

# Opdracht ethiek

Matthijs Soetaert & Jacob Vandererfven

17 Februari 2019

## 1 Deel D

### 1.1 Introductie

In dit deel van de portfolio wordt het artikel 'Waarom hebben we een recht op uitleg wanneer machines over ons beslissen' door Fleur Jongepier & Esther Keymolen. Als eerste wordt een kleine samenvatting geschreven gevolgd ....  
**NOG AANVULLEN MET VERDER**

### 1.2 Samenvatting

Het artikel start met een kort overzicht wat het 'recht op uitleg' is. In de Algemene Verordening Gegevensbescherming (AVG) ofwel de General Data Protection Regulation (GDPR) staat geschreven dat iedere persoon het "recht op uitleg heeft". Dit houdt in dat wanneer een beslissing wordt gemaakt door een machine, de persoon op wie deze beslissing invloed heeft het recht heeft op nuttige informatie over de onderliggende logica bij die beslissing. Nuttige informatie is vaag maar er wordt verondersteld dat dit informatie is die iedereen zou kunnen begrijpen en niet alleen vakexperten. Een voorbeeld hiervan zou kunnen zijn dat wanneer de beslissing wordt gemaakt dat een persoon geen lening krijgt, deze persoon ook mag weten wat de redenen zijn.

Als eerste wordt gekeken waarom juist dat recht op uitleg enkel bij machines geldt en niet bij personen. Door juist het onderscheid tussen deze twee te maken wordt er een normatieve asymmetrie verondersteld ofwel dat de beslissingen in normatief opzicht anders zijn. Om te weten of het wel aan de machine ligt wordt dit onderzocht aan de hand van twee vragen.

Als eerste wordt de vraag gesteld ofdat er wel een normatieve asymmetrie bestaat tussen de twee. Deze vraag kan herleid worden tot de vraag "Bestaat

er een duidelijk onderscheid tussen menselijke en geautomatiseerde beslissingen?” waarbij vermoedelijk verondersteld wordt dat het een de beslissing maakt in afzondering van het ander. Een voorstelling als deze is te simplistisch doordat een puur geautomatiseerde beslissing niet bestaat. Telkens wanneer een beslissing als deze wordt gemaakt is er ergens wel een mens aan te pas gekomen enerzijds in het ontwerp maar anderzijds ook in het maken van de voorkeursregels die de machine moet volgen bij het maken van beslissingen. Deze verwikkeling van mens en machine is juist wat het recht op uitleg zo belangrijk maakt. Wanneer we beroep doen op dit recht doen kunnen we door het menselijke aspect niet gewoon zeggen ”de machine heeft dat besloten” maar zoeken we een morele actor aan de andere kant. Deze morele actor kan geen machine of algoritme zijn en dus zoeken we de mens erachter. Dit is dus in principe wat we doen als we gebruik maken van het recht op uitleg, de mens achter de machine aan het spreken krijgen.

Als tweede vraag wordt er gekeken of het eigenlijk wel principieel iets uitmaakt dat een machine iets over ons beslist. Machines beslissen tegenwoordig al zeer veel over ons: wanneer de verwarming aanmoet, welke reclamespot aan ons getoond wordt, wat het volgende nummer zal zijn dat wordt afgespeeld, uitkeren van een lening.. Indien dit het geval is kan de vraag gesteld worden of het wel zo een groot verschil uitmaakt indien de machine beslist en een automatische email verstuurd met het resultaat ofdat een mens die dezelfde beslissing heeft gemaakt net dezelfde email stuurt. Het is dus niet duidelijk waarom de ene minder problematisch is dan de andere en waarom we dus meer recht op uitleg geeft dan de andere.

De conclusie tot nu toe is dat het er eigenlijk niet zozeer toe doet wie de beslissing maakt maar eerder wat de beslissing is en waarom ze wordt genomen. Dus als de vraag niet te funderen is door het onderscheid mens/machine, hebben we dan wel nog het recht op uitleg? Het voorstel dat door het artikel wordt gedaan is om het specifieke recht op uitleg in de AVG af te leiden uit een meer algemeen moreel recht op uitleg. Voordat dit kan gedaan worden moeten nog enkele afspraken gemaakt worden. Als eerste moet gekeken worden of er wel altijd recht is op uitleg, bijvoorbeeld hebben we het recht op uitleg wanneer een tas koffie meer dan drie euro kost? Om dit onderscheid te maken wordt beroep gedaan op de connectie tussen kennis en handelen en de afbakening te introduceren dat personen een moreel recht hebben op uitleg wanneer ons handelingsvermogen sterk wordt aangetast door de beslissing. Zo heeft het recht op uitleg te maken met het deliberatief handelen of autonoom handelen van een persoon.

Verder wordt er gekeken wat de noodzakelijke condities zijn van het auto-

noom handelen. Zo worden de voor de hand liggende voorbeelden aangehaald zoals fysieke dwang of afwezigheid bij innerlijke dwang. Er zijn ook epistemische condities nodig om vrijwillig te kunnen handelen omdat er een bepaalde mate van kennis of begrip noodzakelijk. Dit kan gaan over zelfkennis maar ook over kennis over de wereld en de beweegredenen van anderen. Dit is nodig omdat de keuzes van anderen soms een grote impact hebben op onze eigen handelingsvrijheid. De sprong is nu snel gemaakt naar het idee dat als iemand jouw autonome handelsvermogen onvoldoende respecteert wanneer er relevante informatie ontnomen is door deze achter te houden. Verder wordt er nog duidelijk gemaakt dat er nog altijd een verschil is tussen liefdadigheid en de morele plicht om informatie te geven.

Door deze vorige redenering van de epistemische condities om autonoom te handelen kunnen we nu kijken op welke manier geautomatiseerde beslissingen deze naleven. De conclusie hieruit is dat de reden dat we denken dat we een recht hebben op uitleg bij geautomatiseerde beslissingsvorming niet zit in het principiële feit dat machines beslissingen moeten nemen maar in het feit dat machines beslissingen nemen die vaak niet voldoen aan de epistemische condities voor autonomie.

Nog laatste stukje eens aanvullen

### 1.3 Normatieve assymetrie

Bij het formuleren van een antwoord op de vraag of het wel door de machine komt dat we juist een recht op uitleg hebben, wordt er gesproken over een normatieve assymetrie tussen mens en machine. Deze duidt erop dat de beslissingen die zijn gemaakt door een mens of door een machine verschillend zouden zijn. In de tekst wordt al snel gezegd dat een menselijke handeling aan de grond ligt van een geautomatiseerde beslissing. Maar als we naar machine learning kijken, de techniek waardoor machines een set voorbeelden krijgt en aan de hand van die set voorbeelden zichzelf leert welke beslissingen te maken, dan lijkt de grens op het eerste zicht te vervagen.

Om dit onderscheid te kunnen maken moet het eerst duidelijk zijn wat het verschil is tussen een geautomatiseerde beslissing en een menselijke beslissing. Dit is niet altijd duidelijk, want wat met een beslissing gemaakt aan de hand van een algoritme wat eigenlijk een set 'als dit, dan dat' regels zijn. Is er dan een verschil als de mens deze regels volgt en een beslissing maakt of als een machine deze regels volgt en een beslissing maakt? Deze manier van werken is enkel toepasselijk wanneer er beslissingen worden genomen volledig gebaseerd op objectieve feiten. Er zou dus ook kunnen gezegd worden dat de machine hier eigenlijk een middel is om de beslissing te automatiseren. De criteria voor de beslissing worden dus ook op voorhand genomen en dan in de machine gestoken. Wanneer we spreken over een geautomatiseerde beslissing als een gegenereerde output door een machine dan kunnen we dus zeggen dat een geautomatiseerde beslissing hetzelfde is als een menselijke beslissing enkel gebaseerd op feiten.

Maar wat met beslissingen gemaakt door machine learning algoritmen (algoritmen die zichzelf leren aan de hand van enkele gelabelde voorbeelden)? Opnieuw geldt hier hetzelfde principe. Machine learning is opnieuw een set van algoritmes die gebruik maken van parameters. Indien een mens deze algoritmes zou uitwerken en gebruik maken van dezelfde parameters dan krijgen we opnieuw dezelfde resultaten al is bij de ene het process naar deze resultaten geautomatiseerd en bij de andere niet.

Door op deze manier naar het verschil geautomatiseerde beslissing te kijken komen we dus eropuit dat een machine enkel beslissingen kan nemen gebaseerd op feiten en dat de mens dit op dezelfde manier dit ook kan. Het verschil zit hem dus dat een mens subjectieve argumenten kan bekijken bij het nemen van een keuze terwijl een machine dit niet kan. Verder zou dus ook kunnen

gezegd worden dat een geautomatiseerde beslissing een automatisering is van een menselijke beslissing enkel gebaseerd op feiten.

Als kijken naar de criteria van het algoritme voor het maken van enkel objectieve beslissingen dan zien we dat deze niet altijd gebaseerd is op enkel feiten en dus soms waardedragend is. Sommige criteria zijn zeer duidelijk objectief zoals bijvoorbeeld „als de temperatuur kleiner dan nul is dan bevriest water bij normale luchtdruk” maar dit is niet altijd het geval. Neem nu het voorbeeld uit het artikel van de lening. Stel dat er een criteria is voor de lening dat te maken heeft met de hoeveelheid geld dat iemand gespaard heeft tot op dat punt. Er zal nooit een punt zijn dat iemand met zekerheid een lening van een bepaald bedrag zal kunnen afbetalen en dus is in deze keuze van deze grens een zekere subjectiviteit.

Rekening houdende met deze argumenten is het mogelijk om een conclusie te vormen rond het verschil tussen een menselijke beslissing en een geautomatiseerde beslissing. In essentie zit het verschil tussen de twee dat een menselijke beslissing een vorm van subjectiviteit met zich meedraagt en dit niet zo is bij een geautomatiseerde beslissing. Verder kunnen we zeggen dat een geautomatiseerde beslissing eigenlijk een automatisering is van een menselijke beslissing die enkel gebaseerd is op feitelijke criteria. Maar er moet ook zeker rekening gehouden worden dat voor het tot stand komen van deze criteria vaak een vorm van subjectiviteit mee gepaard gaat. Deze redenering kan wel verder gebruikt worden zoals in het artikel om een verder moreel recht op uitleg af te leiden.

## 1.4 Moreel recht op uitleg

Aangezien er dus geen duidelijk verschil is tussen de beslissing die een mens maakt of de beslissing die een machine maakt kunnen we zoals aangegeven in het artikel een recht op uitleg afleiden uit een moreel recht op uitleg. In het artikel wordt gesproken over een recht op uitleg wanneer de beslissing iemands vrijwillige keuze of handeling moeilijk of onmogelijk maakt. Deze defenitie vind ik zelf niet volledig genoeg. Stel het voorbeeld uit het voorbeeld van het uitgeven van een lening. Volgens de defenitie uit het boek is er dus een recht op uitleg wanneer de bankier uitleg de lening weigert maar er geen recht op uitleg is over dezelfde beslissing wanneer de lening wordt toegestaan. Er wordt al snel aangevoeld dat er geen reden is waarom een recht op uitleg eerder toepasbaar is bij een nee dan een ja. Het lijkt mij wel logisch dat er eerder beroep zal gedaan worden op het recht op uitleg wanneer er een nee verkregen wordt dan

een ja.

In het artikel wordt er een grens getrokken in verband met de autonomie en het delibartief handelen of met andere woorden ofdat we een reden hebben voor onze (triviale) keuzes. Dit is opnieuw volgens mij een goed vertrekpunt maar wat met dingen die onverwachts gebeuren. Neem nu het voorbeeld van iemand die met de auto aan het rijden is en wordt tegengehouden door de politie en krijgt zomaar een bekeuring maar mag wel doorrijden. Als we de redenering in het boek zouden volgen hebben we hier dus geen recht op uitleg aangezien we zomaar door mogen rijden en onze autonomie dus niet wordt aangetast. In dit voorbeeld valt wel aan te merken dat als de bekeuring te groot wordt waardoor bijvoorbeeld een reis moet worden afgelast, de autonomie terug beperkt wordt en dus het principe van het deliberatief handelen wel terecht is. Het recht op uitleg wanneer onze autonomie beperkt wordt is dus onderdeel van het geheel maar niet voldoende op zich.

## **1.5 Mama**