welke mensen en bedrijven gecontroleerd moeten worden op fraude. De resultaten van deze algoritmen hebben invloed op het leven van belastingbetalers en bedrijven, er wordt namelijk op basis van deze algoritmen besloten of een persoon of bedrijf gecontroleerd moet worden op fraude. Niemand van buiten de Broedkamer wist echter hoe deze uitkomst tot stand is gekomen.

Volgens Zembla hebben de problemen drie jaar lang kunnen voortbestaan omdat de Broedkamer een afgesloten deel binnen de Belastingdienst was. Niemand van buiten de Broedkamer had zicht op wat er gebeurde en de Broedkamer werd niet gecontroleerd.

Notie: de uitzending van Zembla is op basis van verklaringen onder ede van bronnen tot op het directieniveau. Het NRC meldde op 2 februari 2017 dat Staatssecretaris van Financiën Eric Wiebes, gaat onderzoeken of de Belastingdienst tekort is geschoten in de beveiliging van persoonlijke en financiële gegevens van belastingbetalers.

- Ethische knelpunten -

De uitzending van Zembla suggereert dat de Broedkamer een compleet afgeschermd organisatieonderdeel binnen de Belastingdienst was. Niemand van buiten de Broedkamer had, volgens de uitzending, toezicht op wat er gebeurde en de Broedkamer werd volgens Zembla niet extern gecontroleerd. De Belastingdienst heeft alle (financiële) gegevens van alle belastingbetalers in Nederland. Iedere belastingbetaler in Nederland is verplicht deze data met de Belastingdienst te delen en daarmee gaat een zekere mate van vertrouwen gepaard dat het datamanagement binnen de Belastingdienst goed wordt geregeld. Volgens Zembla was dit binnen de Belastingdienst niet het geval. De Belastingdienst gaf de burger geen andere keuze dan hun diensten te vertrouwen, en hebben vervolgens door onzorgvuldige omgang met de data dit vertrouwen geschonden. Ook andere waarden, zoals de privacy van de burger, waren hier in het

geding.

Ook zou de Broedkamer algoritmen hebben bedacht om bijvoorbeeld fraude op te sporen. De resultaten de algoritmen zou invloed kunnen hebben op het leven van belastingbetalers en bedrijven. Deze algoritmen zijn volgens de uitzending echter niet extern gecontroleerd en niemand buiten de Broedkamer zou weten hoe de output tot stand komt.

- Aandachtspunten -

Deze case illustreert het belang van een zorgvuldig datamanagement wanneer u over gevoelige data beschikt. Vragen die gesteld kunnen worden zijn onder andere:

- Welke wetten, voorschriften of richtlijnen zijn van toepassing op uw project?
- Hoe gevoelig is de data op het gebied van privacy?
- Is er een datamanagementplan?
- Wie heeft toegang tot de dataset(s)?
- Hoe wordt de toegang gemonitord?
- Is er iemand in het team die kan uitleggen hoe het gebruikte algoritme werkt?
- Kunt u de werking van het algoritme communiceren met het publiek?
- Wie is verantwoordelijk als er iets mis gaat?
- Wat zijn de mogelijkheden om risico's te beperken?

 De besproken knelpunten hoeven niet de enige te zijn in deze case. Wat zijn, volgens u, nog andere ethische knelpunten of moeilijkheden in deze case?

Gemeente Tilburg - Wifi-tracking
Mei 2019

23 privacy
20 verantwoordelijkheid
21 communicatiestrategieën

- De case -

Op 7 mei 2019 maakte de gemeente Tilburg bekend te stoppen met het aanbieden van een openbaar, gratis wifi-netwerk. De gemeente bood het netwerk sinds 2010 aan in de binnenstad en Spoorzone van Tilburg. Het besluit volgde op de ontdekking dat er via het wifi-netwerk persoonsgegevens werden verzameld. De gemeente Tilburg maakte namelijk gebruik van wifi-tracking, waarbij via het aangeboden netwerk MAC-adressen van mobiele telefoons werden verzameld zonder dat daar in alle gevallen toestemming voor was gegeven. MAC-adressen zijn unieke identificatienummers die worden toegekend aan apparaten in een netwerk. Met wifi-tracking is het mogelijk om het MAC-adres van een telefoon te achterhalen zonder dat die telefoon verbonden is met het wifi-netwerk. Inwoners en passanten van de Tilburgse binnenstad en Spoorzone konden hun MAC-adres alleen afschermen door de wifi-functie van hun smartphone volledig uit te zetten. Wel was het mogelijk gebruik te maken van een opt-out optie. Dit kon echter alleen wanneer het voor inwoners of passanten bekend was wie de leverancier van het trackingnetwerk was. Wanneer van de opt-out optie gebruik werd gemaakt, werd het MAC-adres nog wel opgehaald, maar vervolgens uit de database gefilterd.

De MAC-adressen en locatiegegevens van verschillende burgers werden voor 17.000 euro per jaar verkocht aan de ondernemers van de stad. Bovendien bleek dat de gegevens werden opgeslagen door de leverancier van het netwerk dat de wifi-tracking mogelijk maakte. Daardoor vond de

gemeente Tilburg dat ze niet genoeg controle had over de toegang en verwerking van de persoonsgegevens. Naar aanleiding van de discussie in Tilburg zijn andere gemeentes (tijdelijk) gestopt met wifi-tracking, of overgestapt op geanonimiseerd passententellen.

- Ethische knelpunten -

MAC-adressen zijn persoonsgegevens. De gemeente haalde deze zonder toestemming van bewoners en passanten op. Bovendien had de gemeente de verantwoordelijkheid om gebruikers op de hoogte te stellen van de persoonsgegevens die het netwerk verzamelde en van de manier waarop deze verwerkt werden. Het is onbekend of bewoners en passanten op de hoogte werden gesteld van de opt-out optie, of dat er in de binnenstad en Spoorzone werd aangegeven dat er sprake was van wifi-tracking. Daarnaast kan men zich afvragen of het voor willekeurige passanten niet onevenredig veel moeite kost om gebruik te maken van een opt-out optie die alleen via de leverancier van het netwerk werd aangeboden.

Door persoonsgegevens te verkopen en beschikbaar te stellen aan de leverancier van het netwerk, hield de gemeente de verantwoordelijkheid voor de verzamelde data niet in eigen hand. Zo is het bijvoorbeeld mogelijk om de locatie en MAC-adressen van burgers te combineren met andere data, waardoor er een volledig profiel kan ontstaan. De gegevens die verzameld werden via het netwerk zouden bijvoorbeeld met bestaande klantprofielen (en andere databronnen) gecombineerd kunnen worden, waardoor de passant niet langer anoniem is.

- Aandachtspunten -

Deze casus illustreert het belang van zorgvuldig nadenken over de manier waarop data wordt verzameld, en welke partijen er toegang toe hebben, of toegang zouden kunnen krijgen (tegen betaling).

Vragen die gesteld kunnen worden in het kader van datamanagement zijn:

- Welke partijen hebben toegang tot de data?
- Kunnen burgers het verzamelen van hun data weigeren?
- Is het voor burgers duidelijk dat hun data wordt verzameld?
- Kan de privacy van burgers door het project in geding komen?
- Wie is er verantwoordelijk voor het beheren van de data?
- Is het verstandig de data te verkopen en welke risico's komen hierbij kijken?

 De besproken knelpunten hoeven niet de enige te zijn in deze case. Wat zijn, volgens u, nog andere ethische knelpunten of moeilijkheden in deze case?

Amsterdam Universitair Medisch Centrum - AmsterdamUMCdb
November 2019

13 bron
16 toegang
23 privacy
22 consent

- De case -

Het Amsterdams Universitair Medisch Centrum (UMC) maakte in november 2019 bekend data van intensive care (IC)-patiënten beschikbaar te stellen voor onderzoeksdoeleinden in één centrale database (AmsterdamUMCdb). Hiermee is het Amsterdam UMC het eerste Europese ziekenhuis dat data over een groep patiënten beschikbaar stelt voor onderzoek. Het initiatief voor de database komt van de Europese Vereniging voor Intensive Care Geneeskunde (ESICM) en wordt gesteund door de Nederlandse Vereniging voor Intensive Care (NVIC). In de database zijn ongeveer een miljard datapunten te vinden, verzameld vanaf 2003. Het delen van de data heeft als doel de zorg voor toekomstige patiënten te kunnen verbeteren. Dit kan bijvoorbeeld door de data te gebruiken om algoritmen en modellen te ontwikkelen binnen het domein van kunstmatige intelligentie. Die algoritmen en modellen kunnen de IC zorg verrijken. Data van IC-patiënten was eerder enkel beschikbaar van Amerikaanse zorginstellingen. Zowel de organisatie van de IC-geneeskunde als het type patiënten in Amerika verschilt zoveel van de Europese situatie, dat de algoritmes en modellen die op basis van die data ontwikkeld worden, waarschijnlijk niet toepasbaar zijn in de Europese context. Verschillende Nederlandse ziekenhuizen overwegen zich aan te sluiten bij het initiatief.

- Ethische knelpunten -

Zowel in verschillende kranten als op de website van de AmsterdamUMCdb wordt benadrukt dat er bij het opzetten van het project goed over privacy is nagedacht. Het project voldoet aan de Europese Algemene Verordening Gegevensbescherming (AVG) en patiënten zijn, volgens een toetsing door onafhankelijke privacy-experts, redelijkerwijs niet meer identificeerbaar. Dat betekent niet dat je individuele patiënten niet kan identificeren als je over veel zorgdata beschikt, zoals een zorgverzekeraar. Toch lijkt ook hier goed over nagedacht te zijn. De dataset is alleen toegankelijk voor onderzoekers die bepaalde cursussen hebben gevolgd en referenties binnen de gezondheidszorg hebben. Daarnaast moeten de onderzoekers akkoord gaan met een aantal voorwaarden die de manieren waarop de data gebruikt worden inperken. Zo mag een onderzoeker niet proberen een individu te identificeren, en als dit per ongeluk wel gebeurt, moet het gelijk gemeld worden.

Dat betekent niet dat alle ethische vragen hiermee beantwoord zijn. Een adviseur bij de Privacy Company stelt tegenover de NOS dat hij vindt dat de initiatiefnemers transparanter zouden mogen zijn door de data protection impact assessment (DPIA) te publiceren. Een ander punt van zorg is de mogelijkheid voor IC-patiënten om in te stemmen met het verwerken van hun data in de database. Hoewel in het persbericht vermeld staat dat de data van patiënten die bezwaar maken niet worden opgenomen in de database, is het nog niet duidelijk hoe de mogelijkheid tot bezwaar wordt gefaciliteerd. Zo zou iedere patiënt die de IC verlaat gevraagd kunnen worden of hij of zij akkoord gaat met het delen van de data ten behoeve van wetenschappelijk onderzoek, maar zou het ook kunnen dat patiënten zelf initiatief moeten nemen om bezwaar te maken. Ten slotte is het onbekend wat er met de data van overleden IC-patiënten gebeurt. Hoewel wettelijk gezien gebruik gemaakt mag worden van data van overleden personen, is de status van overleden personen in de medische context omstreden. Vooralsnog blijft het onduidelijk of bijvoorbeeld de familie in moet stemmen met het gebruik van de data,

zoals dat bij orgaandonatie gebruikelijk is. Ten slotte is het ook onduidelijk of er een beleidsmatig en overkoepelend antwoord op dergelijke vragen bestaat, zodat alle ziekenhuizen die zich willen aansluiten bij dit initiatief vergelijkbare richtlijnen kunnen aanhouden.

- Aandachtspunten -

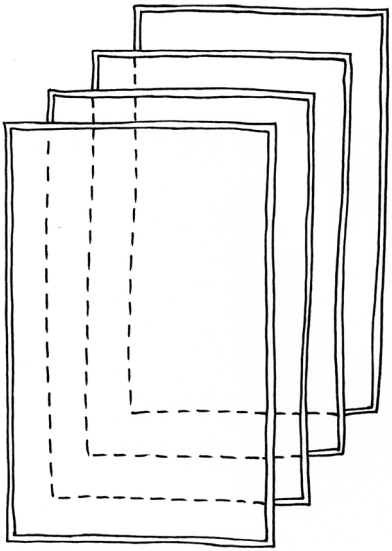
Deze casus illustreert dat goed nadenken over de manier waarop je data verzamelt een belangrijk onderdeel is van datamanagement. Daarnaast maakt deze casus duidelijk dat het voor goed datamanagement belangrijk is om na te denken over de context van het project.

Vragen die gesteld kunnen worden in het kader van datamanagement zijn:

- Voldoet het aan de wettelijke standaarden?
- Zijn er speciale aandachtspunten, gegeven de context van het project?
- Wie beheert de data?
- Hoe veel partijen zijn kunnen (in de toekomst) data aanleveren, en wat zijn hun richtlijnen?
- Hoe is de toegang tot de data gereguleerd?

 De besproken knelpunten hoeven niet de enige te zijn in deze case. Wat zijn, volgens u, nog andere ethische knelpunten of moeilijkheden in deze case?

TRANSPARANTIE



Steeds meer bedrijven en publieke instellingen verzamelen data van hun gebruikers. Het is voor veel mensen onduidelijk wie welke data over hen verzamelt. Dataverzameling en dataprojecten kunnen een impact hebben op de publieke ruimte, sociale interacties, persoonlijke bestaanszekerheid en kunnen zelfs een uitwerking hebben op burgerrechten. Transparantie in dataprojecten houdt in dat men in staat is om de dataset en haar herkomst te verklaren. Accountability (verantwoordelijkheid) betekent dat men verantwoordelijkheid neemt voor de dataverzameling, de analyse en de modellen of algoritmen die gebruikt worden. Het betekent ook dat men transparant is over welke data er wordt verzameld en dat de benodigde informatie wordt verstrekt aan politieke partijen, burgers en experts. De volgende twee cases illustreren het belang van een zekere vorm van transparantie met betrekking tot welke data wordt verzameld en om welke reden.

NS&Exterion - Camera's op NS stations
September 2017

23 privacy
21 communicatiestrategieën
20 verantwoordelijkheid

- De case -

In september 2017 merkte een Twitteraar op dat er camera's in de digitale reclamezuilen van de Nationale Spoorwegen (NS) stations zaten. Op Twitter vroeg hij aan de NS opheldering over het doel van de camera's. Via hun Twitter-account liet NS weten dat de camera's in de reclamezuilen zaten zodat adverteerders konden zien of passanten hun advertentie bekeken en hoe lang er naar de advertentie gekeken werd. Naar aanleiding van die boodschap vroegen andere Twitteraars om meer informatie over de camera's en de gegevens die hiermee verzameld werden. In het antwoord op die vragen deelde de NS mede dat het eigenlijk niet om een camera ging, maar om een sensor. Ook twitterde de NS dat men voor meer informatie over de data bij de exploitant van de reclamezuilen kon informeren, dat er niks illegaals gebeurde en er conform de privacyregeling gehandeld werd.

De reclamezuilen, van exploitant Exterion, bleken uitgerust met een systeem dat statistieken verzamelde over de hoeveelheid mensen die voorbijkwamen, hoeveel mensen er naar de zuilen keken, hoe oud die mensen waren en welk geslacht ze hadden. Hoewel Exterion stelde dat er geen persoonsgegevens werden verzameld, stelde de Autoriteit Persoonsgegevens toch een onderzoek in naar het gebruik van de camera's. Ook was er veel kritiek op de NS, omdat ze als beheerders van de treinstations verantwoordelijk zijn voor wat daar gebeurt. De kritiek op de NS luidde dat ze reizigers er niet van op de hoogte hadden gesteld dat ze gefilmd konden worden voor andere doeleinden dan hun eigen veiligheid.

Naar aanleiding van de kritiek schakelde Exterion op 8 september de camera's uit. Later volgden stickers die de camera's bedekken in de vorm van hun logo, die nu nog steeds op NS stations te herkennen zijn. In 2018 concludeerde de Autoriteit Persoonsgegevens dat burgers niet zonder voorafgaande toestemming gefilmd mogen worden voor commerciële doeleinden, ook niet als de beelden niet worden opgeslagen.

- Ethische knelpunten -

De publieke onrust ontstond vooral om het idee dat reizigers gefilmd konden worden zonder dat zij hierover ingelicht waren, zelfs al werden de beelden niet opgeslagen. In de Trouw werd de situatie op NS-stations vergeleken met douchen: of de beelden worden opgeslagen of niet, je wilt niet dat er iemand meekijkt zonder dat je dat weet. Dit suggereert dat het verzamelen van data (of die nu geanonimiseerd is of niet) controversieel kan zijn, afhankelijk van het doel waarvoor die data wordt verzameld. De kritiek op de NS wijst erop dat er van de NS werd verwacht dat ze als beheerder van dat deel van de publieke ruimte meer verantwoordelijkheid zou nemen voor de privacy van reizigers. Er slechts zorg voor dragen dat er binnen wettelijke grenzen gehandeld werd, was voor veel reizigers niet voldoende.

- Aandachtspunten -

Deze casus toont het belang aan van heldere communicatie over de manier waarop data wordt verzameld en verwerkt. Daarnaast illustreert deze casus hoe belangrijk het kan zijn om over de verantwoordelijkheid van datamanagement na te denken voor organisaties die dit outsourcen aan of faciliteren voor bedrijven.

Vragen die gesteld kunnen worden in het kader van transparantie zijn:

- Is het nodig dat burgers op de hoogte zijn van de manier waarop er data over hen wordt verzameld?
- Wat voor data wordt er verzameld?
- Voor welk doeleinde wordt data verzameld?
- Welke wetten, voorschriften en richtlijnen zijn van toepassing op het project?
- Is het voldoende om aan de geldende wetten, voorschriften en richtlijnen te voldoen, of zijn extra stappen vereist?
- Wie draagt de verantwoordelijkheid voor de communicatie over de verzamelde data?

 De besproken knelpunten hoeven niet de enige te zijn in deze case. Wat zijn, volgens u, nog andere ethische knelpunten of moeilijkheden in deze case?

Facebook - Onderzoek naar emoties van gebruikers
Februari 2017

- De case -

In 2012 heeft Facebook een onderzoek uitgevoerd naar de invloed van de newsfeed van Facebook op de emoties van gebruikers. De aanleiding van het onderzoek was de zorg dat het zien van positieve berichten van vrienden op het sociale medium, zou leiden tot negatieve gevoelens van gebruikers en dat mensen zich buitengesloten zouden voelen. Binnen het onderzoek werd van bijna 700.000 Facebook gebruikers een week lang de newsfeed gemanipuleerd om te kijken welk effect dit had op het gedrag op Facebook. Met het gedrag op Facebook werd binnen dit onderzoek het liken en creëren van berichten bedoeld. Bij één groep gebruikers werden positieve woorden zoals “love” en “nice” uit de newsfeed gefilterd en bij een andere groep gebruikers werden negatieve woorden uit de newsfeed gefilterd zoals “hurt” en “nasty”. Uit het onderzoek bleek dat gebruikers die minder positieve woorden zagen, ook minder positieve berichten creëerden. Dit zou dus impliceren dat wanneer mensen weinig positieve berichten zien op Facebook, zij zich ook minder gelukkig zouden voelen.

Het onderzoek zorgde wereldwijd voor kritiek aangezien de proefpersonen geen toestemming hadden gegeven om mee te doen aan dit onderzoek en hen in zekere zin schade werd aangedaan (namelijk het beïnvloeden van emoties). De proefpersonen hadden geen informed consent gegeven voor deelname aan het onderzoek. Informed consent houdt in dat een persoon toestemming moet geven om mee te doen aan een onderzoek. Informed consent houdt ook in dat de persoon in kwestie de doelstellingen van het onderzoek wordt voorgelegd en de procedure van het onderzoek en de eventuele implicaties en risico’s van het onderzoek weet. Ook wordt er een contactpersoon aangewezen voor het geval de proefpersoon vragen

heeft over het onderzoek of zijn of haar deelname aan het onderzoek wil beëindigen. Informed consent is de ethische en juridische norm voor menselijk onderzoek, om te voorkomen dat de mens schade wordt aangedaan.

- Ethische knelpunten -

De onderzoekers van Facebook verklaren in het onderzoek dat de toestemming die is gegeven met het aanmaken van een Facebook account kan gelden als informed consent. Zij beargumenteren dat het onderzoek in overeenstemming was met de Data Use Policy van Facebook, waar alle gebruikers alvorens een account aan te maken akkoord voor moeten geven, en dat dit akkoord kan gelden als informed consent. James Grimmelmann, professor in de Rechtsgeleerdheid aan de Universiteit van Maryland, kwam als een van de eerste met kritiek op het onderzoek en beargumenteert dat de toestemming die een gebruiker heeft gegeven met het aanmaken van een Facebook account iets heel anders is dan informed consent. De personen in kwestie waren namelijk niet bewust van het feit dat ze hebben meegedaan aan een onderzoek, de procedures van het onderzoek waren niet uitgelegd en de eventuele risico's van het onderzoek waren niet voorgelegd.

Facebook heeft niet open en transparant gehandeld en bracht daarmee privacy en informed consent. Bovendien speelde Facebook met het geluk van hun gebruikers voor hun eigen doeleinden. Het geluk van de gebruikers werd gezien als een middel voor een doel en niet als een doel op zich. Dit schaadde de autonomie van de gebruikers.

- Aandachtspunten -

Wanneer u als bedrijf data verzamelt of een onderzoek uitvoert met de data van mensen dan is het aan te raden om na te denken over hoe u mensen informeert over het onderzoek en/of de betrokken data. Informed consent houdt in dat een persoon toestemming geeft om informatie te

verstrekken aan een onderzoeksproject of er mee instemt om hier aan mee te doen. Het betekent ook dat de persoon geïnformeerd wordt over de doelen van het onderzoek, de procedure en de implicaties die het wellicht heeft voor de persoon in kwestie.

Vragen die gesteld worden wanneer er onderzoek wordt gedaan zijn onder andere de volgende:

- Hoe transparant bent u naar gebruikers over uw project?
- Is het nodig om informed consent te krijgen van de betrokken personen?

 De besproken knelpunten hoeven niet de enige te zijn in deze case. Wat zijn, volgens u, nog andere ethische knelpunten of moeilijkheden in deze case?

DATAGESTUURDE BESLUITVORMING

Data sturen steeds vaker besluitvorming. Deze datagestuurde besluitvorming kent echter een aantal valkuilen. Zo kunnen er in de dataset vooroordelen voorkomen waardoor de besluitvorming op een vertekend beeld gebaseerd kan zijn. Ook kan de werking van de algoritmen die de besluitvorming informeren foutief of onbekend zijn. Datagestuurde besluitvorming heeft meestal invloed op het leven van mensen. Om deze reden is het van belang om een goede kennis te hebben van de dataset en de algoritmen die gebruikt worden alvorens deze in te zetten voor besluitvorming.

Nederlandse Staat - SyRI
Januari 2018

[12] algoritme
[24] bias (vooringenomenheid)
[22] transparantie
[23] privacy

- De case -

In januari 2018 daagde een coalitie van maatschappelijke organisaties, aangevoerd door het Nederlands Juristencomité voor de Mensenrechten en het Platform Bescherming Burgerrechten, de Nederlandse Staat voor de rechter vanwege het gebruik van het Systeem Risico Indicatie (SyRI). Met SyRI kunnen gemeentes onderzoek doen in wijken waarin zij, onder andere, uitkeringsfraude vermoeden. Gemeentes moeten hiervoor toestemming vragen aan het ministerie van Sociale Zaken, omdat SyRI gegevens uit de databases van verschillende overheidsinstanties aan elkaar koppelt, zoals gegevens van de belastingdienst en het UWV. De overheid heeft tot nu toe nog niet bekend gemaakt welke gegevens het algoritme precies verwerkt om te kunnen concluderen dat er sprake is van een verhoogde kans op fraude. Nader onderzoek volgt naar burgers waarvan SyRI aangeeft dat die verhoogde kans er is. In oktober 2019 kwam de zaak voor de rechter. Vanwege de lopende rechtszaak geeft het ministerie van Sociale Zaken aan niet openbaar te maken welk risicomodel SyRI gebruikt.

De gemeente Rotterdam is gestopt met gebruik van SyRI in fraudeonderzoek. In de wijken Hillesluis en Bloemhof leverde het onderzoek een lijst van 1263 verdachte adressen op, wat neerkomt op 1 op de 10 huishoudens. Bewoners geven aan dat ze zich als criminelen behandeld voelen, omdat SyRI hele wijken doorlicht op zoek naar verdachte patronen. Dit is ook één van de redenen van het burgerinitiatief om een rechtszaak aan te spannen. Volgens hen vervaagt de grens tussen

wie wel en niet verdacht is, omdat SyRI het mogelijk maakt burgers door te lichten zonder dat zij dit weten en zonder dat een concrete aanwijzing voor fraude nodig is.

- Ethische knelpunten -

Burgers verstrekken persoonsgegevens aan de overheid wanneer dit wettelijk verplicht is, of wanneer het een voorwaarde is om in aanmerking te komen voor voorzieningen. Door het koppelen van die persoonsgegevens voor risicoprofilering zonder dat hierover openheid wordt gegeven, kunnen burgers het vertrouwen in overheidsinstanties verliezen. Burgers rekenen er immers niet op dat de door hun aangeleverde persoonsgegevens ook tegen hen gebruikt kunnen worden. Omdat er bovendien geen transparantie wordt gegeven over de manier waarop SyRI functioneert, is het voor burgers onduidelijk op welke basis zij als risico worden aangemerkt en waarom er mogelijk nader onderzoek naar hen wordt ingesteld. Het burgerinitiatief dat de Staat voor de rechter heeft gedaagd stelt zelfs dat hiermee onvervreemdbare burgerrechten worden geschonden, zoals de onschuldpresumptie en het inzagerecht dat burgers ten aanzien van hun persoonsgegevens hebben.

- Aandachtspunten -

Vragen die gesteld kunnen worden wanneer computersystemen worden ingezet voor besluitvorming zijn onder andere:

- Is er iemand in het team die kan uitleggen hoe het gebruikte algoritme/model werkt?
- Is er iemand in het team die kan uitleggen hoe de output tot stand komt?
- Welke vooronderstelling liggen besloten in het algoritme/model?
- Wie is er verantwoordelijk als er iets mis gaat?
- Zijn er communicatiestrategieën voor het geval er iets mis gaat?

- Welke bezwaarmogelijkheden zijn er voor degene die getroffen zijn door een besluit?
- Is de procedure proportioneel en haalbaar ook voor mensen met beperkte middelen?
- Is er nagedacht over mogelijke schadeclaims en de kosten hiervan?

 De besproken knelpunten hoeven niet de enige te zijn in deze case. Wat zijn, volgens u, nog andere ethische knelpunten of moeilijkheden in deze case?

Verzekeringen - Risk assessment
Februari 2017

34 bias (vooringenomenheid)
12 algoritmen

- De case -

Op www.insurancecompanies.com is een artikel gepubliceerd over hoe verzekeringsmaatschappijen in de Verenigde Staten de premies berekenen voor polishouders. Verzekeringsmaatschappijen gebruiken een methodologie genaamd risk assessment om de premies te berekenen. Algoritmen berekenen op basis van verschillende datapunten hoe groot de kans is dat een polishouder een beroep doet op zijn verzekering. Op basis van deze berekening wordt de premie bepaald voor een persoon.

De datapunten die hiervoor gebruikt worden zijn onder andere leeftijd, gender, woonplaats en inkomen maar ook de familiegeschiedenis van ziektes en bijvoorbeeld of de polishouder rookt. Al deze datapunten zouden binnen de risk assessment iets zeggen over de mogelijkheid dat de polishouder een claim zal indienen bij de verzekeraar. Wanneer iemand een familiegeschiedenis met erfelijke ziektes heeft, betaalt deze persoon per maand meer dan iemand die geen erfelijke ziektes in de familie heeft. De kans dat de persoon met een familiegeschiedenis van ziektes uiteindelijk in het ziekenhuis belandt zou immers groter zijn dan iemand die geen erfelijke ziektes in de familie heeft.

Volgens insurancecompanies.com zouden de premies voor autoverzekeringen ook op deze manier worden vastgesteld. In Amerika moet iedere chauffeur een verzekering hebben die de kosten dekt van een ongeluk of diefstal. Autoverzekeringsmaatschappijen hebben een lijst van criteria om vast te stellen of het waarschijnlijk is dat de chauffeur een ongeluk zal veroorzaken of dat de chauffeur te maken zal krijgen met

diefstal. Een paar datapunten daarvan zijn de volgende:

- inkomen: mensen met een lager inkomen zouden, volgens dit model, eerder een claim indienen bij de verzekeringsmaatschappij dan mensen met een hoger inkomen.
- leeftijd: jongere chauffeurs zouden, volgens dit model, eerder een ongeluk veroorzaken dan oudere chauffeurs;
- adres: mensen die in een stad wonen betalen meer aangezien de bevolkingsdichtheid groter is en daarmee ook de kans van diefstal.
- burgerlijke staat: de vooronderstelling is dat mensen die getrouwd zijn relatief vaker hun wederhelft in de passagiersstoel zullen hebben en dus voorzichtiger zullen rijden.
- geslacht: mannen zouden meer rijden dan vrouwen en mannen zouden onvoorzichtiger rijden dan vrouwen, wat de kans op een ongeluk verhoogt.

- Ethische knelpunten -

Wanneer bedrijven beslissingen maken met behulp van datasets en algoritmen is het van belang om te realiseren dat geen enkele dataset compleet en geen enkel algoritme neutraal is. Er zijn keuzes gemaakt met welke datapunten een premie wordt berekend, bepaalde dingen zijn wel meegenomen in de risk assessment en andere aspecten niet. In het samenstellen van de dataset liggen aannames en vooroordelen besloten, waardoor (groepen) mensen benadeeld kunnen worden. Het is bijvoorbeeld goed mogelijk dat mannen minder voorzichtig rijden dan vrouwen, maar dat wil niet zeggen dat dit voor iedere man geldt. Wanneer de premie wordt bepaald op basis van de risk assessment zou wel iedere man een hogere premie betalen dan vrouwen. In dit geval kwam de waarde van rechtvaardigheid in het geding. Personen werden niet als persoon, maar op discriminerende wijze behandeld. Ook doet een dergelijk systeem af aan de autonomie van de klanten, aangezien zij worden beoordeeld op iets dat zij niet kunnen veranderen.

- Aandachtspunten -

Deze case illustreert het belang van de kennis van data waarop besluitvorming wordt gebaseerd. Vragen die gesteld kunnen worden zijn onder andere:

- Is de dataset een waarheidsgetrouwe representatie?
- Wat mist er of is er niet zichtbaar in uw dataset?
- Bestaat het gevaar dat bepaalde mensen of groepen gediscrimineerd zouden kunnen worden door uw project?
- Is er iemand in het team die kan uitleggen hoe het gebruikte algoritme werkt?
- Welke aannames liggen besloten in het algoritme?

 De besproken knelpunten hoeven niet de enige te zijn in deze case. Wat zijn, volgens u, nog andere ethische knelpunten of moeilijkheden in deze case?

Predictive policing
Februari 2017

24 bias (vooringenomenheid)
12 algoritmen

- De case -

De Nationale Politie zet steeds meer in op het zogenoemde predictive policing: het voorspellen van crimineel gedrag door middel van grootschalige monitoring en analyse. Daarmee zou de politie al kunnen ingrijpen voordat er een misdaad is gepleegd. Op basis van data analyses kunnen er bijvoorbeeld meer politieagenten worden ingezet in gebieden waar de kans op een nieuw incident (straatroof, inbraak of overval) het grootst is.

Een voorbeeld van predictive policing is het Criminaliteit Anticipatie Systeem (CAS). Het CAS is ontwikkeld door Dick Willems voor de Amsterdamse Politie. Binnen dit systeem is de kaart van Amsterdam opgedeeld in vierkantjes, ieder vierkantje is in het echt 125 bij 125 meter. Per vakje wordt een grote hoeveelheid gegevens verzameld: criminaliteitshistorie, afstand tot bekende verdachten, afstand tot de dichtstbijzijnde snelwegoprit, soort en aantal bedrijven bekend bij de politie, en daarnaast ook demografische en socio-economische gegevens, afkomstig van het CBS. Van ieder vakje wordt op verschillende peilmomenten geregistreerd welke gegevens er op dat moment bekend zijn. Op basis hiervan kan met dit systeem een inschatting gemaakt worden over de waarschijnlijkheid van incidenten in de toekomst. Met deze berekeningen wordt een zogenoemde heatmap van Amsterdam gemaakt, vakjes met een hoge score krijgen een rode kleur, vakjes met een gemiddelde score krijgen een oranje kleur en vakjes met een lage score krijgen een gele kleur. De rood gekleurde vierkantjes zijn hierbij de plaatsen waar de kans op een incident het grootst is. De heatmaps

vormen de basis voor het advies voor surveillanceroutes voor de politie.

De plannen voor het predictive policing blijken veel verder te gaan dan het CAS-systeem. Uit een document van het korps landelijke politiediensten (KLPD) dat begin 2015 lekte wordt de visie van de politie duidelijk. In het document wordt gesteld dat een landelijk camera- en sensorennetwerk de toekomst is van de criminaliteitsbestrijding in Nederland. Met deze zogenoemde slimme camera's zou elke burger in de toekomst in de gaten worden gehouden. Wanneer iemand afwijkt van zijn normale gedrag, bijvoorbeeld door een andere route naar huis te nemen, dan wordt dit als voorbode van criminaliteit gezien en daarmee is deze persoon een verdachte.

- Ethische knelpunten -

Het Criminaliteit Anticipatie Systeem (CAS) maakt voorspellingen over de kans waar een misdaad (straatroof, inbraak of overval) plaats zal vinden het grootst is. Het is belangrijk om te realiseren dat geen enkele dataset compleet en geen enkel algoritme neutraal is. Er zijn keuzes gemaakt welke datapunten worden opgenomen in de besluitvorming en welke datapunten niet worden meegenomen. De criminaliteitshistorie die per vakje wordt verzameld bevat bijvoorbeeld bepaalde misdaden die er in het verleden zijn gepleegd, maar andere misdaden worden niet meegenomen in de voorspelling of er een misdaad zal plaatsvinden. Begrip over welke aannames en vooroordelen in de dataset besloten liggen is van belang alvorens besluitvorming hier op te baseren.

Een ander bezwaar met oog op de predictive policing is dat met predictive policing iedere burger wordt gevolgd. Van oudsher is het Nederlandse wetboek gebaseerd op het zogenoemde daadstrafrecht. Dit houdt in dat er een daad moet zijn gepleegd voordat iemand vervolgd kan worden. Wat er met de slimme camera's in de toekomst zou gaan gebeuren is dat iedere burger gevolgd word. Het innocent until proven quilty-principe[4] wordt daarmee omgedraaid en iedere burger zou vervolgd

[4] Innocent until proven quilty - principe: iemand wordt als onschuldig beschouwd totdat het tegendeel bewezen is

gaan worden, iedere burger is dan een verdachte tot het tegendeel wordt bewezen. Dit zou de aard van het Nederlands strafrecht veranderen, de klassieke daadstrafrecht dan langzaam maar zeker veranderen in een intentiestrafrecht: de intentie om iets crimineels te doen is strafbaar in plaats van de handeling zelf. Het CAS systeem brengt daarmee de waarde van rechtvaardigheid in het geding.

Tenslotte wordt de autonomie van burgers in het geding gebracht, omdat er wordt afgegaan op kenmerken waar de burgers zelf geen invloed op kunnen uitoefenen, zoals je leeftijd, welvaart en wie er in je omgeving door de politie als verdacht worden gezien. Afgaan op deze factoren is bovendien erg discriminerend.

- Aandachtspunten -

Vragen die gesteld kunnen worden wanneer besluitvormingen worden gemaakt op basis van datasets en algoritmen zijn onder andere:

- Is de dataset een waarheidsgetrouwe representatie?
- Wat mist er of is er niet zichtbaar in uw dataset?
- Bestaat het gevaar dat bepaalde mensen of groepen gediscrimineerd zouden kunnen worden door uw project?
- Is er iemand in het team die kan uitleggen hoe het gebruikte algoritme werkt?
- Welke aannames liggen besloten in de dataset en in het algoritme?

 De besproken knelpunten hoeven niet de enige te zijn in deze case. Wat zijn, volgens u, nog andere ethische knelpunten of moeilijkheden in deze case?

Bundesamt für Migration und Flüchtlinge - Taalherkenningssoftware
Maart 2017

[24] bias (vooringenomenheid)
[12] algoritmen en modellen

- De case -

In Duitsland wil het federale bureau voor migratie en vluchtelingen (BAMF: das Bundesamt für Migration und Flüchtlinge) taalherkenningssoftware gaan gebruiken om het land van herkomst van asielzoekers vast te kunnen stellen. De procedure voor asielzoekers om asiel aan te vragen duurde in 2016 gemiddeld 8 maanden. Dit proces wil men in de toekomst gaan versnellen met de taalherkenningssoftware. Zestig procent van de asielzoekers heeft bij aankomst in Duitsland geen enkele vorm van identificatie. Dit maakt het onder andere lastig om het land van herkomst van de asielzoeker te bepalen en daarmee ook de keuze of een persoon asiel toegewezen krijgt. De taalherkenningssoftware zou de plaats van herkomst van asielzoekers gemakkelijk en veilig moeten bepalen.

De software zou als volgt gaan werken; in de software worden fragmenten van verschillende talen en dialecten opgenomen. Vervolgens worden er verschillende spraakfragmenten van de asielzoekers opgenomen welke gespiegeld zouden worden aan de fragmenten in de software. Op deze manier wordt er gekeken of de taal voorkomt in de software en kan er worden vastgesteld waar de asielzoekers vandaan komen.

Vanuit verschillende kanten wordt er kritisch gereageerd op de plannen van de BAMF. Ten eerste op linguïstisch gebied. Joachim Scharloth, professor Taalkunde aan de Universiteit van Dresden, beargumenteert dat taal leeft en voortdurend in ontwikkeling is. Iedere conversatie heeft invloed op iemands taalgebruik. Daarnaast kent een land vele dialecten en vele variaties op dialecten. Zo spreken jongeren anders

dan ouderen en bijvoorbeeld academici anders dan personen zonder opleiding. Bovendien kan vervolging en onderdrukking ook invloed hebben op het dialect van een persoon. Het kan bijvoorbeeld zo zijn dat een bepaald dialect kan leiden tot vervolging en discriminatie waardoor mensen een ander dialect aan kunnen nemen. Het is onmogelijk om alle verschillende dialecten en alle variaties op dialecten op te nemen in de taalherkenningssoftware. Daarnaast zijn de spraakfragmenten die opgenomen worden in de software in een formele setting ingesproken. De levendige en veranderende taal van de asielzoekers zou gespiegeld worden aan deze formele manier van spreken, wat dus zou kunnen leiden tot een vertekende of foutieve uitkomst. Op basis van de uitkomsten van de taalherkenningssoftware zou echter wel een keuze worden gemaakt wat de herkomst van de persoon is en of deze persoon asiel krijgt toegewezen of niet.

Momenteel is de BAMF nog bezig met het onderzoeken of de software in gebruik zal worden genomen. Daarnaast geeft de BAMF aan dat wanneer de software in gebruik zal worden genomen, de uitkomst niet alles bepalend zal zijn maar onderdeel zal worden van de procedure, als “extra, aanvullende verificatie van de identiteit”.

- Ethische knelpunten -

Wanneer bij de aanvraag van asiel door asielzoekers de taalherkenningssoftware wordt ingezet om het land van herkomst vast te stellen, zijn er aantal ethische knelpunten. Technisch gezien kan de taalherkenningssoftware bijvoorbeeld wel afwijkingen opsporen in de vergelijking van de spraakfragmenten, maar de software kan niet verklaren waar deze afwijkingen vandaan komen. Daarnaast worden mensen die een dialect spreken die niet in het systeem is opgenomen benadeeld, omdat de plaats van herkomst niet kan worden vastgesteld. Hierdoor kan er getwijfeld worden aan de rechtvaardigheid van de beoordeling.

Een ander probleem met de software is de aanname van het systeem;

het gaat er namelijk vanuit dat mensen liegen over hun plaats van herkomst totdat het tegendeel bewezen is. Hiermee wordt de waarde van vertrouwen naar de vluchteling toe in het geding gebracht.

- Aandachtspunten -

Vragen die gesteld kunnen worden wanneer er besluitvorming wordt gemaakt op basis van datasets en algoritmen zijn onder andere:

- Is de dataset een waarheidsgetrouwe representatie?
- Wat mist er of is er niet zichtbaar in uw dataset?
- Bestaat het gevaar dat bepaalde mensen of groepen gediscrimineerd zouden kunnen worden door uw project?
- Is er iemand in het team die kan uitleggen hoe het gebruikte algoritme werkt?
- Welke aannames liggen besloten in de dataset en in het algoritme?
- Welke vooronderstelling liggen besloten in het algoritme/model?
- Welke mogelijkheden zijn er voor degene die getroffen zijn door een besluit, om bezwaar te maken?
- Is de procedure proportioneel en haalbaar ook voor mensen met beperkte middelen?

 De besproken knelpunten hoeven niet de enige te zijn in deze case. Wat zijn, volgens u, nog andere ethische knelpunten of moeilijkheden in deze case?

Bristol - Risicopreventie met behulp van een algoritme
Maart 2017

20 verantwoordelijkheid
12 algoritmen

- De case -

Eén op de drie gemeenten in het Verenigd Koninkrijk gebruikt algoritmen die hen kunnen informeren bij het maken van besluiten omtrent de welvaart van hun inwoners, volgens krant The Guardian. Hoewel de verschillende gemeenten verschillende projecten hebben, zet de gemeente Bristol een algoritme in dat alle projecten in andere gemeentes lijkt samen te voegen. Het algoritme van Bristol maakt namelijk gebruik van verschillende databronnen om zo indicaties te krijgen over de werksituatie van hun inwoners, maar ook hun alcohol- en drugsgebruik, psychologische problemen, misdaden, hun afwezigheid van school, mogelijke zwangerschappen onder tieners en zelfs de mogelijkheid dat inwoners mishandeld worden of asociaal gedrag vertonen. Hiervoor gebruiken ze onder andere data van de politie, het nationale instituut voor gezondheidszorg (NHS), het Britse ministerie voor werkgelegenheid en data van lokale autoriteiten. Het systeem van Bristol is zo gemaakt dat het per persoon een score van 0 tot 100 aangeeft bij een categorie, bijvoorbeeld de kans dat die persoon ooit vermist raakt of in aanraking komt met huiselijk geweld. Deze uitkomsten worden enkel met inwoners gedeeld, wanneer deze er om vragen.

De gemeente gebruikt de uitkomsten om besluiten te maken over het inzetten van geld en personeel. Wanneer nieuwe data het risicoprofiel van een inwoner verhoogt, stuurt het systeem deze informatie naar een betrokken sociaal werker zodat deze eventueel het plan van aanpak kan wijzigen. Feedback van medewerkers wordt gebruikt om te bepalen welke indicatoren zwaarder wegen bij een bepaald risico, zodat het systeem

nauwkeuriger wordt. De uitkomsten zijn niet leidend maar worden gebruikt ter aanvulling van het besluit. Wel kan door gebruik te maken van het algoritme informatie veel sneller worden verzameld. Informatie van verschillende instanties die voorheen veel tijd kostte om te achterhalen, wordt nu door het algoritme verzameld en inzichtelijk gemaakt. Met één klik op de knop kunnen ambtenaren een risicoprofiel van inwoners downloaden.

- Ethische knelpunten -

Een interessante vraag voor projecten waarbij algoritmes worden gebruikt die bepaalde risico's in kaart brengen, is in hoeverre je hiermee ook de verantwoordelijkheid op je neemt om te voorkomen dat dat risico zich voordoet. Je kunt je bijvoorbeeld afvragen of de gemeente Bristol preventief sociaal medewerkers af moet sturen op de 10% van de kinderen waarvan de kans dat zij mishandeld gaan worden het hoogst is. Daarnaast is de uitkomst van de data vatbaar voor interpretatie. Als een kind bijvoorbeeld de kans loopt om vermist te raken, betekent dit niet per se dat het kind ook vermist raakt. Als een sociaal medewerker geen indicatie heeft dat er met dat kind iets aan de hand is, maar toch via het algoritme een indicatie krijgt dat er mogelijk iets aan de hand is, dan kan dit een dilemma opleveren voor de desbetreffende medewerker. Moet hij of zij nu wel of niet in actie komen? Als je bovendien het toewijzen van geld en personeel baseert op een algoritme, loop je dan niet het risico om aan een doelgroep voorbij te gaan die het algoritme niet oppikt? De toegang tot de informatie en de interpretatie hiervan zijn lastige vragen bij algoritmes die preventief worden ingezet.

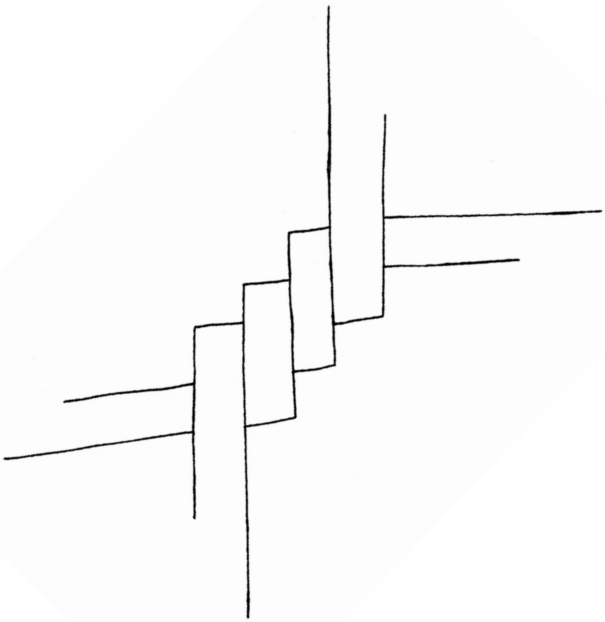
- Aandachtspunten -

Vragen die gesteld kunnen worden wanneer computersystemen worden ingezet voor besluitvorming (in het kader van preventie) zijn onder andere:

- Welke vooronderstellingen liggen besloten in het algoritme/model?
- Welke verantwoordelijkheden hebben we ten aanzien van de (mogelijke) uitkomsten van het algoritme?
- Welke interpretaties zijn er mogelijk ten aanzien van de uitkomsten van het project?
- Staat de actie die ondernomen wordt naar aanleiding van de uitkomst in verhouding tot de aard van het (mogelijke) probleem?
- Gebruiken we het algoritme als aanvullend of leidend?

 De besproken knelpunten hoeven niet de enige te zijn in deze case. Wat zijn, volgens u, nog andere ethische knelpunten of moeilijkheden in deze case?

BELANGENVERSTRENGELING



Steeds meer private en publieke instellingen verkopen data van gebruikers aan derde partijen. Deze derde partijen kunnen onder andere adverteerders zijn die van de data gebruik maken om gerichte advertenties te kunnen plaatsen. Wanneer publieke instellingen bijvoorbeeld financiële data doorvoorkopen kan dit problematisch zijn in verband met de privacy en autonomie van burgers. Vaak is het onduidelijk welke data er wordt doorverkocht aan commerciële partijen en commerciële partijen kunnen misbruik maken van privacygevoelige data. De volgende drie cases illustreren een belangenverstrengeling tussen publieke instellingen en commerciële partijen.

ING – Customer intelligence
Februari 2017

 23 privacy
 20 verantwoordelijkheid

- De case -

Begin maart 2014 kondigde Hans Hagenaars, directeur Particulieren van ING, in Het Financieele Dagblad een proef aan waarbij de bank het betalingsgedrag van klanten zou gaan analyseren. De informatie zou gedeeld worden met adverteerders zodat zij gerichte advertenties aan de klanten van de bank aan konden bieden. Deze zogenoemde customer intelligence zou volgens Hagenaar veel mogelijkheden bieden aan zowel de adverteerders als de klanten. De bank zou de informatie over aan wat en waar klanten geld uitgeven verkopen, adverteerders zouden daardoor gericht reclame kunnen maken en klanten zouden op het juiste moment een relevante aanbieding krijgen. Bijvoorbeeld, wanneer een klant, volgens de bank, teveel betaalt voor energie, dan kon er door de adverteerders een beter aanbod worden gedaan aan de klant.

Vanuit verschillende hoeken werd kritisch gereageerd op het plan. Jacob Kohnstamm van het CBP, College Bescherming Persoonsgegevens (nu A.P.: Autoriteit Persoonsgegevens), besprak in een interview met de NRC dat de bancaire sector vertrouwen verkoopt. Klanten moeten erop kunnen vertrouwen dat er geen privacygevoelige informatie in verkeerde handen valt. Hij stelt voor dat Nederlandse banken goed na moeten denken of ze betaalgegevens van klanten wel willen verkopen aan andere bedrijven. Ook de Consumentenbond had moeite met de plannen van ING. Bart Combée, directeur van de Consumentenbond beargumenteerde: “gegevens over jouw geld zijn zeer privacygevoelig en jouw eigendom.” Hij beschreef op de website van de Consumentenbond dat de relatie met een bank op vertrouwen is gebaseerd. Het delen van klantgegevens door

de bank met andere commerciële partijen staat daar volgens Combée haaks op. Wat Combée betreft kan een partij gegevens die van jou zijn niet verkopen of delen. “Wanneer iemand zelf toestemming geeft met een opt-in[5], dan moet het glashelder zijn waar de klant toestemming voor geeft en hoe een bedrijf voorkomt dat daar misbruik van wordt gemaakt”.

Op 17 maart 2014 maakte ING bekend dat de proef met het commercieel gebruiken van klantgegevens uitgesteld zou worden. In een interview met de Correspondent geeft Bouwe Kuik, IG&H Consultants, aan dat banken in de toekomst customer intelligence wel willen gaan gebruiken. Volgens Kuik zitten banken op het “betaaldatagoud” van hun klanten en zullen zij doorgaan met het verzinnen van manieren om die data te gebruiken. Volgens Kuik is “deze manier van data gebruiken niet per definitie laakbaar, er zullen ook veel mensen zijn die de gepersonaliseerde dienstverlening als een uitkomst zien, of advertenties van derden graag willen ontvangen”.

- Ethische knelpunten -

De kritiek op de plannen van ING waren voornamelijk gericht op de privacygevoeligheid van de data die verkocht zou worden. Wanneer privacygevoelige data wordt doorverkocht aan derde (commerciële) partijen, dan moet het voor het publiek helder zijn wat er met hun data gaat gebeuren en waar zij toestemming voor geven. Door deze helderheid niet te geven, heeft ING de waarden van vertrouwen en transparantie niet voldoende in acht genomen.

[5] Opt-in: de keuze om toe te treden/deel te nemen

- Aandachtspunten -

Vragen die gesteld kunnen worden wanneer publieke instellingen privacy gevoelige data als financiële data doorverkopen aan derde partijen zijn onder andere de volgende:

- Hoe gevoelig is de data op het gebied van discriminatie en privacy?
- Krijgt u inzicht in de persoonlijke levenssfeer van burgers?
- Welke wetten, voorschriften of richtlijnen zijn van toepassing op uw project?
- Hoe informeert u mensen over wat er met hun data gebeurt?
- Hebben mensen de mogelijkheid om medewerking te weigeren?

 De besproken knelpunten hoeven niet de enige te zijn in deze case. Wat zijn, volgens u, nog andere ethische knelpunten of moeilijkheden in deze case?

Deepmind en Royal Free - Streams
Maart 2017

[23] privacy
[20] verantwoordelijkheid
[21] communicatiestrategieën

- De case -

In 2015 ging het Royal Free London NHS Foundation Trust een samenwerking aan met DeepMind. Deepmind is een Brits kunstmatige intelligentie bedrijf en een dochterbedrijf van Google. Binnen deze samenwerking zou DeepMind een app maken om acuut nierfalen van patiënten te monitoren. Dit werd mogelijk gemaakt door patiëntendata die het Royal Free London NHS Foundation Trust beschikbaar stelde. Royal Free is een van de grootste zorgaanbieders in het Verenigd Koninkrijk die gefinancierd wordt door de National Health Service (NHS).

Op 24 februari 2016 maakte DeepMind de samenwerking met de Royal Free bekend. DeepMind zou een smartphone app gaan maken, genaamd Streams, die clinici zou helpen met het monitoren van acuut nierfalen. Acuut nierfalen kan verschillende gevolgen hebben, van kleine nierdisfunctie, tot dialyse, tot transplantie en zelfs tot de dood. Acuut nierfalen zou in Engeland voor 40.000 doden per jaar zorgen. DeepMind claimde dat de app slechts zou functioneren als interface om medische gegevens te controleren. Er werd echter toentertijd geen vermelding gedaan over welke data gebruikt zou worden.

Uit een onderzoek van de New Scientist dat verscheen op 29 april 2016 bleek dat niet alleen de data gerelateerd aan acuut nierfalen, zoals bloedonderzoeken die wezen op diabetes of nierstenen, naar de servers van Google werden doorgestuurd maar ook data die niets met acuut nierfalen te maken had zoals demografische gegevens. Daarnaast werd

niet alleen de data van nierpatiënten naar Google verstuurd, maar van alle patiënten van het Royal Free.

In het Verenigd Koninkrijk moet de patiënt wiens identificeerbare data wordt doorgegeven aan derde partijen expliciet om toestemming worden gevraagd, tenzij de derde partij een relatie van directe zorg heeft met de patiënt in kwestie. Directe zorg wordt gedefinieerd als de preventie, opsporing en behandeling van ziekten en het verlichten van het lijden van de persoon in kwestie. De data van patiënten die te maken hadden met acuut nierfalen zouden onder deze wet doorgestuurd mogen worden naar Google, aangezien Streams zou helpen met de preventie en opsporing van acuut nierfalen. Toch werd ook de data die niets met de aandoening te maken had en ook alle data van patiënten die niet gediagnosticeerd waren met nierfalen doorgespeeld naar de servers van Google omwille van DeepMind. Dit alles zonder expliciete toestemming van de patiënten. Sterker nog, de patiënten werden niet op de hoogte gesteld van het feit dat hun persoonlijke (medische) data nu in handen was van Google.

Toen in november 2015 de data van miljoenen mensen werd gegeven aan DeepMind, werd geen enkele relevante publieke instantie hiervan op de hoogte gesteld. Het Verenigd Koninkrijk heeft een zogenoemde Information Commissioner's Office (ICO) die ervoor verantwoordelijk is dat de Data Protection Act[6] wordt nageleefd. Ook is er de Health Research Authority (HRA) die de belangen van patiënten en het publiek in medisch onderzoek beschermt en behartigt. Dit zijn slechts een paar van de openbare partijen die bij een dergelijke (medische) dataoverdracht op de hoogte hadden moeten worden gebracht. Julia Powles beschrijft echter in een onderzoek naar de samenwerking van DeepMind en de Royal Free dat dit niet is gebeurd. Bovendien is het problematisch in deze dataoverdracht dat het niet bekend is waar Google de (medische) data van alle patiënten voor gebruikt of voor zal gaan gebruiken.

[6] Data Protection Act: een Britse wet inzake de verwerking van data van identificeerbare levende personen, de belangrijkste wetgeving over de bescherming van data

- Ethische Knelpunten -

Bij de samenwerking tussen het Royal Free London NHS Foundation Trust en DeepMind zijn er een aantal ethische knelpunten. Ten eerste heeft de (medische) dataoverdracht plaatsgevonden zonder dat de patiënten hiervan op de hoogte werden gebracht en om toestemming werd gevraagd, bijvoorbeeld in de vorm van informed consent. Ten tweede is de data doorgespeeld zonder dat de relevante instanties op de hoogte werden gebracht. Ten derde is alle medische data nu in handen van Google en weet niemand wat er met deze data wordt gedaan of waar de data in de toekomst voor zal worden gebruikt.

Door deze drie factoren is de autonomie van de klant in het geding gebracht, en zijn de waarden van vertrouwen en transparantie niet gerespecteerd.

- Aandachtspunten -

Wanneer er privacygevoelige data als medische data wordt doorgespeeld van een publieke instelling naar een commercieel bedrijf, zijn er een aantal vragen die gesteld kunnen worden, zoals:

- Welke wetten, voorschriften of richtlijnen zijn van toepassing op uw project?
- Hoe gevoelig is de data op het gebied van privacy?
- Krijgt u inzicht in de persoonlijke levenssfeer van burgers?
- Hoe transparant bent u naar mensen over uw project?
- Hoe informeert u mensen over wat er met hun data gebeurt?
- Hebben mensen de mogelijkheid om medewerking te weigeren?

 De besproken knelpunten hoeven niet de enige te zijn in deze case. Wat zijn, volgens u, nog andere ethische knelpunten of moeilijkheden in deze case?

Nederlandse banken en verzekeraars - Extern Verwijzingsregister

16 toegang
20 verantwoordelijkheid

- De case -

In Nederland hebben zowel banken als verzekeraars een Extern Verwijzingsregister. De registers dienen als lijst waarop zowel banken onderling als verzekeraars onderling gegevens delen over klanten. In principe zijn de registers bedoeld voor de registratie van frauderende klanten, zodat de banken en verzekeraars gewaarschuwd zijn tegen personen die al eerder hebben gefraudeerd. Uit artikelen in het NRC en de Volkskrant blijkt dat de registers op een andere manier worden ingezet. Wanneer je in het Extern Verwijzingsregister van de banken terecht komt kun je geen hypotheek meer afsluiten, geen leningen meer aangaan, geen creditcard meer aanvragen en geen zakelijke rekening openen. Nam je één van deze producten of diensten wel af, dan werden ze geblokkeerd of opgeheven. Is je naam te vinden in het Extern Verwijzingsregister van verzekeraars, dan kan je in praktijk vaak alleen nog een zorgverzekering afsluiten. Zorgverzekeraars hebben namelijk een acceptatieplicht.

Voor beide registers geldt dat er fraudeurs op te vinden zijn en dat je er standaard voor acht jaar in komt te staan. Wat ook voor beide registers geldt, is dat de bank of verzekeraar die je op de lijst plaatst, ook bepaalt of de melding blijft staan wanneer er bezwaar tegen wordt gemaakt. Als wordt besloten dat de melding blijft staan, kan de consument vervolgens enkel nog terecht bij het Klachteninstituut Financiële Dienstverlening (Kifid) of de rechter.

- Ethische overwegingen -

Hoewel het registeren van fraude in een register positief kan zijn wanneer het andere consumenten beschermt tegen de schade die een fraudeur kan veroorzaken, staat de registratie in sommige gevallen niet in verhouding tot de schade die een fraudeur heeft veroorzaakt. Zo wordt er geen rekening gehouden met de omstandigheden waarin iemand een valse claim heeft ingediend. Men kan zich bijvoorbeeld afvragen of acht jaar geen verzekering af kunnen sluiten vanwege een valse claim van 50 euro buitenproportioneel is.

In het artikel in het NRC over het Extern Verwijzingsregister van banken wordt een nog grimmiger beeld geschetst. Zo zouden banken iemand enkel op verdenking van fraude al in het register kunnen zetten. Zelfs na vrijspraak bij de rechter wordt in sommige gevallen nog steeds geweigerd de betreffende persoon uit het register te verwijderen en hier wordt geen extern toezicht op gehouden. Hierdoor kunnen consumenten die geen fraude hebben gepleegd desondanks toch acht jaar zware financiële beperkingen opgelegd krijgen.

- Aandachtspunten -

Vragen die gesteld kunnen worden wanneer data ingezet wordt om instellingen of consumenten te beschermen zijn:

- Wie heeft toegang tot de data?
- Wie heeft de kwaliteit of de rechtvaardigheid van de datasets gecontroleerd?
- Is het nodig dat er toezicht wordt gehouden op de data(sets), eventueel door een onafhankelijke partij?
- Wat welke gevolgen zitten er verbonden aan de uitkomsten van het project?
- Doet het project rekening met de verschillende contexten van de consumenten/burgers die beïnvloed worden door dit project?

De besproken knelpunten hoeven niet de enige te zijn in deze case. Wat zijn, volgens u, nog andere ethische knelpunten of moeilijkheden in deze case?

CONCLUDERENDE NOTITIE

Ondanks de vele mogelijkheden die big data lijkt te brengen, is het belangrijk om stil te staan bij de moeilijkheden in dataprojecten, datamanagement en databeleid. In de besproken cases lijkt data vooroordelen te hebben bij vraagstukken over criminaliteit, fraude en asielzoekers. Aan de andere kant kan data voor commerciële bedrijven veel winst op leveren onder andere door gerichte reclame te kunnen maken.

Uit de cases die besproken zijn wordt echter duidelijk dat in het hele traject van dataprojecten problemen kunnen ontstaan die waarden van mensen, zoals privacy, autonomie, transparantie, rechtvaardigheid en vertrouwen, in het geding brengen. Facetten waarover na moet worden gedacht in dataprojecten zijn onder andere dat men niet blind op data en algoritmen moet vertrouwen. Al bevat een dataset duizenden datapunten, dan nog beschrijft het niet de volledigheid van een mensenleven. Er zijn altijd keuzes gemaakt over welke data wel en niet wordt opgenomen in een dataset. Om bijvoorbeeld discriminatie te voorkomen, is het belangrijk om bewust te zijn over welke aspecten niet zichtbaar zijn in de dataset. Daarnaast kan de werking van algoritmen niet duidelijk zijn of foutief zijn, maar wordt er wel besluitvorming op basis van deze uitkomsten gemaakt. Om deze reden is het belangrijk om te begrijpen hoe de uitkomst van een algoritme tot stand komt.

De besproken knelpunten zijn slechts een deel van de moeilijkheden die zich voor kunnen doen in dataprojecten. Voor een ethische verantwoorde omgang met data in dataprojecten, datamanagement en databeleid raden wij u aan om contact op te nemen met de Utrecht Data School via info@dataschool.nl.

BRONNEN

Amsterdam Universitair Medisch Centrum

- Amsterdam UMC (z.d.) ***Access Request Form and End User License Agreement for AmsterdamUMCdb*** [Formulier.] Geraadpleegd (22-11-2019): https://amsterdammedicaldatascience.nl/arfeula_v1.2.pdf
- Amsterdam UMC, Europese Vereniging voor Intensive Care Geneeskunde, Nederlandse Vereniging voor Intensive Care (2019, 17 november) ***Amsterdam UMC stelt data over intensive care patiënten beschikbaar om levens te redden.*** [Persbericht.] Geraadpleegd (22-11-2019): https://amsterdammedicaldatascience.nl/press/nl.pdf
- Brink, van den, R., Schellevis, J. (2019, 17 november) Slimme algoritmes moeten levens van intensive care-pati:enten redden. ***NOS.*** Geraadpleegd (22-11-2019): https://nos.nl/artikel/2310884-slimme-algoritmes-moeten-levens-van-intensive-care-patienten-redden.html
- Jacobs, A. (2019, 18 november) Newsroom: Amsterdam UMC stelt data over IC-patiënten beschikbaar voor wetenschappelijk onderzoek. ***Smarthealth.*** Geraadpleegd (22-11-2019): https://www.smarthealth.nl/ehealth-mhealth-newsroom-week-47-2019
- (Red. 2019, 17 november) Amsterdam UMC stelt data van IC-patiënten beschikbaar voor onderzoek. ***NU.nl***. Geraadpleegd (22-11-2019): https://www.nu.nl/binnenland/6011628/amsterdam-umc-stelt-data-van-ic-patienten-beschikbaar-voor-onderzoek.html
- (Red. 2019, 22 november) Amsterdam UMC stelt data IC-patiënten beschikbaar. ***ICT&Health***. Geraadpleegd (22-11-2019): https://www.icthealth.nl/nieuws/amsterdam-umc-stelt-data-intensive-carepatienten-beschikbaar/
- Twillert, van, M. (2019) Amsterdam UMC stelt data ic-patiënten beschikbaar. ***Medisch Contact.*** Geraadpleegd (22-11-2019): https://www.medischcontact.nl/nieuws/laatste-nieuws/nieuwsartikel/amsterdam-umc-stelt-data-ic-patienten-beschikbaar.htm
- Webmaster, J. (2019, 18 november) Amsterdam UMC stelt data over IC-patiënten beschikbaar om levens te redden. ***Medicalfacts.*** Geraadpleegd (22-11-2019): https://www.medicalfacts.nl/2019/11/18/amsterdam-umc-stelt-data-over-ic-patienten-beschikbaar-om-levens-te-redden/

Ashley Madison

- MacLellan, Danny. “The Impact Team Manifesto to AshleyMadison.Com.” July 21, 2015. Lamont, Tom. “***Life After the Ashley Madison Affair***.” February 28, 2016. https://medium.com/@dannymack/the-impact-team-manifesto-to-ashleymadison-com-5d4e7225b787#.hwvov8axx

- Lamont, Tom. *"Life After the Ashley Madison Affair."* February 28, 2016. https://www.theguardian.com/technology/2016/feb/28/what-happened-after-ashley-madison-was-hacked
- Titova, Valeria. "*Karma Watch: Ashley Madison.*" July 22, 2016. https://blog.kaspersky.com/ashley-madison-one-year-after/12652/. https://www.reddit.com/r/lgbt/comments/3ebzzj/i_may_get_stoned_to_death_for_gay_sex_gay_man/

De Belastingdienst

- Sondermeijer, Vincent. "**Wiebes: Onderzoek Mogelijk Datalek Belastingdienst.**" February 02, 2017. https://www.nrc.nl/nieuws/2017/02/02/wiebes-start-onderzoek-naar-mogelijk-datalek-belastingdienst-a1544178.
- ZEMBLA - Onderzoeksjournalistiek. ***ZEMBLA - Prutsen En Pielen Zonder Pottenkijkers***. February 1, 2017. https://youtu.be/tI_lJzz2TYQ

Bristol

- Booth, R. (2019, 15 oktober) How Bristol assesses citizens' risk of harm – using an algoritm. ***The Guardian.*** Geraadpleegd (22-11-2019): https://amp.theguardian.com/uk-news/2019/oct/15/bristol-algorithm-assess-citizens-risk-harm-guide-frontline-staff?__twitter_impression=true
- Marsh, S. (2019, 15 oktober) One in three counsils using algorithms to make welfare decisions. ***The Guardian***. Geraadpleegd (22-11-2019): https://www.theguardian.com/society/2019/oct/15/councils-using-algorithms-make-welfare-decisions-benefits

Gemeente Tilburg

- Gemeente Tilburg (2019, 7 mei) ***Geen openbare WiFi meer in de Binnenstad en Spoorzone***. Geraadpleegd (18-09-2019) via: https://www.tilburg.nl/actueel/nieuws/item/geen-openbare-wifi-meer-in-binnenstad-en-spoorzone/
- Kagie, S. (2019, 4 mei) Privacyregels opnieuw overtreden in Tilburg: gemeente bewaart wifi-gegevens van mobieltjes te lang. ***Omroep Brabant.*** Geraadpleegd (18-09-2019) via: https://www.omroepbrabant.nl/nieuws/2992665/Privacyregels-opnieuw-overtreden-in-Tilburg-gemeente-bewaart-wifi-gegevens-van-mobieltjes-te-lang
- Teeffelen, van, K. (2019, 1 maart) Gemeenten stoppen met het wifi-tracken van hun winkelpubliek. ***Trouw.*** Geraadpleegd (18-09-2019) via: https://www.trouw.nl/nieuws/gemeenten-stoppen-met-het-wifi-tracken-van-hun-winkelpubliek~b522858b/
- Teeffelen, van, K. (2019, 4 mei) Tilburg blundert weer met de privacy van bezoekers van de binnenstad. ***Trouw***. Geraadpleegd (18-09-2019) via: https://www.trouw.nl/nieuws/tilburg-blundert-weer-met-de-privacy-van-bezoekers-van-de-binnenstad~b3d25fb7/

- Teeffelen, van, K. (2019, 7 mei) Tilburg stopt met het openbare wifi-netwerk in de binnenstad na privacyfouten. ***Trouw***. Geraadpleegd (18-09-2019) via: https://www.trouw.nl/nieuws/tilburg-stopt-met-het-openbare-wifi-netwerk-in- de-binnenstad-na-privacyfouten~b138820c/
- Verhagen, L. (2019, 31 mei) Tientallen Nederlandse gemeenten volgen u met wifitracking. ***De Volkskrant***. Geraadpleegd (18-09-2019) via: https://www.volkskrant.nl/nieuws-achtergrond/tientallen-nederlandse-gemeenten-volgen-u-met-wifitracking~b5c614ad/

MiDAS

- Egan, Paul. "***False Fraud Cases Against Unemployment Claimants May Hit 50,000***." January 06, 2017. False fraud cases against unemployment claimants may hit 50,000.
- Gross, Allie. "***Update: UIA Lawsuit Shows How the State Criminalizes the Unemployed.***" October 05, 2015. http://www.metrotimes.com/news-hits/archives/2015/10/05/uia-lawsuit-shows-how-the-state-criminalizes-the-unemployed
- Ringler, Doug A. ***Michigan Integrated Data Automated System (MiDAS)***. n.p.: State of Michigan Auditor General, 2016. http://www.audgen.michigan.gov/finalpdfs/15_16/r641059315.pdf
- What to Do If Wrongly Accused of Unemployment Insurance Fraud. ***Michigan Law Unemployment Insurance Clinic,*** 2015. https://www.law.umich.edu/clinical/unemploymentinsurance/Documents/What%20to%20Do%20If%20Accused%20of%20Fraud%20August%202015.pdf

Nederlandse banken en verzekeraars

- Driessen, C., Kooiman, J. (2019, 18 november) Op de zwarte lijst van banken: in de financiële gevangenis. ***NRC.*** Geraadpleegd (22-11-2019): https://www.nrc.nl/nieuws/2019/10/18/op-de-zwarte-lijst-van-banken-in-de-financiele-gevangenis-a3977315
- Eerenbeemt, van den, M. (2016, 22 juli) Hoe kom je van de zwarte lijst van
- verzekeraars af? ***De Volkskrant.*** Geraadpleegd (22-11-2019): https://www.volkskrant.nl/nieuws-achtergrond/hoe-kom-je-van-de-zwarte-lijst-van-verzekeraars-af~b1c3d5eb/

Nederlandse Staat

- Huisman, C. (2019, 27 juni) SyRI, het fraudesysteem van de overheid, faalt: nog niet één fraudegeval opgespoord. ***De Volkskrant***. Geraadpleegd (18-09-2019) via: https://www.volkskrant.nl/nieuws-achtergrond/syri-het-fraudesysteem-van-de-overheid-faalt-nog-niet-een-fraudegeval-opgespoord~b789bc3a/

- Huisman, C. (2019, 16 oktober) Rotterdam stopt omstreden fraudeonderzoek met SyRI. ***De Volkskrant.*** Geraadpleegd (18-09-2019) via: https://www.volkskrant.nl/nieuws-achtergrond/rotterdam-stopt-omstreden-fraudeonderzoek-met-syri~becb336a/
- Kamerstuk 32761-122 (2018, 12 juni) ***Brief van de Staatssecretaris van Sociale Zaken en Werkgelegenheid.*** Geraadpleegd (18-09-2019) via: https://zoek.officielebekendmakingen.nl/kst-32761-122.html
- (Red. 2019, 9 januari) SyRI medio 2019 voor de rechter. ***Platform bescherming burgerrechten.*** Geraadpleegd (18-09-2019) via: https://platformburgerrechten.nl/2019/01/09/syri-medio-2019-voor-de-rechter/
- (Red. 2018, 12 januari) Tommy Wieringa wil niet dat de overheid zijn risico meet. ***Trouw.*** Geraadpleegd (18-09-2019) via: https://www.trouw.nl/nieuws/tommy-wieringa-wil-niet-dat-de-overheid-zijn-risico-meet~b58f21b5/
- (Red. 2019, 13 juni) Buurtbijeenkomst over risicoprofilering van Rotterdam Bloemhof en Hillesluis. ***Platform bescherming burgerrechten.*** Geraadpleegd (18-09-2019) via: https://platformburgerrechten.nl/2019/06/13/buurtbijeenkomst-over-de-risicoprofilering-van-rotterdam-bloemhof-en-hillesluis/
- (Red, zonder datum) Missie. ***Bij voorbaat verdacht.nl*** Geraadpleegd (18-09-2019) via: https://bijvoorbaatverdacht.nl/missie/

NS&Exterion

- Custers, B. (2017, 10 september) Die camera's mogen gewoon niet. ***Trouw.*** Geraadpleegd (18-09-2019) via: https://www.trouw.nl/opinie/die-camera-s-mogen-gewoon-niet~b7a0d982/
- Klaassen, N. (2018, 26 juni) Privacywaakhond: zonder toestemming mag reclamezuil niet filmen. ***Algemeen Dagblad.*** Geraadpleegd (18-09-2019) via: https://www.ad.nl/binnenland/privacywaakhond-zonder-toestemming-mag-reclamezuil-niet-filmen~a5802fb1/
- Niewold, M. (2017, 11 september) Bedrijf volgt treinreizigers met camera, en dat mag misschien niet. ***RTLnieuws.*** Geraadpleegd (18-09-2019) via: https://www.rtlnieuws.nl/economie/artikel/2540266/bedrijf-volgt-treinreizigers-met-camera-en-dat-mag-misschien-niet
- Nsdefect (2017, 3 september) ***@NS_online @ROVER_online wat doen die mini-cameratjes in die reclamezuil! Dit lijkt me geen beveiliging! Station amersfoort perron 6/7*** [Tweet]. Geraadpleegd (18-09-2019) via: https://twitter.com/Nsdefect/status/904582967409987585
- Siedma, T. (2017, 4 september) Tracking op stations. De NS staat erbij en kijkt er naar. ***Bits of Freedom***. Geraadpleegd (18-09-2019) via: https://www.bitsoffreedom.nl/2017/09/04/tracking-op-stations-de-ns-staat- erbij-en-kijkt-er-naar/
- Sondermeijer, V. (2017, 11 september) Exploitant schakelt camera's in reclamezuilen uit. ***NRC***. Geraadpleegd (18-09-2019) via: https://www.nrc.nl/nieuws/2017/09/11/exploitant-schakelt-cameras-in-reclamezuilen-uit-a1573012

Verzekeringen

- "*Insider Information: How Insurance Companies Measure Risk.*" http://www.insurancecompanies.com/insider-information-how-insurance-companies-measure-risk/.
- Rainie, Lee and Janna Anderson. "*Code-Dependent: Pros and Cons of the Algorithm Age.*" February 08, 2017. http://www.pewinternet.org/2017/02/08/code-dependent-pros-and-cons-of-the-algorithm-age/

Predictive Policing

- de Koning, Bart. "De Politie van de Toekomst Houdt Iedere Burger Non-Stop in de Gaten." *De Correspondent.* 2015. https://decorrespondent.nl/3044/de-politie-van-de-toekomst-houdt-iedere-burger-non-stop-in-de-gaten/279163005704-5df91b90.
- KLPD. "*Visie op Sensing binnen de Politie; waarnemen in een genetwerkte maatschappij*." Juni 2011.
- Willems, Dick and Reinder Doeleman. "*Predictive Policing – Wens of Werkelijkheid?*" 2014.https://www.politieacademie.nl/kennisenonderzoek/kennis/mediatheek/pdf/89539.pdf

BAMF

- Biselli, Anna. "*Software, die an der Realität scheitern muss*." Zeit Online. 17 Mar. 2017. http://www.zeit.de/digital/internet/2017-03/bamf-asylbewerber-sprach-analyse-software-computerlinguistik
- "*Dialektsoftware soll Herkunft von Asylbewerbern erkennen.*" Zeit Online. 17 Mar. 2017. http://www.zeit.de/gesellschaft/zeitgeschehen/2017-03/bamf-software-asylverfahren-dialekt-erkennen

Facebook

- Grimmelmann, James. "*As Flies to Wanton Boys*." June 28, 2014. http://laboratorium.net/archive/2014/06/28/as_flies_to_wanton_boys

Vizio

- Fair, Lesley. "*What Vizio Was Doing Behind the TV Screen.*" February 6, 2017. https://www.ftc.gov/news-events/blogs/business-blog/2017/02/what-vizio-was-doing-behind-tv-screen
- Steele, Billy. "*Vizio Tracked and Sold Your TV Viewing Habits Without Consent (updated).*" 02 Jun, 2017. https://www.engadget.com/2017/02/06/vizio-smart-tv-viewing-history-settlement-ftc/

ING

- Heck, Wilmer. "'*Als ING dit plan doorzet, moet wetgever ingrijper*.'" March 14, 2014. https://www.nrc.nl/nieuws/2014/03/14/als-ing-dit-plan-doorzet-moet-wetgever-ingrijper-1355987-a439025
- Klompenhouwer, Laura. "*Consumentenbond: Plannen ING in Strijd Met Privacywetgeving*." March 10, 2014. https://www.nrc.nl/nieuws/2014/03/10/consumentenbond-plannen-ing-in-strijd-met-privacywetgeving-a1426626
- Martijn, Maurits. "*Hoe ABN Amro Weet Dat Jij Een Buggy Nodig Hebt*." 2014. https://decorrespondent.nl/1314/hoe-abn-amro-weet-dat-jij-een-buggy-nodig-hebt/54349265916-f7f236f5
- Redactie. "*ING Stelt Proef Commercieel Gebruik Klantgegevens Uit.*" March 17, 2014. http://www.volkskrant.nl/economie/ing-stelt-proef-commercieel-gebruik-klantgegevens-uit~a3616281/
- *"UPDATE: Zorgen over ING Plannen Met Klantgegevens."* March 10, 2014. https://www.consumentenbond.nl/nieuws/2014/zorgen-over-ing-plannen-met-klantgegevens

Deepmind

- Powles, Julia, and Hal Hodson. "*Google DeepMind and healthcare in an age of algorithms.*" Health and Technology (2016): 1-17.
- Shead, Sam. "*DeepMind's First Deal with the NHS Has Been Torn Apart in a New Academic Study*." Business Insider Deutschland. 16 Mar. 2017. http://www.businessinsider.de/deepmind-royal-free-london-nhs-deal-inexcusable-mistakes-2017-3

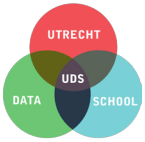
COLOFON

Auteurs: Aline Franzke en Christl de Kloe
Redactie: Danique van der Hoek
onder toezicht van: Mirko Tobias Schaefer
in opdracht van: Gemeente Utrecht

Grafisch ontwerp: Sammy Hemerik

Universiteit Utrecht

Utrecht Data School
Drift 13, kamer 0.01
3512 BR Utrecht


UTRECHT
UDS
DATA
SCHOOL


Gemeente Utrecht


Universiteit Utrecht