

ACHTERGROND

Tekst **Klaas Jan Mollema, Alexander Aarup, Tim van den Berg, Bart van Biezen, Max van Dongen & Martijn van der Weerd**

Klaas Jan Mollema is docent aan de opleiding HBO-ICT en onderzoeker aan het lectoraat Civic Technology van De Haagse Hogeschool. Alexander Aarup, Tim van den Berg, Bart van Biezen, Max van Dongen en Martijn van der Weerd zijn UXD en HBO-ICT studenten aan De Haagse Hogeschool en volgden de minor Innovative Data Visualisation

Slimmer zoeken in Woo-portalen

Hoe studenten van De Haagse Hogeschool werken aan een gebruiksgerichte toegang tot overheidsinformatie

Woo-portalen bevatten vaak duizenden documenten, maar bieden weinig houvast. Vooral journalisten en onderzoekers verdwalen in de eindeloze hoeveelheid zoekresultaten. De Haagse Hogeschool ging samen met rijksambtenaren en studenten op zoek naar een beter ontwerp: een Woo-portaal dat niet alleen zoekt, maar ook slim helpt vinden.

Waar de Wet openbaarheid van bestuur (Wob) aan overheden vraagt om documenten op verzoek te delen, kent de Wet open overheid (Woo) een actievere openbaarmaking:

overheden moeten, behalve reageren op Woo-verzoeken, ook proactief documenten openbaar delen. Deze documenten delen ze via Woo-portalen: websites waarop burgers kunnen zoeken en documenten kunnen bekijken. Deze portalen kenmerken zich vaak door een zoekbalk die toegang biedt tot een oneindige pool aan documenten. Zodra het zoekwoord is ingegeven, krijgen nietsvermoedende burgers een resultatenlijst van vele duizenden documenten (bijvoorbeeld Officiële Bekendmakingen, zoals in de afbeelding) en worden ze aan hun lot overgelaten.

Mensen gebruiken Woo-portalen vooral in de context van onderzoek. Journalisten en onderzoekers controleren en onderzoeken de overheid en gebruiken de openbaar gemaakte documenten als voeding voor

Zoekresultaten 1 - 10 van de 7621 resultaten



hun werk. Zij ervaren deze portalen als onhandige en ondoorgroondelijke vehikels.

Toegankelijker

Het idee voor een toegankelijker Woo-portaal ontstond bij Brechje Schouwerwou, Ines Prakke (BZK Transparant) en Pepijn van Houdt (OCW – RDDI).

'JOURNALISTEN EN ONDERZOEKERS ERVAREN WOO-PORTALEN ALS ONHANDIG EN ONDOORGRONDELIJK'

Geïnspireerd door het aanbevelingssysteem van Spotify vroegen zij zich af: hoe kun je gebruikers in een Woo-documentenbank helpen om snel relevante informatie te vinden, op basis van slimme suggesties? Met deze vraag stapten ze naar I-Lab West (I-Partnerschap) om te verkennen of samenwerking met het onderwijs mogelijk was. De Haagse Hogeschool besloot de handschoen op te pakken en betrok een groep studenten van de opleiding Innovative Data Visualisation bij het project. Samen met de rijksambtenaren en de docent formuleerden de studenten de centrale ontwerpopdracht:

- Er is behoefte aan gecentraliseerde, proactieve toegang tot informatie. Nu zijn Woo-documenten verspreid over meerdere websites. De diverse soorten informatie afkomstig van diverse organisaties toegankelijk maken op een gecentraliseerde plek, zodat de burger makkelijker toegang heeft tot een compleet overzicht.
- Het creëren van een interface die toegankelijk is voor een diverse doelgroep: journalisten, studenten, onderzoekers, burgers en ambtenaren. In de interface op basis van persoonlijke interesse voorzien in aanbevelingen.
- Het visuele ontwerp moet zowel creatief en innovatief zijn, als vallen in de professionele communicatiestijl van de rijksoverheid.

Op zoek naar gebruikers

De groep studenten interviewde journalisten om inzicht te krijgen in wat goed werkt en welke tekortkomingen er zijn in de huidige overheidsplatformen. De journalisten gaven aan dat de huidige overheidswebsites moeilijk navigeerbaar zijn en dat de zoekopties/filters beperkt zijn. Pdf-bestanden zijn vaak moeilijk doorzoekbaar en metadata of context in de vorm van tijdlijnen ontbreken vaak. Bestaande platformen kennen geen notificatiefunctie waardoor proactieve attendering nog altijd ontbreekt.

De journalisten zien kansen voor de inzet van generatieve kunstmatige intelligentie (AI) om documenten massaal toegankelijker te maken. Wel moet dat voorzichtig gebeuren om de kwaliteit van de informatie te blijven garanderen. Herschrijven van documenten doormiddel van AI moet voorkomen worden: de brontekst blijft altijd belangrijk om de integriteit van de informatie te waarborgen.

Op basis van het onderzoek is de Smart Document Finder ontwikkeld als een *proof-of-concept* voor een toegankelijker Woo-portaal. Dit prototype demonstreert hoe een verbeterd *user experience design* en de inzet van AI kunnen bijdragen aan een betere toegankelijkheid. De applicatie is ontworpen om overheidsinformatie transparanter doorzoekbaar te maken voor diverse gebruikers, zoals journalisten, onderzoekers en burgers. De webgebaseerde oplossing biedt een intuïtieve interface die eenvoudige navigatie en verkenning van openbare gegevens mogelijk maakt.

Het platform is gestructureerd rond twee primaire gebruikerspaden: het Zoekpad en het Bladerpad.

- Zoekpad: gebruikers kunnen specifieke zoekopdrachten invoeren (bijvoorbeeld gaswinning Groningen) en krijgen een lijst met relevante documenten. De resultaten worden verfijnd met filters, zoals datum, documenttype en relevantie. Elk document wordt weergegeven met een samenvatting en gebruikers kunnen ervoor kiezen om het volledige document of de samenvatting te bekijken. Deze functie ondersteunt efficiënte, gerichte zoekopdrachten voor gebruikers die op zoek zijn naar specifieke informatie.



- Bladerpad: voor gebruikers die liever informatie verkennen dan direct te zoeken, biedt het Bladerpad een interactief woordweb. Onderwerpen als huisvesting kunnen worden aangeklikt om subonderwerpen te verkennen (bijvoorbeeld verhuur of subsidies). Deze functie stelt gebruikers in staat dieper in specifieke thema's te duiken door een duidelijke, intuïtieve visuele hiërarchie te volgen. Subonderwerpen zijn gekoppeld aan relevante artikelen, waardoor het gemakkelijker wordt om gerelateerde informatie te ontdekken.

'DE BRONTEKST BLIJFT ALTIJD BELANGRIJK OM DE INTEGRITEIT VAN DE INFORMATIE TE WAARBORGEN'



Alle documenten die door de Smart Document Finder worden aangeboden, zijn vooraf ingelezen door generatieve kunstmatige intelligentie (GenAI). Deze technologie vormt de basis om de verschillende functionaliteiten aan te kunnen bieden:

- Het woordweb wordt automatisch gevormd op basis van door GenAI gegenereerde trefwoorden.
- De GenAI-trefwoorden bieden de mogelijkheid ook op inhoudelijke metadata te kunnen filteren.
- Door GenAI bepaalde trefwoorden bij documenten worden voorzien van gerelateerde trefwoorden die bepaald zijn door GenAI.
- Documenten worden integraal overgenomen, maar ook voorzien van een GenAI-samenvatting. Er kan naadloos geschift worden tussen de originele tekst en de samenvatting.
- Er kan een samenvatting gemaakt worden door GenAI die gepersonaliseerd is en voor de gebruiker van toepassing.

Om de beveiliging van het Smart Document Finder-platform te versterken, is een lokaal gehost taalmodel (LLM) geïmplementeerd. Dit voorkomt dat gevoelige gegevens het systeem verlaten, minimaliseert datalekrisico's en waarborgt naleving van gegevensbeschermingsregels. Het LLM verwerkt gebruikersvragen veilig en detecteert dreigingen zoals kwaadwillende invoer of onbevoegde toegang. Dankzij contextueel begrip filtert het efficiënt gegevens, waardoor blootstelling van gevoelige informatie wordt beperkt. Preliminair evaluaties geven aan dat het lokale LLM effectief is in het balanceren van geavanceerde natuurlijke taalverwerkingsmogelijkheden met een hoog niveau van systeembeveiliging.

Gebruikerstest

De Smart Document Finder (SDV, 2025) werd grondig getest met meerdere gebruikersgroepen om de bruikbaarheid van het platform te beoordelen en feedback te verzamelen zodat de functies verder verfijnd konden worden. Wat blijkt?

- Journalisten waardeerden het zoekgemak, maar vroegen om extra filters per documenttype (bijvoorbeeld rapporten, brieven) en een recent bekeken functie voor snellere toegang.
- Onderzoekers waren enthousiast over het Bladerpad, dat complexe thema's als huisvesting en verhuur inzichtelijk maakte. Ze stelden voor academische bronnen en externe referenties toe te voegen.
- Burgers vonden de interface intuïtief, maar vroegen om duidelijkere categorisatie van politieke thema's en eenvoudigere toegang tot documenten over actuele kwesties. Op basis van deze feedback werden verschillende verbeteringen doorgevoerd, waaronder het verfijnen van de structuur van het woordweb om het intuïtiever te maken en het toevoegen van meer geavanceerde zoekfilters.

Doorontwikkeling

De Smart Document Finder is een *proof-of-concept* dat de toegankelijkheid van overheidsdocumenten verbetert. Ons onderzoek toont aan dat een gebruikersgericht aanpak bijdraagt aan betere toegankelijkheid van overheidsdocumenten. De toepassing van generatieve AI voor automatische contextualisatie biedt hierbij een duidelijke meerwaarde. Ook het gebruik van visuele navigatie via trefwoorden vergroot de exploreerbaarheid van de uitgebreide documentencollectie.

Momenteel kan de Smart Document Finder tekstuele documenten koppelen en afstemmen op gebruikersinteresses via een eenvoudig profiel. De volgende stap is ondersteuning van diverse documenttypen, waardoor een rijkere context ontstaat. Integratie van tijdlijnen kan inzicht bieden in politieke processen rondom documenten.

Verder kunnen zoekwoordvoorstellen en een gepersonaliseerd aanbevelingssysteem de gebruikerservaring versterken. Schaalvergroting, met documenten van uiteenlopende overheidsinstellingen, zal het taalmodel verbeteren en de effectiviteit van het platform vergroten. Dit draagt bij aan een completer en gebruiksvriendelijker systeem voor publieke informatievoorziening.

Meer informatie:

- GitHub (2025) DocGov, <https://github.com/Max-van-Dongen/docgov>.
- Officiële Bekendmakingen (2025) Zoek in alle documenten, <https://www.officielebekendmakingen.nl/>.
- SDV (2025) Smart Document Finder, <https://docgov.prsonal.nl/>.
- Spotify (2025) Understanding recommendations on Spotify, <https://www.spotify.com/us/safetyandprivacy/understanding-recommendations>.