

Algoritmische besluitvorming en de Awb

Korte feedback



Vraag 1. Wat zou de uitvoeringspraktijk, de rechtspraak en belanghebbenden in de zin van de Awb helpen om de individuele transparantie van algoritmegebruik bij besluitvorming in de zin van de Awb te bevorderen?

1. Duidelijke documentatie en verantwoording

Uitwerking: Overheidsinstanties zouden een standaardprotocol moeten ontwikkelen voor het documenteren van elke stap in het ontwerp- en implementatieproces van algoritmen. Dit omvat het loggen van beslissingsregels, gebruikte datasets, wijzigingen in het algoritme, en de redenen voor die wijzigingen. Deze documentatie moet toegankelijk zijn voor auditors, rechters, en eventueel betrokken burgers. Pragmatisch omgaan met deze aspecten helpt om de uiteindelijke content efficiënt, kort, duidelijk, en concreet te houden.

Voorbeeld: Finland heeft AI-transparantie richtlijnen ontwikkeld die vereisen dat overheidsorganisaties duidelijke documentatie over hun AI-systemen verstrekken. Dit omvat gedetailleerde beschrijvingen van de algoritmes, data die worden gebruikt, en de besluitvormingsprocessen. Deze richtlijnen bevorderen niet alleen transparantie maar ook publiek begrip van hoe beslissingen worden genomen.

2. Algoritmeregisters

Uitwerking: Een centraal, publiek toegankelijk register waarin alle door de overheid gebruikte algoritmen worden geregistreerd. Dit register moet informatie bevatten over het doel van elk algoritme, de inputdata, de beslissingslogica, de ontwerpers van het algoritme, en eventuele bijgewerkte versies. Het register zou ook links moeten bevatten naar relevante documentatie en auditrapporten.

Voorbeeld: Amsterdam en Helsinki hebben een openbaar algoritmeregister ingevoerd waarin wordt beschreven welke algoritmes door de stad worden gebruikt, hoe deze werken, en welke maatregelen er zijn genomen om bias en discriminatie te voorkomen. Dit houdt echter niet dat alle toegepaste algoritmen ook daadwerkelijk goed werken, maar levert op zo'n minst transparantie en ruimte voor feedback/verbetering,

3. Uitbreiding van de motiveringsplicht

Uitwerking: De huidige motiveringsplicht onder de Awb zou uitgebreid kunnen worden met een expliciete eis dat bij elk algoritmisch genomen besluit, er een duidelijke uitleg wordt gegeven over het gebruikte algoritme. Dit zou informatie moeten bevatten over de werking van het algoritme, hoe het besluit is beïnvloed door het algoritme, en op welke manier de gebruikte data het resultaat hebben beïnvloed.

Voorbeeld: De Europese Unie werkt aan de AI Act, die strenge eisen stelt aan de transparantie van 'high-risk' AI-systemen, inclusief de verplichting om gebruikers te informeren over de logica, significantie en gevolgen van AI-beslissingen. Ook de 'Data Act' is hieraan toegewijd, en zo zijn er meerdere 'Acts' bepaald vanuit de EU beleidsvorming, om zo de burgers beter te kunnen beschermen, en de mogelijkheid geven om te kiezen hoe met hun data wordt omgegaan bij het toepassen van algoritmen.

4. Training en bewustwording

Uitwerking: Ontwikkeling van trainingsprogramma's voor overheidsmedewerkers om hen te onderwijzen over de grondbeginselen van algoritmen en hun impact op besluitvorming. Deze trainingen zouden ook modules moeten bevatten over ethische overwegingen, risico's van bias, en methoden om transparantie te verzekeren.

Voorbeeld: De stad New York heeft een programma gelanceerd om ambtenaren op te leiden in het begrijpen en beheren van AI-tools binnen de overheid. Dit programma omvat trainingen over ethische overwegingen en technieken voor het identificeren van bias.

5. Recht op inzage en correctie

Uitwerking: Implementatie van mechanismen waarbij burgers het recht hebben om informatie te verkrijgen over de algoritmen die mogelijk impact hebben op hun leven. Dit zou ook een procedure moeten omvatten waarbij burgers een verzoek kunnen indienen voor correctie van fouten in de inputdata of de besluitvormingslogica.

Voorbeeld: Onder de Algemene Verordening Gegevensbescherming (AVG) in de EU, hebben individuen het recht om inzage te vragen in persoonlijke gegevens die door algoritmen worden verwerkt en, indien nodig, correctie of verwijdering van deze

gegevens te eisen.

6. Onafhankelijke audits

Uitwerking: Instelling van een onafhankelijke toezichthoudende instantie die regelmatig algoritmen audit. Deze audits zouden zich niet alleen moeten richten op de technische aspecten, maar ook op de naleving van wettelijke en ethische normen, en de impact van algoritmen op de samenleving.

Voorbeeld: In het Verenigd Koninkrijk houdt het Centre for Data Ethics and Innovation (CDEI) toezicht op het gebruik van AI en data-analytische tools binnen de publieke sector, waarbij ze regelmatig audits uitvoeren om compliance met ethische normen te verzekeren.

7. Maatschappelijk debat en participatie

Uitwerking: Organisatie van openbare fora, workshops en consultatiesessies waar burgers, experts en beleidsmakers samen kunnen komen om te discussiëren over het gebruik van algoritmen door de overheid. Dit zou gepaard moeten gaan met het publiceren van gemakkelijk te begrijpen informatie over algoritmen en hun rol in de maatschappij, mogelijk gemaakt door samenwerkingen met educatieve instellingen en mediapartners.

Voorbeeld: De Europese Commissie heeft uitgebreide publieke consultaties gehouden als onderdeel van het voorbereidingsproces voor de AI Act, waarbij burgers, bedrijven en andere belanghebbenden werden uitgenodigd om hun mening en zorgen over AI-regulering te delen.

Vraag 2. Wat wordt naast het huidige juridisch kader eventueel gemist in de Awb, gezien vanuit de waarborgfunctie daarvan en in het licht van de doelstellingen van deze wet?

1. Specifieke bepalingen over algoritmische transparantie en verantwoording

De Awb zou kunnen worden uitgebreid met specifieke bepalingen die eisen stellen aan de transparantie van het gebruik van algoritmen in de besluitvorming. Dit zou

bijvoorbeeld kunnen omvatten dat besluiten die met behulp van algoritmen worden genomen, duidelijk als zodanig moeten worden gekenmerkt en dat een beschrijving van het gebruikte proces en de onderliggende logica wordt meegeleverd.

2. Rechten op uitleg en bezwaar tegen algoritmische beslissingen

Hoewel de Awb voorziet in algemene bezwaar- en beroepsprocedures, zou er specifiek aandacht kunnen zijn voor de rechten van burgers om uitleg te vragen over algoritmische besluiten en om bezwaar te maken tegen besluiten die volledig zijn geautomatiseerd.

3. Richtlijnen voor het testen en valideren van algoritmen

Er zouden richtlijnen of normen kunnen worden ontwikkeld voor het testen en valideren van de algoritmen die gebruikt worden door overheidsinstanties, om te waarborgen dat deze algoritmen eerlijk, accuraat en vrij van discriminatoire bias zijn.

4. Bevordering van Explainable AI en aanpak van het Black Box-dilemma

Ontwikkeling van standaarden voor uitlegbare AI: De Awb zou kunnen vereisen dat algoritmen die gebruikt worden in overheidsbesluitvorming niet alleen transparant zijn, maar ook expliciet uitlegbaar. Dit betekent dat de algoritmen zo ontworpen moeten zijn dat ze inzichtelijk kunnen maken hoe specifieke inputs leiden tot bepaalde outputs.

Juridische vereisten voor XAI: Specifieke juridische vereisten kunnen worden opgesteld om ervoor te zorgen dat besluiten die door algoritmen worden ondersteund, een uitlegniveau bevatten dat voldoet aan de normen van de Awb voor begrijpelijke communicatie. Dit zou bijvoorbeeld kunnen betekenen dat een samenvatting van de besluitlogica moet worden verstrekt aan de betrokken personen.

Publiek toezicht en feedbackloops: Invoeren van mechanismen waarbij burgers feedback kunnen geven op de uitleg die zij ontvangen over algoritmische besluiten. Dit kan helpen om de methoden voor explainable AI continu te verbeteren en te verfijnen.

5. Integratie van Ethische AI Principles in de Awb

Om de waarborgfunctie van de Awb te versterken en de menselijke maat binnen de algoritmische besluitvorming te bevorderen, is het essentieel dat de volgende vijf ethische AI-principes worden geïntegreerd, aanvullend op de regelgeving die wordt ontwikkeld op EU-niveau zoals de AI Act en de Data Act:

- **Transparantie:** Zorg voor volledige openheid over hoe en waarom AI-systemen bepaalde beslissingen nemen, inclusief toegankelijke uitleg voor alle gebruikers en belanghebbenden.
- **Eerlijkheid:** Implementeer strikte maatregelen om ervoor te zorgen dat AI-systemen geen discriminerende bias vertonen en eerlijk omgaan met alle gebruikers, ongeacht hun achtergrond.
- **Privacy:** Bescherm de persoonlijke gegevens van gebruikers door sterke privacyrichtlijnen en beveiligingsmaatregelen te handhaven, en zorg voor naleving van gegevensbeschermingswetten.
- **Verantwoordelijkheid:** Definieer duidelijk wie verantwoordelijk is voor de acties van AI-systemen en stel mechanismen in voor het melden en adresseren van fouten of schade veroorzaakt door deze systemen.
- **Duurzaamheid:** Moedig de ontwikkeling en implementatie van AI-systemen aan die duurzaam zijn voor zowel het milieu als de samenleving, met aandacht voor hun lange-termijneffecten.

5. Voorbeelden van internationale initiatieven:

EU Guidelines on Ethics in Artificial Intelligence: De Europese Commissie heeft ethische richtlijnen opgesteld die aanbevelingen doen voor transparante en uitlegbare AI. Deze richtlijnen benadrukken het belang van transparantie en uitlegbaarheid in AI-systemen, vooral wanneer deze impact hebben op burgers.

Singapore's Model AI Governance Framework: Singapore heeft een raamwerk ontwikkeld dat specifiek gericht is op het beheersen van de risico's van AI, met sterke nadruk op uitlegbaarheid als een kerncomponent voor verantwoordelijk gebruik van AI in de publieke en private sector.

Vraag 3. Het kabinet vindt de Franse wetgeving over algoritmische besluitvorming interessante inspiratie. Bieden ontwikkelingen uit andere landen een goed voorbeeld, zoals de genoemde Franse wetgeving?

1. Australië

Centre for Responsible Technology: Australië heeft initiatieven onderzocht en voorgesteld via het Centre for Responsible Technology. Dit centrum biedt richtlijnen voor de ontwikkeling en implementatie van verantwoorde technologieën, inclusief algoritmen die door de overheid worden gebruikt. Het beleid richt zich op het bevorderen van transparantie, het waarborgen van gelijke toegang, en het tegengaan van discriminatie.

Automated Decision Making Better Practice Guide: Uitgegeven door de Australische overheid, deze gids bevat aanbevelingen voor het implementeren van geautomatiseerde besluitvormingssystemen, waaronder aspecten van transparantie en verantwoording die essentieel zijn voor eerlijkheid en nauwkeurigheid.

2. Zuid-Korea

AI Ethics Guidelines: Zuid-Korea heeft een set ethische richtlijnen voor AI ontwikkeld die principes bevatten zoals transparantie, verantwoordelijkheid, en controleerbaarheid. Deze richtlijnen zijn gericht op zowel private als publieke sector gebruik van AI, met de bedoeling om de ontwikkeling van AI te sturen in een richting die ethische overwegingen respecteert.

3. Singapore

Model AI Governance Framework: Singapore heeft een uitgebreid kader voor AI-governance ontwikkeld dat bedrijven en overheidsorganisaties richtlijnen biedt over hoe ze op verantwoorde wijze AI kunnen inzetten. Het framework benadrukt het belang van menselijke tussenkomst in AI-besluitvorming en het waarborgen van uitlegbaarheid in algoritmische processen.

4. Zweden

Swedish National Strategy for Artificial Intelligence: Zweden heeft een nationale strategie voor kunstmatige intelligentie ontwikkeld die focust op het verhogen van de competenties binnen de publieke sector voor het gebruik van AI. Onderdeel van deze strategie is het bevorderen van transparantie en het stellen van ethische richtlijnen voor AI-gebruik.

5. Verenigde Staten

Algorithmic Accountability Act: Hoewel deze wet nog in voorstel is, illustreert het de beweging in de VS om strengere regulering en controle op algoritmische besluitvorming

te implementeren. De wet zou bedrijven verplichten om impactbeoordelingen en bias-audits uit te voeren voor algoritmen die worden gebruikt in besluitvormingsprocessen.

6. Finland

AuroraAI - Nationaal AI-programma: Finland staat bekend om zijn progressieve benadering van technologie en digitalisering, inclusief het gebruik van kunstmatige intelligentie en algoritmen binnen de overheidsdiensten.

7. Estland

X-Road Platform: Het platform is een nationaal digitaal raamwerk dat veilige gegevensuitwisseling tussen overheidsinstanties, bedrijven en burgers mogelijk maakt. Dit systeem is fundamenteel voor het mogelijk maken van algoritmische processen die toegang tot gegevens uit verschillende bronnen vereisen, terwijl het hoge normen van gegevensbeveiliging en privacy handhaaft.

Vraag 4. Hoe kan de inzet van algoritmen bijdragen aan de waarborgfunctie van de Awb en het versterken van de menselijke maat?

1. Verbetering van de Toegankelijkheid en Snelheid van Diensten

Algoritmen kunnen de efficiëntie en toegankelijkheid van overheidsdiensten verbeteren door het automatiseren van routinematige processen. Dit maakt snellere reactietijden mogelijk en vermindert de wachttijden voor burgers, wat essentieel is voor een mensgerichte dienstverlening. Bijvoorbeeld, geautomatiseerde systemen voor het verwerken van aanvragen kunnen zorgen dat burgers sneller antwoord krijgen op hun vragen.

2. Consistentie en Objectiviteit in Besluitvorming

Algoritmen kunnen helpen om consistentie en objectiviteit te waarborgen in de besluitvorming. Door beslissingen te baseren op vooraf gedefinieerde regels en criteria, kunnen algoritmen de invloed van menselijke vooroordelen verminderen. Dit draagt bij aan de rechtvaardigheid en transparantie van overheidsbeslissingen, wat fundamenteel is voor de waarborgfunctie van de Awb.

3. Identificatie van Patronen en Behoeften

Geavanceerde data-analyse algoritmen kunnen helpen bij het identificeren van patronen en behoeften binnen grote hoeveelheden data die door de overheid worden verzameld. Dit kan overheidsinstanties in staat stellen om proactief in te spelen op de behoeften van burgers, door bijvoorbeeld gerichte diensten aan te bieden aan specifieke bevolkingsgroepen.

4. Versterking van Participatie en Feedback

Algoritmen kunnen ook worden gebruikt om de participatie van burgers te vergroten. Digitale platforms die algoritmische tools gebruiken, kunnen burgers de mogelijkheid bieden om feedback te geven, deel te nemen aan beleidsvorming en transparantie-eisen te stellen. Dit versterkt de democratische betrokkenheid en helpt de overheid om diensten beter af te stemmen op de behoeften van de burger.

5. Ondersteuning van Menselijke Besluitvormers

In plaats van menselijke besluitvorming volledig te vervangen, kunnen algoritmen worden ingezet als ondersteunende tools die besluitvormers voorzien van diepgaande analyses en aanbevelingen. Dit stelt overheidsmedewerkers in staat om meer geïnformeerde beslissingen te nemen en hun aandacht te richten op complexe zaken die een menselijke beoordeling vereisen, waardoor de menselijke maat in de besluitvorming wordt versterkt.

Vraag 5. Zou een eventuele aanpassing of aanvulling van de Awb daarvoor nodig zijn en zo ja welke?

Eventuele wijzigingen zouden gericht moeten zijn op het creëren van een duidelijk juridisch kader dat zowel de implementatie van technologie faciliteert als de rechten en belangen van burgers beschermt. Hier zijn enkele specifieke aanbevelingen voor mogelijke aanpassingen:

1. Expliciete Richtlijnen voor Algoritmische Transparantie en Verantwoording

De Awb kan worden aangepast om expliciete richtlijnen en vereisten voor transparantie in het gebruik van algoritmen te omvatten. Dit zou kunnen betekenen dat overheidsinstanties verplicht worden om duidelijk te maken wanneer en hoe algoritmen worden gebruikt in besluitvormingsprocessen, inclusief de logica achter de algoritmische beslissingen.

Voorstel: Invoeren van een algoritmeregister waarin details over het doel, de werking en de impact van algoritmen worden gepubliceerd.

2. Uitbreiding van Rechten op Inzage en Bezwaar

De wet kan worden uitgebreid om burgers expliciete rechten te geven om inzage te krijgen in de algoritmische processen die invloed hebben op hun leven en om bezwaar te maken tegen beslissingen die op deze manier worden genomen.

Voorstel: Wettelijke bepalingen die het recht op een duidelijke uitleg van algoritmische besluiten en de mogelijkheid om een herziening aan te vragen indien men denkt dat het algoritme fouten heeft gemaakt of vooroordelig heeft gehandeld.

3. Normen voor het Testen en Valideren van Algoritmen

Er kunnen normen worden ontwikkeld voor het testen en valideren van algoritmen om te waarborgen dat ze eerlijk, nauwkeurig en vrij van discriminatoire bias zijn. Dit kan onderdeel zijn van een bredere strategie om de kwaliteit en veiligheid van technologische toepassingen te garanderen.

Voorstel: Verplichte impactbeoordelingen en onafhankelijke audits van algoritmen voordat ze in gebruik worden genomen, vooral als deze impact hebben op belangrijke beslissingen over burgers.

4. Wettelijke Basis voor Menselijke Interventie

Om de menselijke maat te waarborgen, kan de wet expliciet vereisen dat er altijd ruimte

is voor menselijke interventie in besluitvormingsprocessen, vooral in complexe, gevoelige of significante zaken.

Voorstel: Wettelijke vereisten dat besluiten die significant impact hebben op individuen door een mens moeten worden beoordeeld voordat ze definitief worden gemaakt.

5. Ethische Richtlijnen voor het Gebruik van AI en Algoritmen

De Awb kan worden aangevuld met een set ethische richtlijnen of principes voor het gebruik van kunstmatige intelligentie en algoritmen door de overheid. Dit zou helpen om consistentie en verantwoordelijkheid in de gehele publieke sector te waarborgen.

Voorstel: Integratie van een ethisch kader dat richtlijnen biedt over privacy, rechtvaardigheid, non-discriminatie, en transparantie in het gebruik van algoritmen.

6. Gebruik van Synthetische Data

Om privacy te beschermen en tegelijkertijd de ontwikkeling van nuttige algoritmen mogelijk te maken, kan de overheid investeren in technologieën voor het genereren van synthetische data die de kenmerken van echte data repliceren zonder persoonlijke informatie te bevatten.

Voorstel: Ontwikkeling en gebruik van geavanceerde technieken om synthetische datasets te creëren die kunnen worden gebruikt voor het trainen en testen van algoritmen zonder risico op privacy schending.

7. Dynamic Consent Frameworks

In plaats van eenmalige toestemming voor het gebruik van persoonsgegevens, kunnen dynamische toestemmingsframeworks worden ontwikkeld die gebruikers meer controle geven over hun gegevens. Dit kan vooral belangrijk zijn in contexten waar algoritmen leren en zich aanpassen over tijd.

Voorstel: Een systeem waarbij burgers hun toestemming voor het gebruik van hun gegevens kunnen aanpassen afhankelijk van de context, en waarbij duidelijke informatie wordt gegeven over hoe deze gegevens worden gebruikt.

8. Real-time Monitoring en Feedback Systemen

Ontwikkel systemen die real-time monitoring van algoritme-gestuurde besluitvorming mogelijk maken, compleet met feedback mechanismen die burgers direct betrekken. Dit zou kunnen helpen bij het identificeren en corrigeren van fouten of biases voordat ze significante impact hebben.

Voorstel: Implementatie van een digitaal dashboard dat door burgers kan worden geraadpleegd om te zien hoe besluiten die hen beïnvloeden worden genomen, en om feedback te geven op deze processen.

REFERENCES

- AI Now Institute: <https://ainowinstitute.org/>

Een onderzoeksinstituut aan de New York University dat zich richt op de sociale implicaties van kunstmatige intelligentie en technologie. De site biedt toegang tot onderzoeksrapporten, beleidsaanbevelingen en een breed scala aan bronnen over de impact van AI op de samenleving.

- AlgorithmWatch: <https://algorithmwatch.org/en/>

Een non-profit organisatie die analyseert en rapporteert over de gevolgen van algoritmische besluitvormingsprocessen. Ze bieden nieuwsupdates, rapporten en tools om de effecten van algoritmen te begrijpen en te adresseren.

- Data Ethics EU: <https://dataethics.eu/>

Een Europese denktank die zich concentreert op data-ethiek en het verantwoord gebruik van data en AI. De website biedt inzichten, beleidsvoorstellen, en ethische richtlijnen voor het gebruik van data.

- Centre for Data Ethics and Innovation (CDEI):
<https://www.gov.uk/government/organisations/centre-for-data-ethics-and-innovation>

Een Britse overheidsinstelling die zich bezighoudt met het bevorderen van het ethische gebruik van data en AI. De CDEI publiceert onderzoeken, beleidsadviezen en best practices voor het omgaan met nieuwe data-gestuurde technologieën.

- European AI Alliance:

<https://digital-strategy.ec.europa.eu/en/policies/european-ai-alliance>

Een platform opgezet door de Europese Commissie dat gericht is op het bevorderen van discussies en het delen van informatie over kunstmatige intelligentie binnen Europa. De leden van het platform kunnen participeren in een breed scala aan discussies en bijdragen aan de ontwikkeling van AI-beleid in Europa.