



Aanmeldformulier beheer kennisproduct (27-6-2024)

Titel

Duurzame toegankelijkheid van algoritmes en AI

Aanleiding

Steeds meer overheden zien mogelijkheden in het gebruik van algoritmes bij het uitoefenen van hun taken. Daarbij zijn ook steeds geavanceerdere zelflerende modellen (AI) in opkomst. De hoop en verwachting is dat met het toepassen van deze technologie taken efficiënter kunnen worden uitgevoerd. Tegelijkertijd bestaan er ook zorgen over de maatschappelijke impact van het gebruik van algoritmes. Het risico bestaat dat de transparantie van het overheidshandelen in het geding komt als er niet goed verantwoording kan worden afgelegd over de werking van algoritmes. Het archiveren ofwel duurzaam toegankelijk maken van een algoritme is een belangrijke maatregel om te borgen dat door een overheidsorganisatie verantwoording kan worden afgelegd over het gebruik daarvan.

In 2019 is onder de vlag van het KIA-platform Informatiehuishouding Overheden een [handreiking 'Algoritmes & accountability'](#) gemaakt. Deze handreiking biedt al goede handvatten aan organisaties om algoritmes duurzaam toegankelijk te maken. Op dit moment is het echter niet in beheer, het staat op een webpagina van één van de leden van de voormalige werkgroep. Sinds de publicatie hebben er geen grote wijzigingen plaatsgevonden. Een logische stap zou zijn om, in afstemming met de leden van de voormalige werkgroep, deze handreiking als basis te gaan gebruiken voor het kennisproduct.

Status

nieuw te ontwikkelen

Doelgroep

De doelgroep is:

1. Informatieprofessionals die betrokken zijn bij de ontwikkeling van algoritmes.

Te denken valt aan informatiebeheerders, security-/privacy-officers, informatieanalisten, informatiearchitecten, maar ook ontwikkelaars. De handreiking helpt hen om duurzame toegankelijkheid te borgen van algoritmes die worden ontwikkeld of al in gebruik zijn.

2. Eindverantwoordelijken (management)

Deze handreiking helpt informatieprofessionals en eindverantwoordelijken om tot passende maatregelen te komen. De informatieprofessionals kunnen vanuit hun expertise aan de eindverantwoordelijken uitleggen wat het belang is van duurzame toegankelijke algoritmes en inzicht geven in de mate waarin een algoritme daaraan voldoet. De eindverantwoordelijken kunnen op basis daarvan weloverwogen keuzes maken.

3. Professionals kwaliteitsmanagement zoals auditors

De handreiking helpt professionals kwaliteitsmanagement inzicht te krijgen in de gehanteerde methodiek voor het bepalen van duurzame toegankelijkheid van algoritmes en de mate van duurzame toegankelijkheid, zodat zij de passende maatregelen kunnen toetsen.

De handreiking is bedoeld voor alle overheidsorganisaties.

Doel

Het kennisproduct moet overheidsorganisaties helpen passende maatregelen te nemen voor het archiveren van een algoritme. Omdat algoritmes verschillende vormen kunnen aannemen, van beslisboom tot geavanceerde AI, moeten de maatregelen voor de archivering daarmee in overeenstemming zijn. Daarnaast speelt ook de eventuele impact, op de maatschappij of organisatie een rol. Het kennisproduct moet gebruikers helpen te bepalen wat er gearchiveerd moet worden, hoe lang en hoe dit moet gebeuren om te voldoen aan de zes kenmerken van DUTO.

Type kennisproduct

handreiking

Korte beschrijving

De scope van het kennisproduct betreft algoritmes die worden ontwikkeld of ingekocht door overheidsorganisaties. Afhankelijk van het doel waar het algoritme wordt gebruikt en de technologische complexiteit (bijvoorbeeld geprogrammeerde regels vs. AI), worden passende maatregelen voor archivering bepaald.

Het gebruik van een algoritme biedt veel kansen maar brengt ook risico's met zich mee. Zo zijn er zorgen over de wijze waarop overheidsorganisatie hun handelen kunnen verantwoorden. Een goede archivering van een algoritme is een belangrijke voorwaarde om overheidshandelen te reconstrueren waarbij gebruik is gemaakt van algoritmes. Overheidsorganisaties moeten dus al bij de ontwikkeling van een algoritme nadenken over het duurzaam toegankelijk maken hiervan. Het gaat daarbij niet alleen over de code of rekenregels van een algoritme, maar het kan bijvoorbeeld ook gaan over de trainingsdata en ontwerpdocumentatie. Als een algoritme getraind is met een gegevensset dan kan die trainingsdata essentieel zijn om het gedrag te reconstrueren. Ook ontwerpdocumentatie kan nodig zijn om inzicht te verkrijgen in de gemaakte keuzes die van invloed zijn op de werking van een algoritme.

Daarnaast is het ene algoritme het andere niet: een algoritme kan een set geprogrammeerde regels zijn maar ook zelflerend zijn. Verder kunnen maatschappelijke en organisatorische impact en daarmee gepaard gaande risico's uiteenlopen. De DUTO-maatregelen die zorgen voor beheer en toegankelijkheid moeten daarmee ook passend zijn: soms kan volstaan worden met een licht niveau van maatregelen, de andere keer is een zwaar niveau nodig. De handreiking moet informatieprofessionals bij overheidsorganisaties helpen om de impact van een algoritme te bepalen en DUTO-maatregelen te nemen die daarmee overeenstemmen.

De eerder genoemde handreiking 'Algoritmes & accountability' kan zoals gezegd een goede basis vormen. Er zullen wel aanvullende werkzaamheden nodig zoals het in lijn brengen met nieuwe (Europese) wetgeving, een link leggen met het implementatiekader 'Verantwoorde inzet van algoritmen' van het ministerie van Binnenlandse Zaken en het landelijk register voor algoritmes. Daarnaast moet de structuur aangepast worden aan de mogelijkheden van het door het NA gebruikte CMS. Daarbij worden waar mogelijk ook gebruikerservaringen van organisaties meegenomen die al gebruik hebben gemaakt van de 'KIA-handreiking'.

Outline

De handreiking zal modulair worden opgebouwd. Elke module is een afgebakend product dat apart opgeleverd en gepubliceerd zal worden.

Startpagina Introductie en scope

De startpagina bevat (de belangrijkste) informatie over het waarom van de handreiking, wat de handreiking beoogt en voor wie de handreiking is bedoeld. De startpagina wordt aan de hand van de modules gevuld.

Begrippen en uitgangspunten

Deze module verduidelijkt binnen het kader van het kennisproduct de definitie/reikwijdte van een algoritme en gehanteerde uitgangspunten.

Vragenlijst

De vragenlijst bestaat uit een aantal stappen met vragen die helpen inzicht te krijgen in een algoritme. Met de antwoorden op deze vragen wordt de gebruiker in staat gesteld keuzes te kunnen maken voor het op passende wijze archiveren van een algoritme. Het helpt daarbij ook om te bepalen wat er eventueel aan aanvullende informatie nodig is rondom het algoritme om deze duurzaam toegankelijk te maken, zoals ontwerpdocumentatie, trainingsdata etc...

De vragenlijst gaat vergezeld van een bijsluiter met nadere toelichting en achtergronden bij de vragenlijst.

Praktijkvoorbeelden

Door aantal voorbeelden te geven uit de praktijk van archivering van algoritmes bij overheidsorganisaties wordt meer handelingsperspectief verbonden aan de vragenlijst.

Relaties met andere kennisproducten

Op welke manier heeft dit kennisproduct relaties met de andere kennisproducten van het NA? Het kennisproduct heeft een relatie met:

- Het DUTO-raamwerk
- De handreiking risicobeoordeling duurzame toegankelijkheid
- De handreiking archiveren by design

Consultatie

Welke groepen zijn/worden geconsulteerd tijdens de ontwikkeling?

redactiegroep extern

Openbare review

geen openbare review