

Onderwerp: *Onderzoek Anonimiseren Zienswijzen*
Datum: 11 februari 2022
Versie: Definitief

Opdrachtgever: Ministerie van Infrastructuur en Waterstaat
Directie Participatie

Opdrachtnemer: SIS8020

Inleiding

De Directie Participatie van het ministerie van Infrastructuur en Waterstaat (IenW-DP) heeft aan SIS8020 verzocht om een onderzoek uit te voeren naar de mogelijkheden die Artificiële Intelligentie (AI) biedt voor het opsporen en lakken (anonimiseren) van persoonsgevoelige gegevens in zienswijzen. SIS8020 heeft daarvoor een onderzoekteam samengesteld. Het team heeft in de periode van 15 oktober 2021 tot 1 december 2021 het gevraagde onderzoek uitgevoerd en de behaalde resultaten op 29 november 2021 gepresenteerd. Aanvullend op het onderzoek zijn verdere mogelijkheden van de toepassing van AI voor zienswijzen in een sessie met een van de medewerkers van IenW-DP verkend.

Voorliggend rapport is de weerslag van de resultaten van het onderzoek en de aanvullende verkenning naar kansen van AI. Met dit rapport, de verzorgde presentatie op 29 november 2021 en de terug levering van de geanonimiseerde set onderzoeksdocumenten is het onderzoek voltooid en kunnen eventuele vervolgstappen gezet worden.

Aanpak en omvang onderzoek

Het vinden van persoonsgevoelige gegevens met AI kan met verschillende producten. Voor dit onderzoek is gebruik gemaakt van de Smart Privacy Analyzer (SPA) van de SIS8020-techpartner GlobeScope. De SPA wordt al in diverse werkomgevingen ingezet voor enkelvoudig en in bulk lakken van specifieke sets van documenten tot complete WOB-verzoeken. De SPA is goed configureerbaar op nieuwe documenttypen en snel te vullen met grote documentensets. Vandaar dat dit tool bij uitstek geschikt is om in te zetten voor het onderzoek.

Gegeven de onderzoeksdoelstelling, beschikbaar budget en beschikbare tijd, was het niet mogelijk om alle zienswijzen als onderzoeksobject te kiezen. Door IenW-DP is een set van 315 documenten met zienswijzen rondom het project A2 Deil – Vught aangeleverd. Deze set wordt door IenW-DP als representatief gezien voor 1) de wijze waarop de zienswijzen aangeleverd zijn (digitaal, papier, mondeling), 2) het bestandstype van de zienswijze (docx, pdf en jpg) en 3) qua omvang en verscheidenheid van de individuele zienswijzen (van korte stukjes tot uitgebreide documenten)

Voor het bereik van dit onderzoek zijn na afstemming met IenW-DP de volgende gegevens als privacygevoelig aangemerkt:

- Persoonsnaam
- Emailadres
- IBAN
- BSN
- Postcode
- BTW-nummer
- Telefoonnummer

Wel vindbaar door de AI van de SPA maar voor dit onderzoek niet geconfigureerd zijn:

- KVK-nummer (doorgaans niet privacygevoelig)
- Webadres (doorgaans niet privacygevoelig)



- Locatie¹ (doorgaans niet privacygevoelig, nadat de persoonsnaam gelakt is)
- Organisatie (doorgaans niet privacygevoelig)
- Ondertekenaar (is nog van onvoldoende kwaliteit)
- Geldbedrag (lijkt ons niet relevant in dit kader)

NB: de SPA is uitbreidbaar op het vinden van informatie-elementen. Dus als in een later stadium of in een andere context andere informatie-elementen gezocht en gevonden moeten worden, dan kan GlobeScope de AI onder de SPA daar aanvullend op trainen

De documentenset is beschikbaar gesteld via de samenwerkingsruimte van lenW-DP. Vanuit die samenwerkingsruimte zijn de documenten op de speciaal voor dit onderzoek ingerichte beveiligde SPA-server gezet. De documenten zijn eerst allemaal volledig geautomatiseerd en op de achtergrond geanalyseerd op aanwezigheid van persoonsgevoelige gegevens conform bovenstaande lijst.

Na de geautomatiseerde analyse zijn alle 315 documenten door leden van het onderzoeksteam handmatig nagelopen waarbij:

- persoonsgevoelige gegevens die de AI niet heeft gezien (false negatives) alsnog geannoteerd zijn;
- annotatie van gegevens die niet persoonsgevoelig zijn maar waar de AI dat wel dacht (false positives) is weggehaald;

Na correctie zijn alle gevonden persoonsgevoelige gegevens in de zienswijzen weggelakt en zijn de voor publicatie geschikte geanonimiseerde doorzoekbare PDF-bestanden volledig geautomatiseerd gegenereerd. lenW-DP beschikt nu over een set met 315 volledig geanonimiseerde zienswijzen.

De behaalde resultaten en de gevolgde werkwijze is besproken met het team van lenW-DP. Tijdens deze bespreking zijn diverse vragen met betrekking tot de behaalde resultaten toegelicht en is de gevolgde werkwijze van anonimiseren van zienswijzen met behulp van de SPA gedemonstreerd.

Onderzoeksresultaten

Het uitgevoerde onderzoek naar inzet van AI ten behoeve van anonimiseren van zienswijzen heeft geleid tot de volgende inzichten:

1. Bij het beschikbaar stellen van zienswijzen moet rekening gehouden worden met eventueel privacygevoelige informatie in de zienswijzen. Verwerking van informatie moet veilig gebeuren. De gevolgde optie voor dit onderzoek bestond uit het beschikbaar stellen van de zienswijze-bestanden in de samenwerkingsomgeving. Vervolgens zijn deze bestanden handmatig overgezet naar de beveiligde server van de SPA. Dit werkt voor sets in de omvang van ca 500 documenten. Bij grotere volumes of procesmatige verwerking is het verstandig om naar een slimmer uitwisselingsmechanisme te kijken;
2. De kwaliteit van de gescande documenten is sterk van invloed op de kwaliteit van de gevonden persoonsgevoelige informatie. Dit speelt bij bestanden van het type 'docx' en (afgedrukt naar) 'pdf' niet. Documenten die als gescande JPG of als 'plaatje'-pdf zijn aangeleverd bevatten

¹ Locatie in de vorm van bijvoorbeeld straatnaam met huisnummer wordt soms wel en soms niet als (privacy)gevoelige informatie gezien. In zichzelf zijn adressen openbare informatie, maar soms maakt de context dat een adres toch (privacy)gevoelige informatie is. Binnen dit onderzoek is ervoor gekozen om adressen/straatnaam/huisnummer niet te configureren als te anonimiseren informatie.



onzuiverheden waardoor bepaalde gegevens minder goed gevonden worden. Zo gaat soms de punt verloren die vereist is voor het herkennen van een mailadres;

3. Gemiddeld genomen heeft de AI in 81% van de gevallen de persoonsgegevens gevonden. In 19% van de gevallen was een correctie van de medewerker nodig;

Artificial Intelligence (AI)

In de voorbespreking van het onderzoek naar inzet van AI is gekozen voor het leggen van de focus op het vraagstuk rondom anonimiseren van zienswijzen. Andere, bredere en diepere kansen voor de inzet van AI in het domein van zienswijzen werd vooralsnog aangehouden voor een eventueel vervolgonderzoek. Om een vervolg meer handen en voeten te geven en toch de aangewakkerde nieuwsgierigheid naar (on)mogelijkheden van AI te adresseren, heeft het onderzoeksteam in een in samenwerking met lenW-DP, een verkenning op hoofdlijnen uitgevoerd. Op basis daarvan komen we tot de volgende inzichten en aanbevelingen.

Kansen

In onderstaande tabel zijn op hoofdlijnen de kansen benoemd die binnen bereik komen door de directe of indirecte inzet van AI en aanverwante technieken. In de meest rechter kolommen hebben we de indicatie gegeven van de verwachte complexiteit, kosten en tijd die gepaard gaan met het realiseren van die kans. Wat we niet hebben kunnen doen is de indicatie van de mate van toegevoegde waarde die de kans voor lenW-DP en haar stakeholders heeft. Dat is ons inziens noodzakelijk om een verantwoorde keuze te maken uit de ontwikkelingen.

Onderwerp	Toelichting	Complexiteit	Kosten	Tijd
<i>Vergelijkbare zienswijzen ontdekken</i>	lenW-DP ontvangt meerdere zienswijzen voor projecten. Inzicht in overeenkomsten binnen inzichten is vanuit meerdere invalshoeken opportuun: • (op onderdelen) sterk vergelijkbare zienswijzen geven inzicht in de punten waar aandacht voor of onrust over bestaat • Sterk overeenkomende zienswijzen kunnen duiden op 'georganiseerd indienen', niet verboden maar mogelijk wel relevant voor behandeling; Terzijde, het voor de zekerheid indienen van dezelfde zienswijzen via meer kanalen lijkt procedureel voldoende afgevangen te worden	M	M	M
<i>Slimmer/beter classificeren</i>	Zienswijzen worden geclassificeerd in één of meer vooraf gedefinieerde klassen. AI kan bijdragen aan het volledig geautomatiseerd indelen in klassen op basis van een bredere context dan sets van enkelvoudige	L	M	L

Onderwerp	Toelichting	Complexiteit	Kosten	Tijd
	woorden. Dat zou de betrouwbaarheid ten goede komen.			
<i>Onderliggende thema's/ vraagstukken/ probleemgebieden ontdekken</i>	AI kan in combinatie met Data Science (DS)-technieken indicaties afgegeven voor thema's, centrale vraagstukken of gerichte probleemgebieden die uit de ingediende zienswijzen te destilleren zijn.	H	M	M
<i>Consistentie in benadering/behandeling borgen (binnen projecten en over projecten heen)</i>	Als zienswijzen betrekking hebben op een vergelijkbaar onderwerp dan zou behandeling van die zienswijzen ook vergelijkbaar kunnen zijn. Zo is het goed denkbaar dat alle zienswijzen over bijvoorbeeld geluidshinder binnen één project, op een vergelijkbare manier ook behandeld worden. Het is zelfs denkbaar dat ook project overstijgend naar geluidshinder gekeken kan worden. Zowel wat betreft de zienswijzen als de behandeling ervan.	L	M	L
<i>Overstijgende thema's ontdekken (over projecten heen)</i>	Net zoals AI en DS ingezet kan worden om thema's binnen projecten te ontdekken, is het ook mogelijk om project overstijgende thema's te ontdekken	H	M	M
<i>Sentiment/Mood analyse</i>	Zienswijzen verzamelen inbreng van buiten over een bepaald project dat door de overheid gaat worden uitgevoerd. Belanghebbenden kunnen en zullen vandaag de dag hun mening ook op andere platformen kenbaar maken, waaronder de social media platformen. AI kan ingezet worden om de verbinding te zoeken tussen thema's/ vraagstukken / probleemgebieden die uit de zienswijzen blijken (project specifiek én project overstijgend) en hetgeen daarover op social media gedeeld wordt. Groepsbeeldvorming en -beïnvloeding kan zo eerder in zicht komen hetgeen inspelen op sentimenten rondom een thema/ vraagstuk / probleemgebied ondersteunt.	H	M	H
<i>Timelining/Geoplotting</i>	Niet zozeer gedreven door AI maar wel mogelijk als extra inzicht, is het in beeld brengen van de thema's/vraagstukken/probleemgebieden in het verloop van de tijd en/of op locatie. Door tijd/locatie als focuspunt te kiezen kunnen ontwikkelingen in de tijd/in een gebied inzichtelijk gemaakt worden.	L	L	M



Zienswijze als fundament

De in de voorgaande tabel genoemde kansen leunen op verschillende onderliggende technieken. AI is er daar een van, maar ook alle aan DS verwante technieken zijn in meer of mindere mate toepasbaar. Maar het verzilveren van de beschreven kansen begint bij het opbouwen van een fundament dat uit twee onderdelen bestaat:

A. Het herkennen van een zienswijze

Het herkennen van de zienswijze lijkt op het eerste gezicht eenvoudig maar zal dat in de praktijk niet blijken te zijn. Immers zienswijzen kunnen door iedereen op eigen wijze worden geformuleerd. De één begint met een situatieschets, de ander met een grief, een derde begint eerst met hoe het anders zou moeten en een volgende begint een juridisch betoog waarom het toch echt niet kan dat... En waar dat voor mensen soms al moeilijk is om duidelijk begin en eind van een betoog als zienswijze te duiden, is dat voor AI helemaal lastig. Een AI-oplossing die individuele zienswijzen kan herkennen vraagt specialistische en grondige training. Gelukkig zijn documenten met daarin de zienswijzen in ruime mate voorhanden. Om een AI-model te trainen op het herkennen van zienswijzen, moeten deze expliciet geannoteerd worden in een trainingsset van documenten. Hoe groter de trainingsset, desto betrouwbaarder het AI-model wordt. SIS8020 en GlobeScope hebben de expertise en capaciteit om een dergelijk model te trainen. Toch is een 'word of warning' op zijn plaats. Het ontwikkelen van AI-modellen is niet te vergelijken met het ontwikkelen van functionaliteiten in applicaties. Van een functionaliteit kan een specificatie geschreven worden en kan objectief vastgesteld worden dat hetgeen gevraagd is ook geleverd wordt. Met AI ligt dat anders; het is (nog) niet voorspelbaar hoeveel training nodig is om een model te ontwikkelen van voldoende kwaliteit. Daarom ontwikkelen wij modellen altijd in tranches. Iedere tranche levert een werkend model op basis van bepaalde trainingsdata en trainingstechnieken. En een volgende tranche wordt pas opgepakt als a) de verwachte verbetering opweegt tegen de te maken kosten, b) er voldoende trainingsmateriaal beschikbaar is en c) de gevolgde techniek bewezen heeft de kwaliteit van het model te verbeteren.

B. Het vinden van verbanden tussen zienswijzen

Als dan de zienswijzen gevonden zijn, is de volgende cruciale stap om de verbinding tussen de zienswijzen te vinden. Dat kan langs diverse lijnen:

- overeenkomst op basis van inhoud: 100% overeenkomst duidt op doublure, 90% overeenkomst duidt op gelijksoortigheid;
- overeenkomst op basis van kernbegrippen: van iedere zienswijze wordt met behulp van AI de set van kernbegrippen afgeleid (NB dus niet vooraf gedefinieerd!). Zienswijzen die meer kernbegrippen delen, delen meer inhoud dan dat ze met andere zienswijzen doen. Door gebruik te maken van bekende AI- en DS-componenten, vraagt dit weliswaar veel techniek maar blijft zowel tijd als complexiteit beperkt;
- overeenkomst op basis van classificatie: van iedere zienswijze wordt met AI vastgesteld tot welke vooraf bepaalde klasse(n) deze hoort. Door gebruik te maken van bestaande data, is dit een relatief eenvoudige training;

Met het vinden van de zienswijzen (A) en het leggen van verbanden tussen zienswijzen (B) is de weg vrij gemaakt voor vervolg stappen zoals het in beeld brengen van samenhang tussen zienswijzen vanuit verschillende invalshoeken, binnen projecten en over projecten heen, nieuwe thema's te ontdekken en relevante van bestaande bekende thema's inzichtelijk te maken.



Conclusies en aanbevelingen.

De conclusie uit het uitgevoerde onderzoek naar **anonymiseren van zienswijzen** laat zich als volgt samenvatten:

1. De binnen IenW-DP beschikbare documenten met zienswijzen zijn goed te analyseren op aanwezigheid van privacygevoelige gegevens;
2. Uit de uitgevoerde proef blijkt dat er winst te behalen is door het bijtrainen van het AI-model op a) het analyseren van documenten van het type zienswijze en b) het onderscheiden van bedrijfsnaam versus achternaam;
3. Met het huidige model is ca 80% van privacygevoelige informatie gevonden door AI en 20% door de mens. Met extra AI-modeltraining zoals genoemd onder 2 wordt een verdere verbetering van het model verwacht, waardoor de benodigde inzet van medewerkers voor controle en nabewerking verder omlaag gebracht kan worden;

Onze aanbeveling is dan ook:

- I. Laat het AI-model onder de SPA een gerichte trainingsronde op persoonsgegevens binnen zienswijzen ondergaan, met name gericht op het verminderen van de benodigde (controle)inspanning van de medewerker;*
- II. Gebruik een betrouwbaarheidsindicator/drempel voor de keuze of een medewerker de door AI uitgevoerde anonimisering moet controleren of verbeteren;*
- III. Anonimiseer alle zienswijzen die nog gepubliceerd moeten worden;*
- IV. Richt een proces in dat zienswijzen vroeg na ontvangst geautomatiseerd anonimiseert. (Denk ook aan komende Wet Open Overheid (WOO);*

De conclusie met betrekking tot **verdere inzet van AI ten behoeve van zienswijzen** laat zich als volgt samenvatten:

1. Er zijn veel kansen om een beter, breder, dieper inzicht te destilleren uit de zienswijzen
2. De toegevoegde waarde van die kansen ligt niet zozeer in het administratief verwerken van de zienswijzen maar meer bij degene die inzicht in de ingediende zienswijzen nodig hebben;
3. Het verzilveren van die kansen begint bij twee onderdelen a) het herkennen van de individuele zienswijze en b) het leggen van het verband tussen zienswijzen;
4. Het trainen van AI om zienswijzen en hun onderlinge verbanden te herkennen is specialistisch werk, moet in tranches gebeuren, en kan bestaande zienswijze informatie goed gebruiken;
5. Werken met AI-modellen schept vaak te hoge verwachtingen over succes. We kunnen op voorhand niet voorspellen welke betrouwbaarheid in welke tranche gehaald wordt. Dat betekent dat bij het (verder) trainen van AI-modellen de kosten-opbrengsten-balans goed in de gaten moet worden gehouden;

Onze aanbeveling is dan ook:

- I. Onderzoek bij de stakeholders welke kansen op het inzetten van AI bij zienswijzen het meeste waarde toevoegen;*
- II. Start met een eerste tranche training op a) het herkennen van zienswijzen. Voor een eerste indicatie van haalbaarheid zou een set van ca 3.000 documenten voldoende zijn;*
- III. Kies na II of een volgende stap gericht moet zijn op het verbeteren van het herkennen van zienswijzen of juist gericht moet zijn op het herkennen van verbanden tussen zienswijzen;*