HRIA van Facebookpagina's van de overheid

Human Rights Impact Assessment van de verwerking van persoonsgegevens op Facebookpagina's van de overheid

Datum	18 november 2022
Status	Openbare versie

Colofon

DPIA door **Ministerie van Binnenlandse Zaken en Koninkrijksrelaties**

Turfmarkt 147
2511 DP Den Haag
Postbus 20011
2500 EA Den Haag
www.rijksoverheid.nl/bzk

Contact

Naam project **HRIA-rapport** over de verwerking van persoonsgegevens op Facebookpagina's van de overheid

Auteurs Winfried Tilanus
https://www.privacycompany.eu/

Inhoudsopgave

Samenvatting 4

1. Inleiding 6
   - 1.1 Facebookpagina's 6
   - 1.2 Human Rights Impact Assessment 6
   - 1.3 Mensenrechten 7
   - 1.4 Reikwijdte van deze HRIA 8

2. Analytisch kader 9
   - 2.1 Persoonlijke autonomie, identiteit en personalisering 9
   - 2.2 Vertekening in algoritmen 10
   - 2.3 Schendingen van mensenrechten buiten de EU 13

Assessment volgens het IAMA-model 15

3. Waarom – de beoogde effecten, doelen en doelstellingen 15
   - 3.1 Aanleiding en probleemdefinitie 15
   - 3.2 Doeleinden 15
   - 3.3 Publieke waarden 16
   - 3.4 Wettelijke grondslag 16
   - 3.5 Betrokkenen en verantwoordelijkheden 17

4. Het algoritme 18
   - 4.1 Deel A Data-input 18
   - 4.2 Deel B Algoritme-throughput 20

5. Implementatie, gebruik en toezicht – output 22
   - 5.1 Beslissingen op basis van output algoritme 22
   - 5.2 De rol van de mens in de beslissing 23
   - 5.3 Effecten van het algoritme 24
   - 5.4 Procedures 24
   - 5.5 Context 24
   - 5.6 Communicatie 25
   - 5.7 Evaluatie, auditing en borging 25

6. Mensenrechten 26
   - 6.1 Inleiding 26
   - 6.2 Analyse per aspect van mensenrechten 28

Conclusie 43

Samenvatting

Dit rapport in opdracht van de Nederlandse Ministerie van Binnenlandse Zaken en Koninkrijksrelaties beoordeelt de risico's voor de mensenrechten van het gebruik van Facebookpagina's door de Nederlandse overheid. Deze Human Rights Impact Assessment (HRIA) is gericht op de risico's voor de rechten van betrokkenen op non-discriminatie, vrijheid van gedachte, geweten en godsdienst en vrijheid van meningsuiting en informatie. Deze HRIA moet gelezen worden in combinatie met de (uitgebreide) Data Protection Impact Assessment (DPIA) die is uitgevoerd op dezelfde gegevensverwerking.

De conclusie van deze beoordeling is dat het gebruik van Facebookpagina's door Nederlandse overheidsorganisaties mogelijk veel invloed heeft op mensenrechten. Door het gebrek aan transparantie (waaronder zinvolle toegang tot de werkelijke verwerking van big data door Facebook) was het niet mogelijk de werkelijke invloed te beoordelen. Nederlandse overheidsorganisaties beschikken niet over de middelen om toegang tot en zeggenschap te krijgen over de door Facebook gebruikte algoritmen en deze te corrigeren/minimaliseren of uit te leggen in verband met bezoeken aan de content die zij publiceren op Facebookpagina's. Door het ontbreken van transparantie en middelen om de potentiële negatieve gevolgen te minimaliseren voor bezoekers van de pagina's van overheidsorganisaties moet worden aangenomen dat het gebruik van Facebookpagina's hoge risico's meebrengt voor mensenrechten.

Wanneer iemand een Facebookpagina van een Nederlandse overheidsinstantie bezoekt, kan dat bezoek leiden tot verandering van de personalisering van de tijdlijn en de advertenties die deze persoon ziet. Al het gedrag op pagina's wordt door Facebook opgeslagen, zoals het openen van een post, tijd besteed aan het lezen van een post, deze leuk vinden en/of doorsturen naar een vriend. Omdat het voor gebruikers onmogelijk is om alle updates van hun vrienden, van vrienden van vrienden en content van gevolgde pagina's op Facebook te lezen, gebruikt Facebook algoritmen om de content die het aan gebruikers toont te selecteren. Deze HRIA beschrijft zes oorzaken waardoor deze algoritmen vertekend kunnen zijn en concludeert dat er sprake is van vertekening door eenzijdige suggesties en dat vertekening door feedbackloops (doorlopende terugkoppeling) en door optimalisatiecriteria zeer waarschijnlijk aanwezig is. Deze conclusie is gebaseerd op de observaties van de suggesties van Facebook tijdens het (beperkte) onderzoek voor deze DPIA en HRIA: de algoritmen selecteerden anti-overheidsstandpunten die de meeste vind-ik- leuks kregen. Het eerste criterium om berichten met de meeste vind-ik-leuks voor te stellen, zorgt sowieso al voor een vooroordeel. Bovendien stuurt het in één bepaalde richting. Op den duur kan dit zichzelf versterken omdat de als eerste voorgestelde berichten de meeste kans maken op nieuwe kliks en vind-ik-leuks.

De algoritmen kunnen gebruikers sturen in de richting van minder diversiteit of vertegenwoordiging van minderheden en aanzetten tot bepaalde acties, meningen of gedachten, terwijl ze andere groepen discrimineren of buiten beeld laten.

Facebook heeft geen volledige informatie gegeven over de gegevens die het bedrijf gebruikt voor zijn algoritmische beslissingen en ook geen inzicht in de gehanteerde logica in antwoord op inzageverzoeken van individuele betrokkenen. Dit gebrek aan transparantie maakt het onmogelijk te beoordelen welke impact de personalisering van Facebook heeft op mensenrechten als de overheid gebruikmaakt van Facebookpagina's.

Buitenlandse overheden van buiten de EU kunnen impact hebben op de mensenrechten van bezoekers van een Facebookpagina van een Nederlandse overheidsinstantie. Overheidsinstanties kunnen openbaarmaking afdwingen door het gebruik van OSINT of ADINT voor het monitoren van individuele bezoeken aan Facebookpagina's van de overheid. Ze kunnen Facebook ook dwingen (of Facebook-medewerkers hacken/omkopen) om gegevens te verstrekken die Facebook uit bezoeken aan een overheidspagina heeft afgeleid. Dit kan leiden tot intimidatie of (cyber)aanvallen tegen individuele bezoekers van een overheidspagina of tot mensenrechtenschendingen wanneer deze mensen naar derde landen reizen. Zoals beschreven in de DPIA volgt uit de Schrems-II-jurisprudentie van het Europese Hof van Justitie dat er een reëel risico bestaat dat overheidsinstanties in de VS toegang krijgen tot door Facebook verzamelde persoonsgegevens.

Daarom heeft de Ierse toezichthouder een verbod uitgevaardigd op toekomstige doorgifte van gegevens van Facebookgebruikers uit de EU naar de VS.

De conclusie van deze HRIA is dat overheidsorganisaties moeten stoppen met het gebruik van Facebookpagina’s als communicatiemiddel, vanwege de grote risico's voor de mensenrechten van paginabezoekers. Dit kan veranderen als Facebook zich als zeer groot online platform moet houden aan nieuwe Europese regels over digitale diensten uit de Digitaledienstenverordening, maar de naleving door Facebook van deze regels moet worden beoordeeld.

1. Inleiding

Dit rapport in opdracht van het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties is een Human Rights Impact Assessment (HRIA) van de verwerking van persoonsgegevens op Facebookpagina's van de overheid. Deze HRIA is gericht op de risico's voor de rechten van betrokkenen op non-discriminatie, vrijheid van gedachte, geweten en godsdienst en vrijheid van meningsuiting en informatie. Het is een aanvulling op de Data Protection Impact Assessment (DPIA) van dezelfde gegevensverwerking en is gebaseerd op hetzelfde onderzoek als voor de DPIA is gedaan. Met het oog op de bondigheid wordt in dit rapport waar mogelijk verwezen naar de DPIA.

In oktober 2021 heeft Facebook zijn bedrijfsnaam gewijzigd in Meta. In dit rapport wordt toch nog de naam 'Facebook' gebruikt voor het socialemediaplatform, om verwarring te voorkomen met andere apps die Meta aanbiedt, zoals Instagram en WhatsApp.

1.1 Facebookpagina's

Vroeger stonden Facebookpagina's voor organisaties bekend als *Fan Pages*. Overheidsorganisaties kunnen en willen Facebookpagina's gebruiken om een breed publiek te bereiken. Facebook stelt overheidsinstanties in staat direct te communiceren met mensen op een manier waaraan zij gewend zijn, via een platform waar ze al veel tijd aan besteden.

Een Facebookpagina van een overheidsorganisatie kan worden bekeken door Facebookgebruikers en niet-Facebookgebruikers. Facebookgebruikers die een pagina leuk vinden of volgen, krijgen updates van die organisatie in hun *Nieuwsoverzicht*. Zelfs als ze de overheidspagina niet volgen, zien ze mogelijk een aanbeveling als hun vrienden de pagina volgen of een bericht op een dergelijke pagina leuk vinden.

1.2 Human Rights Impact Assessment

Een Human Rights Impact Assessment (HRIA) beoordeelt welke mensenrechten worden beïnvloed door het gebruik van een bepaald systeem in een bepaalde context. Een van de methoden hiervoor is een model voor de beoordeling van de impact van algoritmen op mensenrechten, 'Impact assessment voor mensenrechten bij de inzet van algoritmen' (IAMA).[1] Dit model is onderdeel van een toolbox voor ethisch verantwoorde innovatie, de 'Toolbox Ethisch Verantwoorde Innovatie'[2], gepubliceerd door de Nederlandse centrale overheid. Dit IAMA-model is uitsluitend gericht op de impact op mensenrechten van algoritmen en niet op de impact op mensenrechten door andere aspecten van een systeem. Deze HRIA volgt het model van de IAMA, maar breidt dit uit naar andere mechanismen die mogelijk gevolgen hebben voor mensenrechten, zoals de toegang van overheden van derde landen op door Facebook verwerkte persoonsgegevens als gevolg van bezoeken aan een pagina van de overheid.

[1] Impact Assessment Mensenrechten en Algoritmes (IAMA); https://www.rijksoverheid.nl/documenten/rapporten/2021/02/25/impact-assessment-mensenrechten-en-algoritmes

[2] Toolbox Ethisch Verantwoorde Innovatie; https://www.digitaleoverheid.nl/overzicht-van-alle-onderwerpen/nieuwe-technologieen-data-en-ethiek/publieke-waarden/toolbox-voor-ethisch-verantwoorde-innovatie/

Dit IAMA-model heeft een gestructureerde vragenlijst, onderverdeeld in 4 groepen:

Het doeleinde van de verwerking
De kenmerken van het algoritme en de gegevens die het gebruikt
De toepassing van het algoritme en het gebruik van de resultaten ervan
De door het algoritme beïnvloede mensenrechten

Deze HRIA volgt deze vragen. Omdat het IAMA-model in het Nederlands is, zijn de vragen en enkele andere delen van het model (voor het oorspronkelijk in het Engels geschreven onderzoek) vertaald. Nadere toelichtingen en de juridische achtergrond zijn te vinden in het IAMA-model.

1.3 Mensenrechten

Mensenrechten worden beschermd door verschillende verdragen en wetten.[3] Deze HRIA is voornamelijk gericht op het Europees Verdrag voor bescherming van de rechten van de mens[4], maar ook op het Handvest van de grondrechten van de Europese Unie[5] ('Handvest').

Het Europees Verdrag voor de bescherming van de rechten van de mens is toepasselijk recht in de 46 lidstaten van de Raad van Europa. Nederland behoort daartoe. Iedereen die meent dat een lidstaat deze rechten schendt, kan direct een zaak aanspannen bij het Europese Hof voor de mensenrechten. Binnen de EU zijn de fundamentele rechten gewaarborgd in het Handvest van de grondrechten van de Europese Unie. Dit Handvest geldt voor de Europese instellingen en de toepassing van Europese regelgeving in nationaal recht. Het Europese Hof van Justitie interpreteert het Handvest. Deze twee verdragen zijn belangrijk voor deze HRIA omdat zij voorgaan op het nationale recht.

Het IAMA-model vat de mensenrechten uit de verschillende verdragen samen in de volgende groepen rechten:

Aan een persoon toebehorende rechten
1 Persoonlijke identiteit / persoonlijkheidsrechten / persoonlijke autonomie
2 Sociale identiteit / relationele privacyrechten / relationele autonomie
3 Fysieke en geestelijke integriteit
4 Gegevensbescherming / informationele privacyrechten
5 Communicatierechten
6 Ruimtelijke privacyrechten
7 Eigendomsgebonden privacyrechten
8 Reputatierechten (recht op bescherming van eer en goede naam)
9 Gezonde leefomgeving (milieubescherming)
10 Sociale en economische rechten

Vrijheden
11 Vrijheid van meningsuiting
12 Vrijheid om informatie te ontvangen
13 Godsdienstvrijheid

[3] https://www.rijksoverheid.nl/onderwerpen/mensenrechten/mensenrechten-nederland
[4] https://www.coe.int/en/web/conventions/full-list?module=treaty-detail&treatynum=005
[5] https://eur-lex.europa.eu/legal-content/NL/TXT/HTML/?uri=OJ:C:2012:326:FULL

14 Vrijheid van vergadering en demonstratievrijheid
15 Vrijheid van vereniging
16 Politieke rechten/vrijheden

Gelijkheidsrechten
17 Gelijkheid voor de wet
18 Verbod van direct onderscheid op bepaalde gronden
19 Verbod van indirect onderscheid op bepaalde gronden
20 Verbod van discriminatoir gemotiveerd handelen
21 Recht op materieel onderscheid/maatwerk
22 Recht op redelijke accommodatie / positieve actie
23 Verbod van profilering
24 Verbod van segregatie

Procedurele rechten
25 Recht op behoorlijk bestuur (precontentieuze fase)
26 Recht op een effectief rechtsmiddel en toegang tot de rechter (contentieuze fase)
27 Recht op een onafhankelijke en onpartijdige rechter
28 Recht op beoordeling binnen redelijke termijn
29 Recht op een eerlijk proces
30 Recht op een eerlijk proces: strafrechtelijke waarborgen
31 Strafrechtelijk legaliteitsvereiste (geen straf zonder wet)

Bij het analyseren van de impact op mensenrechten volgt de HRIA deze lijst met mensenrechten uit het IAMA-model.

Zowel het Europees Verdrag voor de rechten van de mens (art. 14)[6] als de Nederlandse Grondwet (art. 1)[7] verbiedt discriminatie, maar geen van deze artikelen beperkt de mogelijke gronden voor discriminatie. Tegelijkertijd wordt onderscheid maken op een van deze gronden alleen aangemerkt als discriminatie als daarvoor "geen objectieve en redelijke rechtvaardiging" bestaat[8]. Het hangt dus af van de context of onderscheid maken op een bepaalde grond discriminatie is. Bij het analyseren van mogelijke discriminatoire effecten van een systeem moet eerst bepaald worden welke gronden voor het maken van onderscheid discriminatoir is binnen de context van de toepassing van dat systeem. In deze HRIA wordt de term 'beschermde kenmerken' gebruikt voor alle gronden voor discriminatie binnen de context van het geanalyseerde systeem.

1.4 Reikwijdte van deze HRIA

Dit HRIA-rapport beoordeelt de impact op mensenrechten wanneer een gebruiker een Facebookpagina bezoekt die is aangemaakt door een Nederlandse overheidsorganisatie. De reikwijdte omvat zowel geregistreerde gebruikers van een Facebookpagina als bezoekers van een pagina van de overheid zonder Facebookaccount (niet-gebruikers).

Het onderzoek omvat het verzamelen van gegevens van buiten het platform over niet-gebruikers (met cookies), aangezien het mogelijk is dat deze mensen op zoek zijn naar overheidsinformatie die alleen te vinden is op Facebook of dat ze onbedoeld terechtkomen op een openbare Facebookpagina als gevolg van een zoekopdracht, zonder de voorwaarden van Facebook geaccepteerd te hebben. Deze vorm van

[6] https://www.echr.coe.int/documents/convention_nld.pdf
[7] https://wetten.overheid.nl/jci1.3:c:BWBR0001840&hoofdstuk=1&artikel=1&z=2018-12-21&g=2018-12-21
[8] Zie bijvoorbeeld J.D. en A. tegen het Verenigd Koninkrijk, ECLI:CE:ECHR:2019:1024JUD003294917

gegevensverwerking valt onder het onderzoek omdat de gegevensverwerking (met het cookie) het gevolg is van een bezoek aan een overheidspagina. Dit is een relevante verwerking van persoonsgegevens van bezoekers van een pagina.De reikwijdte omvat de verwerking die voor Facebook noodzakelijk is om aan bezoekers van de Facebookpagina van de overheid aanbevelingen te tonen, waaronder aanbevelingen gecreëerd door het algoritme van Facebook om content te rangschikken op basis van afgeleide voorkeuren. Dit valt onder het onderzoek, omdat de gegevensverwerking plaatsvindt op een pagina van de overheid en het gevolg is van algoritmische verwerking van persoonsgegevens van bezoekers van de pagina.

Zoals beschreven in de DPIA werd ten behoeve van deze DPIA en HRIA een testpagina aangemaakt, het (fictieve) Ministerie van Privacy.

Deze HRIA is gebaseerd op het onderzoek en de analyse van de interacties met de testpagina die zijn beschreven in de DPIA van Facebookpagina's van de overheid.

2. Analytisch kader

Voor een goed begrip van de potentiële impact op mensenrechten van het gebruik van Facebookpagina's door overheidsorganisaties is het noodzakelijk het IAMA-model aan te vullen met enige theoretische achtergrond.

Mensenrechten zijn op de eerste plaats gekoppeld aan persoonlijke identiteit en autonomie. Intuïtief gezien is het duidelijk dat er een verband bestand tussen identiteit en autonomie aan de ene kant en personalisering aan de andere kant. Om echter de potentiële impact van personalisering op de persoonlijke identiteit en autonomie te beoordelen, is er een theorie nodig over de relatie tussen deze concepten. Het eerste deel van dit analytisch kader biedt een dergelijke theorie.

Op de tweede plaats kan vertekening in algoritmen een grote impact hebben op mensenrechten. Het is echter moeilijk te beoordelen of een algoritme vertekend is. Voordat de impact op mensenrechten wordt beoordeeld, is het dus nodig te weten waar deze vertekeningen in algoritmen vandaan komen, hoe ze te achterhalen zijn en hoe de impact kan worden beperkt. In het tweede deel van dit analytisch kader komen de verschillende vormen van vertekening in algoritmen aan de orde.

Op de derde plaats ligt de oorzaak van andere potentiële schendingen van mensenrechten in de wereldwijde toegankelijkheid van op Facebook gepubliceerde informatie. Instanties en bedrijven in derde landen kunnen informatie verzamelen over bezoeken aan een Facebookpagina van een overheidsorganisatie en deze informatie gebruiken tegen Europese burgers. In het derde deel van dit analytisch kader worden de verschillende manieren besproken om dergelijke schendingen van mensenrechten door niet-EU-landen te analyseren.

2.1 Persoonlijke autonomie, identiteit en personalisering

Om de relatie tussen persoonlijke autonomie, identiteit en personalisering te begrijpen, wordt in deze HRIA de analyse gebruikt die Stuart Hall heeft gemaakt van 'culturele identiteit'.[9] Stuart Hall verdeelt 'identiteit' in twee conflicterende processen, een foucaultiaans proces van het opleggen van identiteiten aan mensen en een freudiaans proces van zelfidentificatie. Het foucaultiaanse proces is een proces van machtsuitoefening, het freudiaanse proces is de bron van menselijk handelen.

[9] Stuart Hall; Who Needs 'Identity'? in: Questions of Cultural Identity; bewerkt door Stuart Hall en Paul du Gay (Londen, 1996).

De ontwikkeling van menselijke autonomie en menselijke identiteit hangen direct samen met dit freudiaanse proces van zelfidentificatie. Zelfidentificatie is altijd een proces van jezelf identificeren met iemand of iets waarin je jezelf herkent. Dit is alleen mogelijk als er voorbeelden beschikbaar zijn waarmee je je kunt identificeren, in het persoonlijke leven of openbaar in de media. Voor de ontwikkeling van menselijke autonomie en identiteit is de zichtbaarheid van verschillende groepen en minderheden in het persoonlijke leven en in de media dus essentieel.

Een sociaal netwerk biedt in potentie een platform waarop minderheden zichtbaar en vertegenwoordigd kunnen zijn. Tegelijkertijd kan de personalisering van een sociaal netwerk de diversiteit van mensen die zichtbaar wordt sterk beperken. Dit kan zelfs leiden tot een foucaultiaans proces van het opleggen van een identiteit aan iemand. Wanneer bijvoorbeeld het personaliseringsalgoritme van het sociale netwerk er een sterke voorkeur voor heeft om aan vrouwen achterin de twintig nieuwsoverzichten en afbeeldingen te tonen van gezinnen met jonge kinderen, wordt aan deze vrouwen de identiteit van 'gehuwde moeder' opgelegd. Andere identiteiten, zoals ongehuwde werkende vrouwen, verdwijnen uit de vertegenwoordiging. Dit verhindert vrouwen achterin de twintig zich te identificeren met andere identiteiten dan 'gehuwde moeder' en beperkt zo hun mogelijkheden om hun eigen identiteit te vormen en autonome keuzes te maken.

2.2 Vertekening in algoritmen

Bezoeken aan een Facebookpagina van de overheid kunnen impact hebben op sommige mensenrechten als de door Facebook gebruikte algoritmen vertekend zijn. Vertekening of *bias* is een neiging of vooroordeel voor of tegen een persoon of groep, met name op een manier die als oneerlijk wordt beschouwd.[10] Om de potentiële impact op mensenrechten van de toepassing van algoritmen te begrijpen, is het van belang de mogelijke bronnen van vertekening te herkennen. Voor deze HRIA is de analyse van datawetenschapsprocessen, algoritmen en vooringenomenheid van Tobias Baer gebruikt als analytisch kader voor vooringenomenheid in algoritmen.[11]

In deze HRIA wordt geen duidelijk onderheid gemaakt tussen algoritmen die zijn bedacht door datawetenschappers en algoritmen die gebruikmaken van machine-learning. Machine-learning algoritmen kunnen diversere en complexere data verwerken dan datawetenschappers, maar beide methoden zijn kwetsbaar voor dezelfde bronnen van vertekening en dezelfde systematische fouten. Mache-learning kan gevoeliger zijn voor bepaalde bronnen van vertekening door de grotere hoeveelheden verwerkte gegevens en de mindere mate van menselijk toezicht.

2.2.1 Bronnen van vooringenomenheid in algoritmen

Zowel wanneer een algoritme volledig wordt ontwikkeld door datawetenschappers als wanneer deze gebruikmaken van machine-learning, hebben ontwikkelaars data nodig om te bepalen welke keuzes het algoritme moet maken. Als deze data een bias, wordt die vertekening meegenomen in het algoritme, wat leidt tot vertekende resultaten van het algoritme. Vertekeningen hebben meerdere oorzaken. Hieronder worden zes oorzaken wat uitgebreider toegelicht.

Op de eerste plaats bestaan er in onze maatschappij talrijke vooroordelen, bijvoorbeeld op grond van geslacht of leeftijd. Deze vooroordelen worden meegenomen in de data die worden gebruikt voor het ontwikkelen van het algoritme en daardoor zal het algoritme ook vooringenomen zijn. Het tegengaan van die vertekening betekent het tegengaan van vooroordelen in onze maatschappij.

[10] David Marshall; "Recognizing your unconscious bias," Business Matters, www.bmmagazine.co.uk/in-business/recognising-unconscious-bias/ (22 oktober 2013)

[11] Tobias Baer; Understand, Manage, and Prevent Algorithmic Bias: A Guide for Business Users and Data Scientists (Kaufbeuren, Duitsland, 2019)

Een voorbeeld van een dergelijk vooroordeel is dat vrouwen moeder worden als ze een bepaalde leeftijd bereiken. Dit vooroordeel kan leiden tot een vertekening in het algoritme, wat leidt tot een voorkeur voor aan moederschap gerelateerde onderwerpen in de aan Facebookgebruikers voorgestelde content. Om deze vertekening tegen te gaan, zou het algoritme opzettelijk berichten kunnen voorstellen die gerelateerd zijn aan vrouwen zonder kinderen. Het tegengaan van deze vorm van vertekening is een politieke beslissing.

Op de tweede plaats lopen de data die worden gebruikt voor het ontwikkelen van algoritmen achter op de ontwikkelingen in de maatschappij, wat ertoe leidt dat algoritmen enigszins ingesteld zijn op het handhaven van de bestaande situatie, in plaats van op het volgen van nieuwe trends. Deze vertekening kan worden tegengegaan door een algoritme continu nieuwe data te voeren en het te ontwerpen op het oppakken van nieuwe trends.

Op de derde plaats kan het algoritme vertekend raken als er een feedbackloop is tussen de resultaten van het algoritme en de data gebruikt voor het creëren van het algoritme. Dit gebeurt bijvoorbeeld als een lijst met suggesties wordt gerangschikt op basis van het aantal kliks op elk onderdeel. Wanneer het algoritme wordt aangepast op basis van het huidige aantal kliks op elk onderdeel, ontstaat er een feedbackloop, omdat de lijst al gerangschikt was en mensen de neiging hebben voornamelijk te klikken op de onderdelen die bovenaan staan. Dit leidt tot een vertekening die een kleine trend steeds meer uitvergroot bij elke toepassing van het algoritme. Wanneer de rangschikking van suggesties op individueel niveau plaatsvindt, kan iedere gebruiker steeds verder in zijn/haar eigen bubble worden geduwd. Dit is alleen tegen te gaan door de dataset voor de ontwikkeling van een algoritme geheel los te verzamelen van de toepassing ervan.

Wanneer de voor het creëren van algoritmen gebruikte datasets vrij zijn van vertekening heeft een algoritme gewoonlijk hetzelfde resultaat bij vergelijkbare input. Wanneer een algoritme echter gevoelig is voor de situatie of bepaalde eigenschappen van een persoon, zal het de meeste suggesties echter in een bepaalde richting doen. Dit is een vierde bron van vertekening. Suggesties in een bepaalde richting doen kunnen zorgen voor het ontstaan van een 'bubble': een omgeving die neigt naar een bepaald onderwerp of mening, terwijl andere onderwerpen en meningen uit beeld raken. Tijdens het onderzoek voor de DPIA en HRIA werd deze vorm van vertekening zichtbaar, doordat testgebruiker B alle politieke partijen met een Facebookpagina volgde, maar het Nieuwsoverzicht van testgebruiker B toonde een toenemende hoeveelheid anti-regeringsberichten (DPIA, paragraaf 1.1.3 personalisering). Deze vorm van vertekening kan worden tegengegaan door opzettelijk suggesties toe te voegen over verschillende onderwerpen en verschillende meningen te tonen. In de praktijk is dit lastig uit te voeren, omdat het algoritme nog steeds moet kiezen welke andere onderwerpen of meningen het zal tonen.

Datawetenschappers moeten bij het ontwerpen van een algoritme keuzes maken over de gebruikte criteria voor optimalisering. Wanneer bijvoorbeeld het aantal kliks wordt gebruikt om de rangschikking van een link te bepalen, krijgen links met pakkende, click-baitachtige titels de voorkeur boven meer beschrijvende links. Dit zorgt ervoor dat het algoritme de voorkeur gaat geven aan click-baits en rage-baits. Wanneer de kwaliteit van een suggestie is gedefinieerd als 'buiten het gebruikelijke van de gebruiker, maar gekozen door de gebruiker', geeft het algoritme de voorkeur aan wijziging. Omdat de meeste mensen een gematigde mening hebben, zou een dergelijk algoritme gebruikers naar extremere meningen aan beide kanten van het midden gaan duwen. Deze vorm van vertekening is fundamenteel voor het ontwerpen van algoritmen en kan alleen worden tegengegaan als onderdeel van het ontwerpproces van het algoritme in zijn bredere context.

Tijdens de ontwikkeling van een algoritme moeten datawetenschappers tot slot kleinere keuzes maken die vertekening kunnen veroorzaken. Als onderdeel van het normale proces van de voorbereiding van data voor gebruik wordt deze opgeschoond. Dat betekent dat 'uitschieters' (punten met extreem afwijkende waarden) worden gewist. Maar wanneer heeft een punt extreem afwijkende waarden? De datawetenschappers moeten beslissen welke punten ze behouden en welke worden weggegooid. Door dergelijke beslissingen kunnen datawetenschappers nieuwe vertekening introduceren.

Het is ook gebruikelijk om het algoritme verder af te stemmen op basis van een beoordeling van de resultaten. Tijdens dit proces moet de datawetenschappers beslissen welke optimalisaties correct zijn, waardoor mogelijk opnieuw vertekening wordt geïntroduceerd. Dit soort vertekening tegengaan is lastig, omdat er altijd sprake kan zijn van vertekening in de keuzes van datawetenschappers. Zorgvuldige keuzes door datawetenschapers die goed inzicht hebben in het algoritme en de context waarin dit wordt toegepast kunnen daarbij helpen, evenals publieke verantwoording van deze keuzes.

2.2.2 Directe en indirecte discriminatie

Van directe discriminatie is sprake als het kenmerk dat wordt gebruikt om twee mensen verschillend te behandelen, bv. geslacht of een fysiek kenmerk dat veel voorkomt bij mensen van een bepaalde afkomst[12] direct wordt geregistreerd en gebruikt in het algoritme om onderscheid te maken. Van indirecte discriminatie is sprake als andere informatie, die niet het kenmerk zelf bevat, maar daar nauw mee samenhangt, wordt gebruikt om onderscheid te maken. Een voorbeeld van dergelijke indirecte discriminatie, of discriminatie *by proxy,* is wanneer groepen mensen anders behandeld worden op grond van hun postcode. Omdat sommige woonwijken min of meer gescheiden zijn door inkomens en fysieke kenmerken, leidt het gebruik van de postcode tot indirecte discriminatie op basis van deze kenmerken. Datawetenschappers kunnen hun kennis proberen te gebruiken om alle variabelen te negeren die indirecte discriminatie kunnen veroorzaken. Bij het gebruik van machine-learning wordt dit ingewikkelder:

> "... wanneer de vertekening al aanwezig is in de data die is gebruikt voor het trainen van het model (bv. omdat het maatschappelijke vooroordelen weerspiegelt), probeert het machine-learning algoritme indicatoren voor de vertekening te bepalen. Wanneer je directe indicatoren (bv. de postcode) verwijdert, vindt het indirecte (bv. de afstand tot Piet Patat). Als je de indirecte indicatoren verwijdert, vindt het zelfs nog meer indirecte (bv. het aantal bedrijven in de buurt van de sollicitant waarvan de naam begint met "J"). Daarom is het niet altijd mogelijk de vertekening uit het model te krijgen..."(Bear, p. 156)

Indirecte discriminatie kan worden achterhaald door middel van statistische analyse. Daarvoor moeten het beschermde kenmerk (bv. een fysiek kenmerk dat veel voorkomt bij mensen van een bepaalde afstamming) en de dataset of het resultaat van het algoritme aan elkaar gekoppeld worden. Dit beschermde kenmerk wordt vaak niet verzameld. Het kan zelfs verboden zijn dit te verzamelen.[13] Zonder toegang tot het beschermde kenmerk is het eigenlijk onmogelijk indirecte discriminatie te achterhalen.

[12] De term 'ras' of 'raciale herkomst' wordt niet gebruikt ter vermijding van de suggestie van aanvaarding van theorieën die erop gericht zijn het bestaan van afzonderlijke menselijke rassen vast te stellen. Zie overweging 51 van de AVG.

2.2.3 Beoordeling van vertekening

Beoordelen in hoeverre een algoritme vertekend is, kan gebeuren door nauwlettend toezicht en streng onderzoek bij het ontwikkelen van het algoritme of door observatie van het gedrag van het algoritme. In beide gevallen moet het vooraf duidelijk zijn welke vormen van vertekening onwenselijk zijn en welke gewenst. Dit is een politieke keuze, hoewel mensenrechten al een goed uitgangspunt vormen. Om het algoritme te beoordelen, moet dit worden beoordeeld los van de systemen en gebruikersinterfaces waarmee het is verbonden. Het algoritme zou herhaaldelijk moeten worden gedraaid in een testomgeving, met kleine veranderingen in de inputdata. Toegang tot de onderliggende data, afleidingen en beslissingen is ook nodig voor de analyse of de dataset kenmerken bevat die indirect discriminerend zijn. Zonder een dergelijke laboratoriumopstelling is het onmogelijk volledig inzicht te krijgen in het gedrag van een algoritme.

2.3 Schendingen van mensenrechten door niet-EU-landen

De meest vergaande manier om mensenrechten af te dwingen binnen de EU is een beroep op het Europees Verdrag voor bescherming van de rechten van de mens en het Handvest van de grondrechten van de Europese Unie.

Aangezien Facebook wereldwijd actief is, zijn door Facebook verwerkte gegevens mogelijk toegankelijk in landen die geen lid zijn van de Raad van Europa. Deze wereldwijde toegankelijkheid kan op verschillende manieren gevolgen hebben voor de mensenrechten van bezoekers van pagina's van de Nederlandse overheid:

- Wanneer een bezoeker van een pagina van een Nederlandse overheidsinstelling naar een land reist dat geen lid is van de Raad van Europa, kan die bezoeker op grond van het plaatselijke strafrecht vervolgd worden. Een man die bijvoorbeeld openlijk berichten van de overheid over homoseksualiteit leuk vond, kan worden vervolgd in een land waar homoseksuele gedragingen verboden zijn.

- Wanneer een vijandige overheidsactor van buiten de Raad van Europa direct een bezoeker van een pagina van een Nederlandse overheidsinstelling stalkt, intimideert of daartegen (cyber)aanvallen uitvoert, bijvoorbeeld tegen een dissident die naar Nederland is gevlucht en interactie heeft met een pagina van de IND (Immigratie- en Naturalisatie Dienst).

- Wanneer derde landen regels of normen opleggen aan Nederlandse bezoekers van een Facebookpagina van de overheid op een manier die in strijd is met mensenrechten, bijvoorbeeld door het niet voorstellen van berichten die volgens het Europese Verdrag voor bescherming van de rechten van de mens beschermd zijn door de vrijheid van meningsuiting. Bijvoorbeeld door het verwijderen van content van de overheid over seksuele gezondheid uit een tijdlijn, omdat deze als 'ongepast' wordt beschouwd in andere culturen.

Overheidsinstanties in derde landen kunnen Facebook gebruiken om informatie te verzamelen die nodig is om bezoekers van Nederlandse Facebookpagina's te vervolgen of aan te vallen. Dit kan op verschillende manieren gebeuren:

[13] Op grond van het algemene verbod op verwerking van bijzondere categorieën van persoonsgegevens in de AVG (Algemene verordening gegevensbescherming), artikel 9.

- Door verstrekking af te dwingen van persoonsgegevens die buiten de lidstaten van de Raad van Europa zijn opgeslagen of door wetten met extraterritoriale reikwijdte. Deze gegevensbeschermingsrisico's zijn beschreven in de DPIA.

- Door middel van openbare informatie, bijvoorbeeld wanneer de douane nagaat welke overheidspagina's een reiziger volgt als deze de grens van een land passeert.

- Door middel van 'Advertisement Intelligence' (ADINT)[14]. ADINT stelt veiligheidsdiensten en andere overheidsactoren in staat persoonlijke aspecten van mensen af te leiden door middel van de advertentie-interfaces van Facebook. Deze interfaces lekken mogelijk gevoelige gegevens. De informatie die Facebook opslaat en afleidt, kan informatie bevatten over het bekijken of volgen van een overheidspagina op Facebook.

Door deze methoden kunnen sommige mensenrechten van bezoekers van Facebookpagina's van Nederlandse overheden nog steeds worden beïnvloed door niet-lidstaten.

[14] https://adint.cs.washington.edu/, ADINT wordt bijvoorbeeld gebruikt door de Nederlandse veiligheidsdiensten AIVD en MIVD, zie: https://www.ctivd.nl/documenten/rapporten/2022/02/08/rapport-74

Assessment volgens het IAMA-model

3. Waarom – de beoogde effecten, doelen en doelstellingen

Deel 1 van het IAMA-model gaat over het "Waarom" van het voornemen om een algoritme te ontwikkelen, in te kopen, aan te passen en/of in te zetten (hierna, voor de bondigheid: inzet van een algoritme). Wat zijn de redenen, de achterliggende ratio en de beoogde effecten van de inzet van een algoritme? Welke onderliggende waarden bepalen de benadering van het gebruik van het algoritme? Deze overkoepelende vragen moeten eerst beantwoord worden in een besluitvormingsproces over de inzet van algoritmen voordat bijvoorbeeld de vragen aan de orde komen over randvoorwaarden of mogelijke impact op fundamentele rechten.

3.1 Aanleiding en probleemdefinitie

Licht uw voorstel voor het gebruik/de inzet van een algoritme toe. Wat is de aanleiding hiervoor geweest? Voor welk probleem moet het algoritme een oplossing bieden?

Om goed te kunnen functioneren, moeten overheden actief communiceren met de bewoners van hun grondgebied. Het bereik kan enorm verbeteren wanneer de communicatie plaatsvindt via een platform dat pushberichten mogelijk maakt en via een platform met een groot publiek. Facebook is een dergelijk platform. De hoeveelheid informatie op Facebook is te groot voor een individuele gebruiker om gemakkelijk te verwerken. De algoritmen van Facebook helpen gebruikers de prioriteit te bepalen van de content op hun tijdlijn. De algoritmen helpen ook door voorstellen te doen voor andere content.

3.2 Doeleinden

Wat is het doel dat bereikt dient te worden met de inzet van het algoritme? Wat is hierbij het hoofddoel en wat zijn subdoelen?

De doeleinden voor gegevensverwerking van Facebook zijn uitgebreid beschreven in de DPIA (hoofdstuk 4, Doeleinden van de verwerking). De doeleinden van overheidsorganisaties voor het aanmaken van een Facebookpagina zijn beschreven in het antwoord op de eerste vraag.

De algoritmen worden gebruikt voor personalisering. In de DPIA worden verschillende vormen van personalisering met betrekking tot de Facebookpagina van de overheid beschreven:

1. Berichten getoond in het *Nieuwsoverzicht* van de specifieke mensen en organisaties die werden gevolgd door de twee testgebruikers, waaronder de testpagina van de overheid.
2. Berichten getoond in het *Nieuwsoverzicht* gebaseerd op de aanbevelingen van het algoritme van Facebook in verband met de dagelijkse bezoeken aan de testpagina van de overheid.
3. Advertenties getoond als 'gesponsorde inhoud'
4. Advertenties getoond als 'gesponsorde berichten'
5. Advertenties getoond als 'vergelijkbare pagina's' bij een bezoek aan de testpagina van de overheid

Deze vormen van personalisering worden bepaald door algoritmen.

3.3 Publieke waarden

Wat zijn de publieke waarden die de inzet van het algoritme ingeven? Indien meerdere publieke waarden de inzet van het algoritme ingeven, kan daar een rangschikking in aangebracht worden?

Facebook en zijn algoritmen worden ingezet om de enorme hoeveelheden informatie, waaronder de berichten van de overheid, toegankelijk te maken voor bestaande en potentiële Facebookgebruikers. Deze waarde is direct afgeleid van het mensenrecht van vrijheid om informatie te ontvangen.

Als neveneffect kan het gebruik van Facebook de volgende rechten ondersteunen:

1. Persoonlijke identiteit / persoonlijkheidsrechten / persoonlijke autonomie
2. Vrijheid van meningsuiting
3. Vrijheid van vergadering en demonstratievrijheid

(zie deel 4 van deze HRIA voor een uitgebreide analyse)

Wat zijn de publieke waarden die mogelijk in het gedrang komen door de inzet van het algoritme?

Van de 31 mensenrechten die zijn opgesomd in het IAMA-model is in deze HRIA vastgesteld dat het gebruik van overheidspagina's mogelijk een impact heeft op de volgende 21 mensenrechten (zie paragraaf 6.2 voor een uitgebreide analyse):

1. Persoonlijke identiteit / persoonlijkheidsrechten / persoonlijke autonomie
2. Fysieke en geestelijke integriteit
3. Gegevensbescherming / informationele privacyrechten
4. Communicatierechten
5. Ruimtelijke privacyrechten
6. Eigendomsgebonden privacyrechten
7. Vrijheid van meningsuiting
8. Vrijheid om informatie te ontvangen
9. Vrijheid van demonstratie
10. Verbod van indirect onderscheid op bepaalde gronden
11. Verbod van discriminatoir gemotiveerd handelen
12. Recht op redelijke accommodatie / positieve actie
13. Verbod van profilering
14. Verbod van segregatie
15. Recht op behoorlijk bestuur (precontentieuze fase)
16. Recht op een effectief rechtsmiddel en toegang tot de rechter (contentieuze fase)
17. Recht op een onafhankelijke en onpartijdige rechter
18. Recht op beoordeling binnen redelijke termijn
19. Recht op een eerlijk proces
20. Recht op een eerlijk proces: strafrechtelijke waarborgen
21. Strafrechtelijk legaliteitsvereiste (geen straf zonder wet)

3.4 Wettelijke grondslag

Wat is de wettelijke grondslag van de inzet van dit algoritme en van de beoogde besluiten die genomen zullen worden op basis van dit algoritme?

Zie hoofdstuk 11 van de DPIA. Facebook legt in zijn algemene Privacyverklaring (voor het laatst bijgewerkt op 26 juli 2022) aan zijn gebruikers uit dat het zich voor personalisering baseert op de rechtsgrond

noodzaak om een overeenkomst uit te voeren.[15] Er bestaan twee uitzonderingen: Facebook baseert zich op toestemming voor cookies voor het personaliseren van advertenties op en buiten Facebook en op uitdrukkelijke toestemming als het bijzondere categorieën gegevens die gebruikers actief verstrekt hebben aan Facebook.[16] Deze rechtsgrond van uitdrukkelijke toestemming is niet van toepassing op waargenomen of afgeleide bijzondere categorieën gegevens. In de DPIA is geconcludeerd dat Facebook geen rechtsgrond heeft voor de verwerking van afgeleide gevoelige gegevens.

Tegelijkertijd is in de DPIA geconcludeerd dat Facebook geen rechtsgrond heeft voor het plaatsen en lezen van volgcookies wanneer gebruikers of niet-gebruikers een overheidspagina bezoeken en geen toestemming geven voor marketingcookies.[17]

3.5 Betrokkenen en verantwoordelijkheden

Welke partijen en personen zijn er bij de ontwikkeling/het gebruik/het onderhoud van het algoritme betrokken?

Facebook ontwikkelt, gebruikt en onderhoudt de algoritmen. In de DPIA is beschreven dat beheerders en gebruikers de gegevensverwerking niet actief kunnen beïnvloeden of beperken. Gebruikers kunnen het algoritme alleen 'voeden' door pagina's te volgen/niet meer te volgen, content leuk of niet leuk te vinden en te klikken op links, maar dergelijke acties hebben geen directe invloed op de rangschikking in het algoritme. Beheerders van overheidspagina's kunnen verhinderen dat Facebook andere aanbevolen pagina's toont, maar beschikken niet over de mogelijkheid om bezoeken aan een overheidspagina uit te sluiten van het personaliseringsalgoritme.

Hoe zijn de verantwoordelijkheden belegd ten aanzien van de ontwikkeling en de inzet van het algoritme, ook nadat het algoritme eenmaal is afgerond?

De algoritmen worden ontwikkeld en ingezet onder verantwoordelijkheid van Facebook. Facebook biedt juridisch een gezamenlijke-verwerkingsverantwoordelijke-overeenkomst aan voor het produceren van statistieken over bezoeken aan een overheidspagina. Facebook splitst derhalve eenzijdig de verantwoordelijkheid voor deze specifieke output van de algoritmische gegevensverwerking. In de DPIA is beschreven dat deze overdracht onvolledig is, aangezien Facebook de verzamelde persoonsgegevens over bezoeken aan overheidspagina's zelfstandig blijft verwerken voor zijn eigen doeleinden, om zijn rangschikkingsalgoritme te 'voeden' en content te personaliseren.

---

[15] Meta Privacycenter, Privacyverklaring, Wat is onze rechtsgrond voor het verwerken van je gegevens, URL: https://www.facebook.com/privacy/policy/?section_id=7-WhatIsOurLegal. Facebook beschrijft dat deze rechtsgrond geldt voor:

- Het personaliseren van functies en inhoud (zoals je nieuwsoverzicht en verhalen);
- Het personaliseren van getoonde advertenties, en
- Het doen van aanbevelingen voor jou (zoals mensen die je mogelijk kent, groepen of evenementen waarin je mogelijk bent geïnteresseerd of onderwerpen die je wilt volgen) in en buiten onze producten.

[16] Meta Privacyverklaring, voor het laatst bijgewerkt op 26 juli 2022, Wat is onze rechtsgrond, URL: https://www.facebook.com/privacy/policy/?subpage=7.subpage.1-WhatIsOurLegal

[17] In de beschrijving van zijn doeleinden met een 'gerechtvaardigd belang' sluit Facebook een categorie gebruikers in genaamd '*Als je een apparaat gebruikt dat we niet kunnen koppelen aan een geregistreerde gebruiker van de Meta-producten*.' Facebook beschrijft inderdaad dat het zich baseert op deze rechtsgrond voor de verwerking van cookies, maar vermeldt het doeleinde van gedragsgerichte advertenties niet.

Wie is eindverantwoordelijk voor het algoritme?

Zie hoofdstuk 5 van de DPIA, verwerker of (gezamenlijke) verwerkingsverantwoordelijke.

4. Het algoritme

Als eenmaal vaststaat waarom een algoritme zal worden ingezet en hoe de organisatorische borging van publieke waarden en belangen vorm zal krijgen, is het zaak om na te denken over de vormgeving van het in te zetten algoritme. Daar gaat dit deel van het IAMA-model over, waarin de "Wat?" van het project aan de orde komt. Dit gedeelte is in twee delen gesplitst: deel A betreft de input van het algoritme: de data (of digitaal vastgelegde gegevens) die gebruikt gaat worden en de randvoorwaarden daarvoor. Deel B betreft het algoritme zelf, dus de throughput van het project. Net als het vorige deel omvat dit deel een aantal vragen en punten die in ieder besluitvormingsproces over de inzet van een algoritme moeten worden meegenomen.

4.1 Deel A Data-input

4.1.1 Inschatting: type algoritme

Wat voor type algoritme wordt gebruikt, of wat voor type algoritme gaat ontwikkeld worden?

Het gebruik van Facebook bestaat uit een cluster van algoritmen/taken dat interactie heeft met zijn gegevensopslagen (Hive en andere databases) en met elkaar. Facebook maakt gebruik van zelflerende en niet-zelflerende algoritmen. Facebook beschouwt de interne werking van zijn systemen als een bedrijfsgeheim. Zie hoofdstuk 8 van de DPIA voor meer informatie.

4.1.2 Databronnen en kwaliteit

Wat voor type data gaat gebruikt worden als input voor het algoritme en uit welke bronnen is de data afkomstig?

Facebook beschouwt de exacte interne werking van zijn systemen als een bedrijfsgeheim en maakt niet bekend welke input wordt gebruikt voor elk algoritme.

Is de kwaliteit en betrouwbaarheid van de data voldoende voor de beoogde datatoepassing? Leg uit.

Facebook beschouwt de exacte interne werking van zijn systemen als een bedrijfsgeheim en verstrekt geen informatie over de kwaliteit van de gebruikte data.

4.1.3 Bias en aannames in de data

Welke aannames en bias liggen in de data besloten en hoe wordt de invloed daarvan op de output van het algoritme gecorrigeerd of anderszins ondervangen of gemitigeerd?

Facebook beschouwt de exacte interne werking van zijn systemen als een bedrijfsgeheim en verstrekt geen informatie over biases in de data en verleent aan klanten of onafhankelijke wetenschappelijke onderzoekers geen toegang voor onderzoek naar bias en aannames in de data.[18] Facebook stelt wel beperkte data beschikbaar aan onderzoekers uit een archief van (politieke) advertenties (dat geldt ook voor Twitter en Google). Volgens onderzoekers van de Universiteit van Amsterdam was het in deze archieven echter "*lastig op schaal te definiëren en*

achterhalen wat een 'politieke advertentie' is. Op de tweede plaats, verificatie: advertentiearchieven bleken kwetsbaar voor niet-authentiek gedrag, vooral van advertentiekopers die hun ware identiteit of de herkomst van hun middelen trachten te verbergen. Op de derde plaats, doelgegevens: advertentiearchieven leggen niet op een zinvolle manier vast hoe advertenties gericht worden of verdeeld."[19] De dataset met politieke advertenties die Facebook verstrekt, is onvolledig omdat adverteerders zelf moeten melden als een advertentie een politieke advertentie is. Dit zorgt voor een onderrapportage van politieke advertenties.

De komende maanden, wanneer zeer grote platforms er formeel van op de hoogte zijn gesteld dat [zij] moeten voldoen aan specifieke transparantieregels uit de EU Digitaledienstenverordening (DSA), zal Facebook een advertentie-inventaris openbaar moeten maken.[20] Deze inventaris moet de gebruikte parameters bevatten voor het selecteren van groepen ontvangers en de gebruikte parameters voor het uitsluiten van een of meer van dergelijke bepaalde groepen.[21]

Indien gebruik wordt gemaakt van trainingsdata: is de data representatief voor de context waarin het algoritme ingezet gaat worden?

Facebook beschouwt de exacte interne werking van zijn systemen als een bedrijfsgeheim en verstrekt geen informatie over in hoeverre de data representatief zijn voor de context waarin de algoritmen worden gebruikt en het verleent evenmin toegang om dit te onderzoeken.

4.1.4 Beveiliging en archivering

Is de data voldoende beveiligd? Maak hierin onderscheid tussen de inputdata en de outputdata.

---

[18] Medium, Researchers, NYU, Knight Institute Condemn Facebook's Effort to Squelch Independent Research about Misinformation, 4 August 2021, URL: https://medium.com/cybersecurity-for-democracy/researchers-nyu-knight-institute-condemn-facebooks-effort-to-squelch-independent-research-about-59cec0793939

[19] Joren Vrancken, Theme Analysis of Political Facebook Ads in the 2021 Dutch General Election, January 2022, URL: https://www.researchgate.net/publication/357790703_Theme_Analysis_of_Political_Facebook_Ads_in_the_2021_Dutch_General_Election. Zie ook: Ausloos, J., Helberger, N., Leerssen, P., Vreese, C.H. de, Zarouali, B, Platform ad archives: promises and pitfalls, in: Internet Policy Review, vol. 8, no. 4 2019, URL: https://policyreview.info/articles/analysis/platform-ad-archives-promises-and-pitfalls

[20] De DSA is op 5 juli 2022 aangenomen door het Europees Parlement, geconsolideerde tekst via de URL: https://https://eur-lex.europa.eu/legal-content/NL/TXT/?uri=CELEX:32022R2065. Artikel 30 van de DSA bevat een verplichting voor zeer grote platforms zoals Facebook een register openbaar te maken van via hun netwerken verkochte advertenties met metadata over de demografische samenstelling van de doelgroep en hoe de advertentie gericht was. Zie: Paddy Leersen, Platform ad archives in Article 30 DSA, 25 May 2021, URL: https://dsa-observatory.eu/2021/05/25/platform-ad-archives-in-article-30-dsa/.

[21] Europese Commissie, Vragen en antwoorden: Digitaledienstenverordening. *"Na hun aanwijzing door de Commissie hebben aanbieders van zeer grote platforms en zeer grote onlinezoekmachines vier maanden tijd om aan de DSA te voldoen. Aanwijzing door de Commissie vindt plaats op basis van de door deze dienstverleners opgegeven gebruikersaantallen en dienstverleners krijgen na de inwerkingtreding van de DSA drie maanden de tijd om deze te verstrekken."* URL: *https://ec.europa.eu/commission/presscorner/detail/nl/QANDA_20_2348*

Facebook verstrekt geen nadere gegevens over zijn beveiligingsbeleid en evenmin over beveiligingsaudits en beveiligingscertificeringen. Zie de DPIA met betrekking tot pseudonimisering.

Is er controle op de toegang tot de data? Maak hierin onderscheid tussen de inputdata en de outputdata.

Facebook beschouwt de exacte interne werking van zijn systemen als een bedrijfsgeheim en verstrekt geen informatie over de toegangscontrole voor de gebruikte inputdata. In de privacyverklaring van Facebook staat: *"We delen de informatie die wij verzamelen wereldwijd, zowel intern in onze kantoren en datacentres, als extern met onze partners, leveranciers, dienstverleners en derde partijen."*[22] Facebook beschrijft wel welke gegevens het deelt met bepaalde partners, zoals adverteerders, maar in andere gevallen legt Facebook alleen uit "we delen informatie met..." Facebook vermeldt niet welke toegangscontroles worden uitgevoerd wanneer er toegang wordt verleend aan derde partijen.

Hoe worden relevante regels over archivering in acht genomen, zoals die in de Archiefwet zijn vastgelegd?

Facebook beschikt niet over faciliteiten om te voldoen aan de Archiefwet. Overheidsorganisaties zijn zelf verantwoordelijk voor het bewaren van kopieën van hun communicatie, voor zover dit verplicht is op grond van de specifieke archiveringsregels voor hun organisatie.[23] Zie hoofdstuk 10 van de DPIA voor algemene informatie over de door Facebook gehanteerde bewaartermijnen.

4.2 Deel B Algoritme-throughput

4.2.1 Type algoritme:

Type algoritme: wat voor soort algoritme wordt gebruikt of gaat worden gebruikt? Hoe werkt het? Onderscheid tussen: Een niet-zelflerend algoritme, waarin de mens de regels specificeert die de computer moet volgen en een zelflerend algoritme, waarin de machine zelf leert over de patronen in de data (machine learning)

Zie ook het antwoord op 4.1.1. Facebook beschouwt de exacte interne werking van zijn systemen als een bedrijfsgeheim en verstrekt geen informatie over de gebruikte algoritmen.

Waarom wordt voor dit type algoritme gekozen?

Waarom is dit type algoritme het meest geschikt om de bij vraag 3.2 geformuleerde doelstellingen te bereiken?

Welke alternatieven zijn er en waarom zijn die minder passend of bruikbaar?

Omdat het onbekend is welke algoritmen Facebook gekozen heeft, kunnen deze vragen niet worden beantwoord.

---

[22] Meta Privacycenter, Privacyverklaring, Hoe delen we gegevens met Partners, leveranciers, serviceproviders en derden? URL: https://www.facebook.com/privacy/policy/?section_id=4-HowDoWeShare .

[23] Zoals vermeld in een Basis Selectie Document voor elk ministerie.

4.2.2 Eigenaarschap en controle

Indien het algoritme is ontwikkeld door een externe partij: zijn er heldere afspraken gemaakt over eigenaarschap en beheer van het algoritme? Wat zijn die afspraken?

De algoritmen zijn eigendom en onder controle van Facebook. Overheidsorganisaties hebben geen controle over de updates van Facebook en zijn niet in staat uit te leggen hoe de algoritmen werken.

4.2.3 Accuraatheid van het algoritme

Wat is de accuraatheid van het algoritme, en op basis van welke evaluatiecriteria is deze accuraatheid bepaald?

Is de mate van accuraatheid (vraag 2B.3.1) acceptabel voor de manier waarop het algoritme ingezet zal worden?

Facebook maakt geen gegevens bekend over de accuraatheid van zijn algoritmen, waardoor deze twee vragen niet kunnen worden beantwoord.

Hoe wordt het algoritme getest?

Facebook maakt geen informatie bekend over zijn testprocedures.

Welke maatregelen kunnen worden getroffen om de risico's van reproductie of zelfs versterking van biases tegen te gaan (bijv. andere samplingstrategie, feature modification, ...)?

Welke aannames liggen ten grondslag aan de selectie en weging van de indicatoren? Zijn die aannames terecht? Waarom wel/niet?

Omdat het onbekend is welke algoritmen Facebook gekozen heeft en op welke wijze, kunnen deze vragen niet worden beantwoord.

Hoe vaak/erg zit het algoritme ernaast? (bijv. in termen van aantal false positives, false negatives, R-squared, ...)

Facebook maakt geen gegevens bekend over de accuraatheid van zijn algoritmen.

Dit zal veranderen op grond van de DSA. Zowel de aangewezen digitaledienstencoördinator als de Commissie en gescreende onderzoekers krijgen recht op toegang tot bepaalde data, met inbegrip van data met betrekking tot algoritmen. Dit omvat de data die noodzakelijk is om de risico en mogelijke schade te beoordelen die de systemen van het platform of zoekmachine veroorzaken, "*nauwkeurigheid, de werking en het testen van algoritmische systemen voor inhoudsmoderatie, aanbevelingssystemen of reclamesystemen, in voorkomend geval met inbegrip van trainingsgegevens en -algoritmen.*"[24]

4.2.4 Transparantie en uitlegbaarheid

Is het duidelijk wat het algoritme doet, hoe het dit doet, en op basis waarvan (welke data) het dit doet? Leg uit.

[24] Zie overweging 64 van de DSA, URL: https://digital-europe-website-v1.s3.fr-par.scw.cloud/uploads/2021/03/FINAL-DSA-Paper-March-2021-1.pdf.

De algoritmen zorgen voor de personalisering, maar alleen Facebook weet hoe ze dat doen en op basis van welke data. Artikel 29 DSA geeft het recht om gepersonaliseerde profilering te weigeren. Dit artikel luidt als volgt: *"Naast de in artikel 27 uiteengezette vereisten bieden **aanbieders van zeer grote onlineplatforms die gebruikmaken van aanbevelingssystemen ten minste één optie aan** voor elk van hun aanbevelingssystemen dat **niet gebaseerd is op profilering** zoals gedefinieerd in artikel 4, punt 4), van Verordening (EU) 2016/679."*
Indien Facebook deze optie invoert, met inbegrip van het gebruik van volgcookies en paginabezoekers eerlijk en duidelijk de keuze biedt om de optie zonder profilering te kiezen, kan Facebook het probleem van gebrek aan transparantie en uitlegbaarheid oplossen.

Voor welke personen en groepen binnen en buiten de eigen organisatie wordt de werking van het algoritme transparant gemaakt en hoe gebeurt dit?

Alleen Facebook heeft inzicht in de algoritmen. De recente collectieve rechtszaak in Californië (voorlopig geschikt zonder vonnis in augustus 2022[25]) suggereert dat het vanwege de complexiteit van de interne werking van de Facebooksystemen zelfs voor medewerkers van Facebook lastig is de werking van de algoritmen te begrijpen. Dit is onlangs bevestigd tijdens een hoorzitting van technici van Facebook georganiseerd door de Californische rechtbank als onderdeel van de schikking.[26]

Voor welke doelgroepen moet het algoritme uitlegbaar zijn?

1 Voor gebruikers met een Facebookaccount die een overheidspagina op Facebook bezoeken.
2 Voor gebruikers zonder een Facebookaccount die een overheidspagina op Facebook bezoeken.
3 Voor de overheidsorganisatie die de Facebookpagina publiceert, zowel vanuit het oogpunt van naleving van de wet en vanuit het oogpunt als werkgever (verantwoordelijkheid voor de beheerders van de overheidspagina).

Kan de werking van het algoritme voor de bij vraag B.4.3 geïdentificeerde doelgroepen op een voldoende begrijpelijke manier worden uitgelegd?

Nee, op dit moment niet. Mogelijk verandert dit wanneer de DSA-regels voor zeer grote platforms van kracht worden.

5. Implementatie, gebruik en toezicht – output

Een algoritme als zodanig zorgt niet voor ongewenste effecten. Die worden altijd veroorzaakt door implementatie, inzet of toepassing van het algoritme, door de context waarin het algoritme wordt ingezet en door de beslissingen en maatregelen die worden gekoppeld aan de output van het algoritme.

Dit deel gaat dan ook over de implementatie en het gebruik van het betreffende algoritme, dus over de (omgang met) de 'output' van het algoritme.

5.1 Beslissingen op basis van output algoritme

Wat gebeurt er met de uitkomsten van het algoritme? Welke beslissingen worden daarop gebaseerd?

[25] The Register, Facebook settles Cambridge Analytica class action for undisclosed amount, 29 August 2022, URL: https://www.theregister.com/2022/08/29/facebook_settles_cambridge_analytica_lawsuit/

[26] The Intercept, Facebook Engineers: We Have No Idea Where We Keep All Your Personal Data, 7 September 2022, URL: https://theintercept.com/2022/09/07/facebook-personal-data- no-accountability/

Facebook gebruikt de resultaten van de algoritmen voor de personalisering van:

1. Berichten getoond in het Nieuwoverzicht van personen of organisaties die de gebruiker wil volgen, bijvoorbeeld de plaats van de overheidspagina in het Nieuwsoverzicht.
2. Berichten getoond in het Nieuwsoverzicht gebaseerd op de aanbevelingen van het algoritme van Facebook, bijvoorbeeld aan wie de overheidspagina wordt aanbevolen of welke andere pagina's worden aanbevolen op de overheidspagina.
3. Advertenties die op de overheidspagina worden getoond als 'gesponsorde inhoud'.
4. Advertenties die op de overheidspagina worden getoond als 'gesponsorde berichten'.
5. Advertenties getoond als 'vergelijkbare pagina's' bij een bezoek aan de overheidspagina.

De overheidsorganisatie baseert geen beslissingen op de uitkomst van deze algoritmen.

5.2 De rol van de mens in de beslissing

Welke rol spelen mensen bij het nemen van beslissingen op basis van de output van het algoritme ('human in the loop') en hoe worden zij in staat gesteld om die rol te spelen?

Er zijn geen 'humans in the loop' wanneer Facebook de personalisering bepaalt.

5.2.1 Is er nu en in de toekomst voldoende gekwalificeerd personeel aanwezig om het algoritme te beheren, te herzien en aan te passen indien gewenst/nodig?

Facebook beschikt over voldoende gekwalificeerd personeel om de algoritmen te wijzigen. Volgens het laatste jaarverslag zijn bij Meta 71.970 mensen in dienst.[27]

5.3 Effecten van het algoritme

Wat zullen de effecten zijn van de inzet van het algoritme voor burgers en hoe wordt rekening gehouden met de 'menselijke maat' bij het nemen van beslissingen op basis van de output van het algoritme?

Er is te weinig bekend over de inzet van algoritmen door Facebook om op betrouwbare wijze de effecten te beoordelen van de inzet van algoritmen voor burgers. Facebook beweert nergens dat het rekening houdt met de 'menselijke maat' bij het nemen van beslissingen op basis van de output van het algoritme.

Welke risico's voor stigmatiserende, discriminerende of anderszins schadelijke of nadelige effecten zijn er voor de burger en hoe zullen die worden bestreden of gemitigeerd?

Er is te weinig bekend over de inzet van algoritmen door Facebook om op betrouwbare wijze de risico's te beoordelen van stigmatiserende, discriminerende of anderszins schadelijke of nadelige effecten voor burgers. Facebook verstrekt geen informatie over hoe het deze effecten mitigeert.

Dit zal veranderen op grond van de DSA. Zie bijvoorbeeld overweging 81 over een van de risico's die moet worden gemitigeerd: *"de werkelijke of voorzienbare impact van de dienst op de uitoefening van de grondrechten, zoals die zijn beschermd door het Handvest, met inbegrip van maar*

[27] https://www.annualreports.com/HostedData/AnnualReports/PDF/NASDAQ_FB_2021.pdf

niet beperkt tot de menselijke waardigheid, de vrijheid van meningsuiting en van informatie, met inbegrip van de vrijheid en pluriformiteit van de media, het recht op een privéleven, het recht op gegevensbescherming, ***het recht op non-discriminatie****, de rechten van het kind en de consumentenbescherming. Dergelijke risico's kunnen bijvoorbeeld voortvloeien* ***uit het ontwerp van de algoritmische systemen die door het zeer grote onlineplatform*** *of de zeer grote onlinezoekmachine worden gebruikt of uit het misbruik van hun dienst door het indienen van onjuiste meldingen of andere methoden om de vrijheid van meningsuiting te beknotten of de concurrentie te belemmeren."*

Hoe zullen de verwachte effecten bijdragen aan de oplossing van het probleem dat de aanleiding is voor de ontwikkeling/inzet van het algoritme (zie vraag 1.1) en het bereiken van de daarbij gestelde doelen (zie vraag 1.2)?

Er is te weinig bekend over de inzet van algoritmen door Facebook om op betrouwbare wijze hun bijdrage te beoordelen aan de oplossing van het aanvankelijke probleem en het bereiken van de doeleinden.

Hoe verhouden de verwachte effecten zich tot de waarden die worden gediend (zie vraag 1.3)? Welke risico's zijn er dat bepaalde waarden onder druk komen te staan en hoe wordt daarmee dan omgegaan?

Er is te weinig bekend over de inzet van algoritmen door Facebook om op betrouwbare wijze de effecten daarvan te beoordelen op de gediende waarden en de risico's dat waarden onder druk komen te staan. Omdat deze onbekend zijn, is het onmogelijk te zeggen hoe daarmee wordt omgegaan.

5.4 Procedures

Via welke procedures zullen beslissingen op basis van het algoritme worden genomen?

De exacte procedure die Facebook hanteert voor het nemen van beslissingen op basis van zijn algoritmen is onbekend.

Hoe worden verschillende relevante actoren (bestuurlijke en politiek verantwoordelijken, burgers) bij de besluitvorming betrokken?

Overheidsorganisaties en paginabezoekers zijn niet betrokken bij het besluitvormingsproces.

Hoe wordt gegarandeerd dat in deze procedures wordt voldaan aan de eisen van goed en behoorlijk bestuur en – waar nodig – rechtsbescherming? Hebben burgers een effectieve mogelijkheid om een klacht in te dienen of bezwaar te maken? Zo ja, op welke manier?

Overheidsorganisaties hebben geen zeggenschap over de procedures en geen middelen om te garanderen dat wordt voldaan aan de eisen van goed en behoorlijk bestuur. Facebook heeft aan de onderzoekers geen zinvolle informatie verstrekt over de algoritmen en besluitvorming in antwoord op de voor de DPIA ingediende inzageverzoeken van betrokkenen (paragraaf 2.5). Burgers beschikken derhalve niet over een effectieve manier om inzicht te krijgen in de genomen beslissingen en geen effectieve mogelijkheid om een klacht in te dienen.

5.5 Context

Tijd/periode: wanneer gaat het algoritme ingezet worden? Hoe lang is de periode dat het ingezet wordt?

Veel Nederlandse overheidsorganisaties maken al gebruik van Facebookpagina's. Daarvoor is geen einddatum bepaald.

Plaats: waar vindt inzet van het algoritme plaats? Is dat in een bepaald geografisch gebied, is dat bij een bepaalde groep personen of dossiers?

De algoritmen worden ingezet wanneer een persoon ook maar een overheidspagina op Facebook bezoekt.

Kan het algoritme ook nog worden ingezet als contextfactoren veranderen of als het algoritme gebruikt wordt in een andere context dan waarvoor het is ontwikkeld?

Er is geen documentatie over de context waarvoor de algoritmen zijn ontwikkeld. Facebook beschouwt de details over het algoritme als bedrijfsgeheim. Daardoor is het onmogelijk te beoordelen wat een verandering van de context is en of het algoritme nog steeds kan worden gebruikt.

5.6 Communicatie

Hoe open kunt u zijn over de werking van het algoritme in het licht van de doelstellingen en context van de inzet ervan?

Facebook beschouwt de werking van de algoritmen als een bedrijfsgeheim. Overheidsorganisaties met Facebookpagina's hebben geen informatie over de werking van de algoritmen. Het is niet mogelijk open te zijn over de werking van de algoritmen.

Op welke manier beoogt u te communiceren over de inzet van het algoritme?

In de bijgewerkte Facebook Privacyverklaring[28] staat niets over het gebruik van algoritmen, maar meerdere keren wordt vermeld dat er sprake is van geautomatiseerde verwerking, zonder nadere gegevens daarover. Overheidsorganisaties met Facebookpagina's kunnen niet zinvol communiceren over de werking van het algoritme.

Wordt de output van het algoritme gevisualiseerd, bijvoorbeeld in een tabel, grafiek of dashboard? Zo ja: is de vorm van visualisatie of weergave een correcte representatie van de output van het algoritme? Is de visualisatie makkelijk te lezen voor verschillende gebruikersgroepen?

De output van de algoritmen wordt niet gevisualiseerd.

5.7 Evaluatie, auditing en borging

Is voorzien in goede instrumenten voor evaluatie, auditing en borging van het algoritme?

Overheidsorganisaties met Facebookpagina's beschikken niet over goede instrumenten voor evaluatie, auditing en borging van de algoritmen.

Zijn er voldoende mogelijkheden om rekenschap en verantwoording af te leggen over het algoritme?

Overheidsorganisaties met Facebookpagina's beschikken niet over voldoende mogelijkheden om verantwoording af te leggen over de algoritmen.

[28] Facebook Privacyverklaring (van kracht sinds 26 juli 2022), URL: https://www.facebook.com/privacy/policy

Welke mogelijkheden zijn er voor auditors en toezichthouders om (formele) consequenties te verbinden aan de inzet van een algoritme door de overheid (bijv. terugkoppeling van bevindingen, doen van aanbevelingen, budgettaire consequenties, ...)?

Overheidsorganisaties met Facebookpagina's zijn niet in staat om op een zinvolle manier consequenties te verbinden aan feedback, bevindingen of aanbevelingen. Budgettaire consequenties zijn mogelijk wanneer aan de overheidsorganisatie een boete wordt opgelegd.

Dit zal veranderen na de inwerkingtreding van de DSA-regels voor zeer grote platforms zoals Facebook. Nationale digitaledienstencoördinatoren die samenwerken in een raad en de Europese Commissie zijn bevoegd tot handhaving. Zie hoofdstuk IV van de DSA, *Uitvoering, samenwerking, sancties en handhaving.* Daarnaast legt afdeling 5 (aanvullende verplichtingen voor aanbieders van zeer grote onlineplatforms en van zeer grote onlinezoekmachines om systeemrisico's te beheersen, art. 33 en verder) aan aanbieders van zeer grote platforms veel verplichtingen op met betrekking tot risicobeoordeling en onafhankelijke controle.

6. Mensenrechten

6.1 Inleiding

De eerste drie delen van het IAMA bevatten vragen en aandachtspunten die voor alle algoritmes relevant zijn en waarvan serieuze en goede bespreking eraan kan bijdragen dat algoritmes op een zorgvuldige, doordachte en goed ingebedde manier worden ingezet. Gelet op het grote belang van de bescherming van grondrechten, en de bijzondere risico's die kunnen bestaan voor de aantasting van die grondrechten door inzet van algoritmen, is het van belang om aan dit onderwerp afzonderlijk aandacht te besteden.

Dit deel van het IAMA-model bevat daarom een 'stappenplan grondrechten', dat een tweeledig doel heeft:

1 Het dient als hulpmiddel om te identificeren of het in te zetten algoritme grondrechten zal raken;
2 Zo ja, dan maakt het een gestructureerde discussie mogelijk over de vraag of er mogelijkheden zijn om deze grondrechtenaantasting te voorkomen of te mitigeren, en of er gronden zijn waarom de (al dan niet gemitigeerde) grondrechtenaantasting toch aanvaardbaar moet worden gevonden.

Verschillende stappen van het IAMA-model zijn ook relevant bij de beantwoording van de vraag welke publieke waarden door een algoritme kunnen worden gediend en geraakt.

Voor elk mensenrecht waarop de algoritmen impact hebben, stelt het IAMA-model zeven vragen:

1 **Mensenrecht:** welk (aspect van een) mensenrecht wordt geraakt?
2 **Specifieke wetgeving en richtlijnen:** is er ten aanzien van het betreffende grondrecht specifieke wetgeving van toepassing waarmee rekening moet worden gehouden? Zo ja, is daaraan voldaan?
3 **Zwaartebepaling:** welk aspect van het mensenrecht wordt geraakt en is de verwachte interferentie verreikend, beperkt of ligt deze daartussenin? Wat is de bijbehorende kleurcode (rood, oranje, groen)?
4 **Doelstellingen:** welke doelen worden nagestreefd? Zie het antwoord op vraag 1.2.
5 **Doeltreffendheid/effectiviteit/geschiktheid:** is inzet van dit specifieke

algoritme een doeltreffend, geschikt en effectief middel om deze doelen te bereiken?

6 **Noodzakelijkheid/subsidiariteit:** is inzet van dit specifieke algoritme noodzakelijk, d.w.z. zijn er geen andere of mitigerende maatregelen beschikbaar om dit te doen?

7 **Afweging/proportionaliteit (in strikte zin):** zijn de doelen – alles bij elkaar genomen – zwaarwegend genoeg om de aantasting van het mensenrecht te kunnen rechtvaardigen?

In deze HRIA is tot nu toe gebleken dat veel aspecten van de gebruikte algoritmen onbekend zijn en dat het niet mogelijk is de impact daarvan te beoordelen. Daardoor is het onmogelijk om vraag 2 tot en met 7 van dit model te beantwoorden. Daarom is voor deze HRIA gekozen voor een andere benadering en worden voor elk aspect van mensenrechten de volgende vragen beantwoord:

1 **Mogelijke impact:** heeft het gebruik van Facebookpagina's door de overheid mogelijk impact op dit aspect van mensenrechten? Zo ja, op welke manier?

2 **Hoe te beoordelen:** hoe kan de mogelijke impact op dit aspect van mensenrechten worden beoordeeld?

6.2 Analyse per aspect van mensenrechten

.1 Aan een persoon toebehorende rechten

Cluster	Voorbeelden	Mogelijke impact	Hoe te beoordelen?
Persoonlijke identiteit / persoonlijk-heidsrechten / persoonlijke autonomie	- Recht op zelfontplooiing - Vrijheid om het eigen gedrag te bepalen - Vrijheid om eigen uiterlijk vorm te geven (vrijheid van kleding) - Vrije beroepskeuze, studiekeuze, beroepskeuze etc. (vrijheid van beroepskeuze) - Respect voor de eigen identiteit (genderidentiteit / seksuele identiteit etc.)	Deze processen komen tot stand door identificatie (zie analytisch kader). Identificatie is alleen mogelijk wanneer verschillende groepen en minderheden zichzelf kunnen identificeren met, zichtbaar en vertegenwoordigd zijn. De personalisering van Facebook kan de diversiteit in de vertegenwoordiging vergroten en verkleinen en leiden tot verbetering of verslechtering van dit recht. Het beleid van Facebook dat gebruikers verplicht zijn hun account te registreren onder hun wettelijke naam leidt ertoe dat gebruikers betrokken worden op basis van kenmerken die mogelijk duidelijk samenhangen met die naam, zoals gender of afkomst en niet met hoe zij zichzelf identificeren of betrokken willen zijn bij andere mensen. Dit kan impact hebben op de manier waarop bezoekers op overheidspagina's reageren.	Test in een laboratoriumomgeving of het volgen van een overheidspagina leidt tot verschillen in de vertegenwoordiging van minderheden. Onderzoek naar hoe interacties op Facebook veranderen als mensen onder pseudoniem posten.

Cluster	Voorbeelden	Mogelijke impact	Hoe te beoordelen?
	- Reproductieve rechten - Recht op kennis van de eigen afstamming (recht om te weten wie de eigen biologische familie is) - Naamrechten (recht op respect voor iemands naam) - Contractvrijheid	Geen directe impact.	
Sociale identiteit / relationele privacyrechten / relationele autonomie	- Recht op respect voor familierelaties / gezinsleven (recht op respect voor privé- en gezinsleven) - Recht om te huwen - Recht op gezinsvorming - Recht om seksuele relaties aan te gaan - Recht om professionele/zakelijke relaties aan te gaan - Recht op toegang tot de arbeid/professie (recht om te werken) - Recht op toegang tot een land / verblijfsrechten - Recht op onderwijs	Geen directe impact.	

Cluster	Voorbeelden	Mogelijke impact	Hoe te beoordelen?
Fysieke en geestelijke integriteit	- Gewetensvrijheid / vrijheid van gedachte	Dit recht is afhankelijk van een sociale omgeving die openstaat voor andere standpunten over de wereld en andere benadering van ethiek en deze steunt. Een sociaal netwerk biedt in potentie een platform voor uiteenlopende meningen, maar de personalisering kan mensen in de richting van hun eigen bubble sturen. Het is onduidelijk in hoeverre het bezoeken of volgen van een Facebookpagina bijdraagt aan een dergelijke bubble.	Test in een laboratoriumomgeving of het volgen van een overheidspagina leidt tot toename of afname van vertegenwoordiging van verschillende standpunten in de personalisering.
	- Recht op leven - Verbod op foltering / onmenselijke of vernederende behandeling en bestraffing - Verbod op fouilleren: lijfsvisitatie (recht om gevrijwaard te zijn van onrechtmatige lijfsvisitaties)	Een impact hierop is mogelijk in extraterritoriale gevallen.	DTIA voor alle mogelijke gegevensoverdrachten en wetten van buiten de EU.
	- Verbod op refoulement - Toestemmingsvereiste bij medische behandeling en onderzoek - Recht op toegang tot de gezondheidszorg - Respect voor handelingsbekwaamheid - Recht op vrijwillige levensbeëindiging - Recht op abortus - Verbod op (moderne) slavernij/dienstbaarheid/gedwongen arbeid/mensenhandel/uitbuiting	Geen directe impact.	

Cluster	Voorbeelden	Mogelijke impact	Hoe te beoordelen?
Gegevensbe-scherming / informationele privacyrechten	- Bescherming tegen ongeautoriseerde/onzorgvuldige gegevensverwerking - Recht op toegang tot gegevens - Recht op rectificatie - Recht op vergetelheid	Zie DPIA	
Communicatie-rechten	- Briefgeheim (geheimhouding van correspondentie) - Bescherming tegen afluisteren/aftappen/interceptie - Verbod op ongeautoriseerde doorgifte van communicatie-gegevens - Vertrouwelijkheid communicatie met advocaat, arts etc.	Zie DPIA	
Ruimtelijke privacyrech-ten	- Bewegingsvrijheid - Habeas-corpusrechten (verbod van vrijheidsbeneming, huisarrest etc.) - Vrije woonplaatskeuze (vrijheid van woonplaats)	Een impact hierop is mogelijk in extraterritoriale gevallen.	DTIA voor alle mogelijke gegevensoverdrachten en wetten van buiten de EU.
	- Verbod op ongeautoriseerd volgen van personen (GPS-tracker)	Zie DPIA	
	- Vrijverkeersrechten (EU-recht) (Bewegingsvrijheid) - Recht om het land te verlaten - Verbod op ongeautoriseerd cameratoezicht	Geen directe impact.	

Cluster	Voorbeelden	Mogelijke impact	Hoe te beoordelen?
Eigendomsgebonden privacyrechten	- Bescherming tegen doorzoeking kleding/tassen/laptop/computer etc.	Een impact hierop is mogelijk in extraterritoriale gevallen.	DTIA voor alle mogelijke gegevensoverdrachten en wetten van buiten de EU.
	- Huisrecht (bescherming tegen invallen/doorzoekingen) - Vrije beschikking over eigendom (Recht op eigendom) - Bescherming tegen onteigening - Intellectuele eigendomsrechten	Geen directe impact.	
Reputatierechten (recht op bescherming van eer en goede naam)	- Verbod op strafbare belediging / smaad / laster - Bescherming van eer en goede naam	Hoewel een sociaal netwerk hierop grote impact kan hebben, lijkt het bezoeken of volgen van een Facebookpagina hierop geen impact te hebben.	
Gezonde leefomgeving (milieubescherming)	- Recht op duurzame ontwikkeling - Recht op milieubescherming - Bescherming tegen uitstoot van schadelijke stoffen - Recht op schoon drinkwater - Recht op toegang tot sanitatie (riolering) - Recht op toegang tot energie	Geen directe impact.	
Sociale en economische rechten	- Recht op een minimum bestaansniveau - Recht op sociale zekerheid en bijstand	Geen directe impact.	

Cluster	Voorbeelden	Mogelijke impact	Hoe te beoordelen?
	- Toegang tot onderwijs (Recht op onderwijs)		

.2 *Vrijheden*

Cluster	Voorbeelden	Mogelijke impact	Hoe te beoordelen?
Vrijheid van menings-uiting	- Persvrijheid/journalistieke vrijheid - Artistieke vrijheid - Vrijheid van wetenschap/academische vrijheid - Vrijheid van keuze van uitingsmiddel (mondeling/schriftelijk etc.) - Klokkenluiden (vrijheid om misstanden te melden) - Bescherming van journalistieke bronnen/verschoningsrecht	Een sociaal netwerk kan een platform zijn voor vrijheid van meningsuiting. Tegelijkertijd kan het eigen beleid van het platform de vrijheid van meningsuiting beperken. Ook kan een vertekening in de personalisering gebruikers in een bepaalde richting duwen. De zichtbaarheid van direct identificeerbare gebruikers kan zowel voor gebruikers als voor instellingen ook leiden tot een *beperkend* effect. Berichten kunnen ook gekleurd zijn door wat gebruikers denken te moeten doen om bovenaan te komen te staan. Het bezoeken of volgen van een Facebookpagina kan impact hebben op de personalisering en daardoor op het recht op vrijheid van meningsuiting.	Houd toezicht op gevallen waarin het eigen beleid van Facebook botst met de vrijheid van meningsuiting en leg deze vast. Test in een laboratoriumomgeving of het volgen van een overheidspagina leidt tot toename of afname van verschillende vertegenwoordigde standpunten in de personalisering. Gedragsonderzoek naar mogelijke afkoelende effecten.

Cluster	Voorbeelden	Mogelijke impact	Hoe te beoordelen?
Vrijheid om informatie te ontvangen	- Passieve informatieverzameling (recht op toegang tot bestaande informatie)	Door informatie te publiceren op een Facebookpagina waarop mensen al actief zijn, steunt de Nederlandse overheid dit recht actief. Aan de andere kant neemt de vrije toegang af, omdat Facebookpagina's slechts gedeeltelijk toegankelijk zijn zonder Facebookaccount.	
	- Actieve informatieverzameling (recht op toegang tot overheidsinformatie / openbaarheid van bestuur)	Geen directe impact zolang overheidsorganisaties alle informatie ook beschikbaar stellen op openbare internetpagina's, buiten het Facebookplatform.	
	- Plicht tot voorzien in pluriforme informatie	De personalisering van Facebook kan leiden tot eenzijdigere informatie.	Toegang voor gescreende wetenschappers tot de daadwerkelijke door Facebook verwerkte data met betrekking tot populaire overheidspagina's om te onderzoeken of het volgen van een overheidspagina door een gebruiker of door zijn/haar vrienden leidt tot toename of afname van verschillende vertegenwoordigde standpunten in de personalisering. Daarnaast moeten onderzoekers in staat zijn A/B-testen uit te voeren in een geïsoleerd laboratorium met modelaccounts. Op dit moment verbiedt Facebook het gebruik van testaccounts.

Cluster	Voorbeelden	Mogelijke impact	Hoe te beoordelen?
	- Recht op vrije toegang tot het internet (Recht op internettoegang)	Geen directe impact.	
Gods-dienst-vrijheid	- Vrijheid om een religie te hebben of niet te hebben - Religieuze uitingsvrijheid (symbolen, rituelen) - Vrijheid om samen te komen met andere gelovigen - Vrijheden van kerkgenootschappen/religieuze gemeenschappen - Scheiding van staat en religie (religieuze neutraliteit van de staat) - Respect voor religieuze/filosofische overtuigingen in het onderwijs	Geen directe impact.	
Vrijheid van vergadering en demonstratievrij-heid	- Vrijheid van vergadering - Samenkomsten, protestmarsen etc. - Vrije keuze van onderwerp, tijd, plaats en middelen - Bescherming tegen 'hostile audiences'	Sociale netwerken spelen een belangrijke rol bij het organiseren van protesten. Een bericht posten op een Facebookpagina maakt het gemakkelijker demonstraties over een bepaald onderwerp te organiseren. De personalisering kan gebruikers in de richting van radicalisering en radicalere demonstraties duwen.	Toegang voor gescreende wetenschappers tot de daadwerkelijke door Facebook verwerkte data met betrekking tot populaire overheidspagina's om te onderzoeken of het volgen van een overheidspagina leidt tot toename of afname van verschillende vertegenwoordigde standpunten in de personalisering. Daarnaast moeten onderzoekers in staat zijn A/B-testen uit te voeren in een geïsoleerd laboratorium met modelaccounts. Op dit moment verbiedt Facebook het gebruik van testaccounts.

Cluster	Voorbeelden	Mogelijke impact	Hoe te beoordelen?
Vrijheid van vereniging	- Vrijheid om al dan niet lid te zijn van een vereniging - Interne verenigingsvrijheid (eigen keuze leden, activiteiten) - Vrijheid van politieke partijen (vrijheid van vergadering en vereniging in politieke kwesties) - Vrijheid van vakbonden (vrijheid van vergadering en vereniging in vakbondskwesties) - Collectieve actierechten - Stakingsrecht	Geen directe impact.	
Politieke rechten/vrijheden	- Recht op periodieke organisatie van vrije en geheime verkiezingen (Recht op vrije verkiezingen) - Actief en passief kiesrecht - Petitierecht	Geen directe impact.	

.3 Gelijkheidsrechten

Cluster	Voorbeelden	Mogelijke impact	Hoe te beoordelen?
Gelijkheid voor de wet	- Gelijke toepassing van algemene wetgeving op iedereen die onder haar bereik valt - Verbod van willekeur - Consistentievereiste - Rechtszekerheidsbeginsel	Geen directe impact.	

Cluster	Voorbeelden	Mogelijke impact	Hoe te beoordelen?
Verbod van direct onderscheid op bepaalde gronden	Beslissingen of regels mogen niet in doorslaggevende mate zijn ingegeven door of gebaseerd op beschermde persoonskenmerken	Geen directe impact.	
Verbod van indirect onderscheid op bepaalde gronden	- Beslissingen of regels mogen niet disproportioneel benadelend uitpakken voor personen die behoren tot groepen met beschermde persoonskenmerken	De bezochte of gevolgde pagina of de daarop gelezen berichten kunnen samenhangen met beschermde kenmerken. Wanneer de bezochte pagina's gebruikt worden voor het bepalen van de personalisering, kan dit leiden tot indirect onderscheid.	Toegang voor gescreende wetenschappers tot de daadwerkelijke door Facebook verwerkte data met betrekking tot populaire overheidspagina's om te onderzoeken of het bekijken van bepaalde informatie op een overheidspagina leidt tot verschillen wanneer het algoritme kenmerken afleidt of groepen selecteert, bijvoorbeeld voor advertenties. Daarnaast moeten onderzoekers in staat zijn A/B-testen uit te voeren in een geïsoleerd laboratorium met modelaccounts. Op dit moment verbiedt Facebook het gebruik van testaccounts. Toegang voor gescreende onderzoekers tot daadwerkelijk door Facebook verwerkte gegevens om te analyseren of er een verband bestaat tussen door Facebook verwerkte niet-beschermde kenmerken en beschermde kenmerken.
Verbod van discriminatoir gemotiveerd handelen	- Verbod van racistisch/xenofoob etc. gemotiveerd handelen (bijv. geweldpleging) - Verbod van het geven van opdracht tot discriminatie	Geen directe impact.	

Cluster	Voorbeelden	Mogelijke impact	Hoe te beoordelen?
Recht op materieel onderscheid/maatwerk	- Plicht om rekening te houden met verschillen tussen mensen en groepen	Geen directe impact.	
Recht op redelijke accommodatie / positieve actie	- Recht op voorzieningen voor mensen met een handicap die hen in staat stellen gelijkwaardig in de samenleving te participeren	Facebook investeert in de toegankelijkheid van zijn pagina's. Facebook heeft geen toegankelijkheidscertificering. Communicatie met Facebookpagina's is niet toegankelijk voor mensen met geen of beperkte toegang tot het internet of technische vaardigheden.	Test op naleving van WCAG 2.1.
	Recht op compenserende maatregelen voor in het verleden ontstane structurele maatschappelijke ongelijkheid	Het is niet mogelijk de algoritmen van Facebook aan te passen om te compenseren voor ongelijkheden waarvoor de Nederlandse overheid besluit te compenseren. Het is mogelijk dat personalisering dergelijke ongelijkheden versterkt, bijvoorbeeld door mogelijkheden voor bepaalde groepen op een lagere plaats te zetten als deze gepubliceerd worden op een overheidspagina.	
Verbod van profilering	Verbod van het creëren van categorieën of profielen op basis van beschermde persoonskenmerken, die vervolgens de basis vormen voor besluitvorming of beleid (het recht niet te worden onderworpen aan besluitvorming die uitsluitend gebaseerd is op profilering)	Facebook vraagt gebruikers beschermde kenmerken zoals leeftijd en geslacht op te geven. Het is niet duidelijk in hoeverre deze beschermde kenmerken de personalisering van Facebook beïnvloeden.	Analyseren of er een verband bestaat tussen door Facebook verwerkte niet-beschermde kenmerken en beschermde kenmerken. Daarvoor is toegang tot de databases van Facebook noodzakelijk. In een laboratoriumomgeving testen of deze beschermde kenmerken en samenhangende

Cluster	Voorbeelden	Mogelijke impact	Hoe te beoordelen?
			kenmerken de personalisering beïnvloeden.
Verbod van segregatie	- Verbod van ruimtelijke of andere vormen van scheiding van groepen die daarbij wel een vergelijkbare behandeling krijgen	In potentie kan personalisering ertoe leiden dat groepen die een vergelijkbare personalisering binnen de groep zien, maar die afwijkt van de personalisering bij andere groepen omdat zij niets zien van andere groepen. Dit kan daadwerkelijk leiden tot segregatie van groepen. In hoeverre dit gebeurt en wat de invloed daarop is van het bezoeken of volgen van een overheidspagina is onbekend.	Maak een netwerkanalyse van de suggesties die worden gedaan aan een grote groep Facebookgebruikers. Test in een laboratoriumomgeving de impact van het volgen of bekijken van een overheidspagina op de netwerken. Voor het uitvoeren van deze testen moeten onderzoekers in staat zijn A/B-testen uit te voeren met modelaccounts. Op dit moment verbiedt Facebook het gebruik van testaccounts.

.4 Procedurele rechten

Cluster	Voorbeelden	Mogelijke impact	Hoe te beoordelen?
Recht op behoorlijk bestuur (precontentieu-ze fase)	Recht op transparantie en informatievoorziening	Door publicatie van informatie op een Facebookpagina maken Nederlandse overheidsorganisaties zichzelf transparant voor Facebookgebruikers. Tegelijkertijd kan personalisering ertoe leiden dat mensen niet op de hoogte worden gesteld van nieuwe informatie, waardoor de transparantie afneemt.	Analyseer welke groepen voorstellen voor de overheidspagina wel en niet zien als gevolg van personalisering.

Cluster	Voorbeelden	Mogelijke impact	Hoe te beoordelen?
	- Participatie- en verdedigingsrechten (bijv. hoorrechten) - Recht op zorgvuldige besluitvorming - Motiveringsplicht (recht op een gemotiveerde beslissing) - Verbod van willekeur - Verbod van misbruik/détournement de pouvoir	In geval van vervolging buiten de EU (zie analytisch kader) worden de door Facebook verzamelde gegevens, met inbegrip van de bezochte en gevolgde overheidspagina's mogelijk gebruikt voor vervolging in strijd met deze rechten.	DTIA voor alle mogelijke gegevensoverdrachten en wetten van buiten de EU.
Recht op een effectief rechtsmiddel en toegang tot de rechter (contentieuze fase)	- Bevoegdheid tot bieden van effectief rechtsherstel	Het is onduidelijk in hoeverre Facebook in staat is samen te werken als zijn betrokkenheid noodzakelijk is voor de naleving van dit mensenrecht.	Monitor en documenteer gevallen waarin Facebook onterechte beslissingen al dan niet corrigeert.
	- Verbod van hoge drempels (griffierecht, bijstand door advocaat, immuniteiten, termijnen)	De reactie op het inzageverzoek van betrokkene is niet volledig (zie DPIA), dus mogelijk moet informatie noodzakelijk voor een eerlijk proces worden opgevraagd door middel van een andere procedure; dit houdt mogelijk een schending van dit mensenrecht in.	
	- Recht op toegang tot een overheidsrechter - Recht op inhoudelijke behandeling - 'Full jurisdiction' (gehele zaak moet kunnen worden beoordeeld) - Ius de non evocandi (recht om niet te worden afgehouden van de rechter die iemand in een bepaald geschil toekomt) - Recht op gefinancierde rechtsbijstand	In geval van vervolging buiten de EU (zie analytisch kader) worden de door Facebook verzamelde gegevens, met inbegrip van de bezochte en gevolgde overheidspagina's mogelijk gebruikt voor vervolging in strijd met deze rechten.	DTIA voor alle mogelijke gegevensoverdrachten en wetten van buiten de EU.

Cluster	Voorbeelden	Mogelijke impact	Hoe te beoordelen?
	- Recht op effectieve tenuitvoerlegging rechterlijke uitspraak		
Recht op een onafhankelijke en onpartijdige rechter	- Persoonlijke onafhankelijkheid (bijv. ambtsduur) - Zakelijke/functionele afhankelijkheid (bescherming tegen druk van buitenaf) - Institutionele onafhankelijkheid - Subjectieve onpartijdigheid (geen betrokkenheid bij een van de partijen) - Objectieve onpartijdigheid (geen legitieme twijfel mogelijk over onbevooroordeelde beoordeling)	In geval van vervolging buiten de EU (zie analytisch kader) worden de door Facebook verzamelde gegevens, met inbegrip van de bezochte en gevolgde overheidspagina's mogelijk gebruikt voor vervolging in strijd met deze rechten.	DTIA voor alle mogelijke gegevensoverdrachten en wetten van buiten de EU.
Recht op beoordeling binnen redelijke termijn	- Gehele procedure (inclusief voorprocedure) mag niet te lang duren - Recht op mogelijkheden tot versnelling van de procedure - Recht op compensatie bij te lange duur van procedure	In geval van vervolging buiten de EU (zie analytisch kader) worden de door Facebook verzamelde gegevens, met inbegrip van de bezochte en gevolgde overheidspagina's mogelijk gebruikt voor vervolging in strijd met deze rechten.	DTIA voor alle mogelijke gegevensoverdrachten en wetten van buiten de EU.
Recht op een eerlijk proces	- Procedure op tegenspraak - Equality of arms (voorbereidingstijd, toegang tot dossiers/documenten) - Hoor en wederhoor - Evenwichtige en faire bewijsregels - Motiveringsplicht voor rechter	In geval van vervolging buiten de EU (zie analytisch kader) worden de door Facebook verzamelde gegevens, met inbegrip van de bezochte en gevolgde overheidspagina's mogelijk gebruikt voor vervolging in strijd met deze rechten.	DTIA voor alle mogelijke gegevensoverdrachten en wetten van buiten de EU.

Cluster	Voorbeelden	Mogelijke impact	Hoe te beoordelen?
	- Gelijke mogelijkheden tot horen getuigen / deskundigen - Openbaarheid zitting/vonnis - Rechtszekerheid - Verschoningsrechten (vertrouwelijkheid)		
Recht op een eerlijk proces: strafrechtelijke waarborgen	- Onschuldpresumptie - Zwijgrecht/recht om niet mee te werken - Uitlokverbod - Ne bis in idem (recht om niet tweemaal terecht te staan of te worden gestraft) - Recht op bijstand door een advocaat - Recht op bijstand door een tolk	In geval van vervolging buiten de EU (zie analytisch kader) worden de door Facebook verzamelde gegevens, met inbegrip van de bezochte en gevolgde overheidspagina's mogelijk gebruikt voor vervolging in strijd met deze rechten.	DTIA voor alle mogelijke gegevensoverdrachten en wetten van buiten de EU.
Strafrechtelijk legaliteitsvereiste (geen straf zonder wet)	- Verbod op terugwerkende kracht van strafwetgeving - Lex mitior-beginsel	In geval van vervolging buiten de EU (zie analytisch kader) worden de door Facebook verzamelde gegevens, met inbegrip van de bezochte en gevolgde overheidspagina's mogelijk gebruikt voor vervolging in strijd met deze rechten.	DTIA voor alle mogelijke gegevensoverdrachten en wetten van buiten de EU.

Conclusie

Het doel voor Nederlandse overheidsorganisaties van het gebruik van Facebookpagina's is vergroting van het bereik van hun communicatie. Wanneer iemand een Nederlandse overheidspagina bezoekt, gebruikt Facebook informatie over dit bezoek en mogelijke interacties met de inhoud op de pagina om een gepersonaliseerd *Nieuwsoverzicht* en gepersonaliseerde advertenties te tonen. Overheidsorganisaties hebben geen macht over deze personalisering. Facebook gebruikt een geautomatiseerd systeem met verschillende algoritmen, waarschijnlijk zelflerend en niet-zelflerend, voor het selecteren en rangschikken van de informatie die getoond wordt in het persoonlijke Nieuwsoverzicht en advertenties van iedere gebruiker. Facebook verstrekt geen gedetailleerde informatie over de ontwikkeling en interne werking van deze algoritmen, noch in openbare communicatie, noch in reactie op de voor de DPIA ingediende inzageverzoeken van betrokkenen. De conclusie van deze HRIA is dat veel vragen over de algoritmen niet beantwoord kunnen worden, omdat Facebook de benodigde informatie niet verstrekt.

Facebook is niet transparant over:

1. De gebruikte types algoritmen.
2. De gebruikte gegevensbronnen voor het samenstellen en/of trainen van algoritmen, de kwaliteit van die data, mogelijke vertekeningen daarin en in hoeverre die data representatief zijn.
3. Hoe zowel de input- als de outputdata beveiligd zijn en hoe lang deze worden bewaard.
4. De tijdens het ontwerp van de algoritmen gemaakte ontwerpkeuzes.
5. De accuraatheid van de algoritmen.
6. De testprocedures voor de algoritmen.

Facebook biedt overheidsorganisaties evenmin een interface om de beslissingen van de algoritmen structureel of incidenteel opzij te zetten. Het ontbreken van een interface maakt het voor overheidsorganisaties onmogelijk beslissingen te corrigeren of 'een menselijke schakel toe te voegen'. Facebook biedt ook geen hulpmiddelen voor overheidsorganisaties om de algoritmen te evalueren of te auditeren. Zonder dergelijke hulpmiddelen kunnen overheidsorganisaties aan hun bezoekers niet uitleggen hoe de algoritmen worden gebruikt. Zij kunnen er ook niet voor instaan dat de personalisering volgt op bevestigende acties van overheidsorganisaties of bewijs leveren wanneer een beslissing wordt aangevochten bij de rechter.

Dit gebrek aan transparantie en het ontbreken van auditmiddelen verhindert overheidsorganisaties de positieve dan wel schadelijke effecten van de door Facebook gebruikte algoritmen te beoordelen. Overheidsorganisaties kunnen niet beoordelen of mitigerende acties noodzakelijk zijn om schadelijke effecten tegen te gaan en zo ja, welke acties.

Wanneer een overheidsorganisatie Facebookpagina's gebruikt, heeft de personalisering van Facebook mogelijk impact op verschillende mensenrechten. Oorzaken van deze potentiële impacten zijn de mogelijkheid dat de algoritmen sturen in de richting van minder diversiteit of vertegenwoordiging van minderheden, de mogelijkheid dat de algoritmen gebruikers aanzetten tot bepaalde acties, meningen of denkwijzen of gebruikers richting extremisme duwen of omdat de personalisering mogelijk discriminerend is en groepen segregeert. Deze effecten hebben mogelijk impact op de volgende mensenrechten:

1. Persoonlijke identiteit / persoonlijkheidsrechten / persoonlijke autonomie
2. Gewetensvrijheid / vrijheid van gedachte
3. Vrijheid van meningsuiting
4. Vrijheid om informatie te ontvangen
5. Vrijheid van vergadering en demonstratievrijheid
6. Verbod van indirect onderscheid op bepaalde gronden

7. Verbod van profilering
8. Verbod van segregatie
9. Recht op transparantie en informatievoorziening

Overheidsorganisaties uit derde landen krijgen mogelijk data van bezoekers van Nederlandse overheidspagina's in handen. Overheidsinstanties kunnen openbaarmaking afdwingen door het gebruik van OSINT of ADINT voor het volgen van bezoeken aan of bezoekers van Facebookpagina's van de overheid. Ultiem kan dit leiden tot intimidatie of (cyber)aanvallen of oneerlijke processen of tot mensenrechtenschendingen wanneer deze bezoekers naar derde landen reizen. Mensenrechten die mogelijk geraakt worden door deze extraterritoriale effecten zijn:

1. Fysieke en geestelijke integriteit
2. Ruimtelijke privacyrechten
3. Bescherming tegen doorzoeking kleding/tassen/laptop/computer etc.
4. Recht op behoorlijk bestuur (precontentieuze fase)
5. Recht op een effectief rechtsmiddel en toegang tot de rechter (contentieuze fase)
6. Recht op een onafhankelijke en onpartijdige rechter
7. Recht op beoordeling binnen redelijke termijn
8. Recht op een eerlijk proces
9. Recht op een eerlijk proces: strafrechtelijke waarborgen
10. Strafrechtelijk legaliteitsvereiste (geen straf zonder wet)

Deze potentiële gevolgen voor mensenrechten staan los van de in de DPIA beoordeelde gegevensbeschermingsrisico's.

Uit de observaties in de DPIA blijkt dat er sprake is van vertekening door eenzijdige suggesties en dat er zeer waarschijnlijk sprake is van vertekening door feedbackloops en optimalisatiecriteria. De optimaliseringsvertekening wordt mogelijk veroorzaakt door het gebruik van het aantal vind-ik-leuks als optimaliseringscriterium. Dit effect kan zichzelf versterken omdat de belangrijkste voorgestelde berichten de meeste kans maken op nieuwe kliks en vind-ik-leuks.

Al met al is de conclusie van deze HRIA dat het gebruik van Facebookpagina's door Nederlandse overheidsorganisaties een grote kans heeft op impact op ten minste 9 mensenrechten, en nog eens 10 extra rechten wanneer data over bezoeken aan een pagina worden gebruikt door overheidsinstanties buiten de EU. Gelet op het gebrek aan transparantie van Facebook en het ontbreken aan middelen om te beoordelen, is het onmogelijk de werkelijke impact te meten. Aangezien overheidsorganisaties niet beschikken over middelen om de gebruikte algoritmen te beoordelen, te controleren, te corrigeren of uit te leggen, moet het gebruik van Facebookpagina's door de overheid aangemerkt worden als een hoog risico voor mensenrechten.

Deze situatie zal hopelijk begin 2023 verbeteren wanneer specifieke regels uit de Digitaledienstenverordening van kracht worden voor zeer grote platforms zoals Facebook. Voor een goede beoordeling van de rechtmatigheid, ethische impact en robuustheid[29] van de door Facebook gebruikte algoritmen voor personalisering is het cruciaal dat de Europese Commissie als aangewezen toezichthoudende instantie voor Facebook, evenals gescreende onderzoekers, toegang krijgt tot het ontwikkelingsproces van Facebook en de reeds gemaakte keuzes, tot de door Facebook opgeslagen data over bezoeken aan overheidspagina's, met inbegrip van de data gebruikt voor het produceren van Page Insights, evenals toegang tot een gecontroleerde omgeving waarin het gedrag van de algoritmen geobserveerd kan worden.

[29] https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai