
DIGITALE TOOLS EN BEGRIJPELIJKE BRIEVEN

EINDRAPPORT
GEBRUIKER CENTRAAL

Korte titel	Digitale tools en begrijpelijke brieven
Opleverdatum	31 – 03 – 2026
Opdrachtgever	Ministerie van BZK – Afdeling Dienstverlening
Opdrachtnemer	ICTU – Gebruiker Centraal
Versie	V1

Inhoud

Inhoud.....	2
1. Inleiding.....	3
2. Hoe hebben we gewerkt?	3
3. Welke digitale tools voor begrijpelijke brieven zijn er?	7
4. Waar moeten de tools aan voldoen?.....	9
5. Welke behoeftes hebben gebruikers voor deze tools?	12
6. Hoe test je tools met gebruikers?.....	14
7. Hoe test je de teksten met eindgebruikers?	14
8. Hoe kies je de juiste tool?	15
9. Wat zouden we aanbevelen voor het vervolg?	16
Bronnen	17
Bijlagen	18
Bijlage 1 – Lijst met tools (alfabetische volgorde).....	18
Bijlage 2 – Toetsingskader.....	21
Bijlage 3 – Testen van het model of algoritme	26
Bijlage 4 – Testplan	28
Bijlage 5 – Stappenplan Test je tekst.....	30
Bijlage 6 – Gesprekspartners	32
Bijlage 7 - Over de taal in dit rapport	33

1. Inleiding

Steeds meer van het dagelijks leven speelt zich digitaal af. Ook steeds meer diensten van de overheid zijn zowel digitaal als fysiek beschikbaar. Het ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK) wil dat eindgebruikers al deze diensten makkelijker kunnen bereiken en gebruiken.

Een belangrijk onderdeel is duidelijke communicatie. Niemand houdt van ingewikkelde brieven en informatie. Daarom wil BZK publieke dienstverleners helpen bij het schrijven van begrijpelijke brieven. Dit kan door gebruik te maken van nieuwe technologieën, zoals AI.

Momenteel zijn er veel digitale hulpmiddelen op de markt die helpen met begrijpelijk schrijven. En er komen er steeds meer. Maar welke zijn er allemaal? Waar moeten die digitale tools aan voldoen? En wat hebben publieke dienstverleners eigenlijk nodig? De antwoorden op deze vragen vind je in dit rapport.

Proces

Het eerste wat we hebben gedaan is een overzicht maken van zo veel mogelijk digitale tools op het gebied van begrijpelijk schrijven. Daarnaast hebben we aan gebruikers gevraagd waar de tools in ieder geval aan moeten voldoen. Op basis daarvan hebben we een toetsingskader opgesteld. Hierin staan criteria waar een tool aan moet voldoen. Bijvoorbeeld of het voldoet aan wet- en regelgeving. En of de tool voldoet aan ethische afwegingen. Op basis van dit kader kunnen ook toekomstige tools worden beoordeeld.

Maar een overzicht en toetsingskader zeggen niets over de kwaliteit van de tools. Wil je weten of de output begrijpelijk is, dan moeten de tools eerst getest gaan worden. Op basis van het toetsingskader in dit rapport kan een eerste selectie van tools gemaakt worden. Deze tools kunnen dan getest worden op basis van het plan dat ook in dit rapport staat.

Op deze manier kan het ministerie in de toekomst een selectie verspreiden van tools die aan alle eisen voldoen. En die begrijpelijke teksten leveren. Zodat iedere bewoner weet wat er van hem/haar verwacht wordt bij het lezen van de brief.

NB: Omdat dit rapport gaat over digitale hulpmiddelen en begrijpelijk schrijven, leek het ons logisch om deze hulpmiddelen te gebruiken voor het herschrijven van dit rapport. Zie voor meer informatie [Bijlage 7 - Over de taal in dit rapport](#).

2. Hoe hebben we gewerkt?

Hier lees je hoe we deze opdracht hebben uitgevoerd. We beschrijven:

- welke tools we hebben gevonden

- wat de kerngroep deed en wie erin zat
- met welke experts we gesproken hebben
- wat er uit de inputsessies kwam
- hoe we het toetsingskader ontwikkeld hebben
- waarom we een deel van het onderzoek anders hebben aangepast.

Welke tools hebben we gevonden?

We hebben op verschillende manieren gezocht naar tools die ambtenaren helpen bij het schrijven van begrijpelijke teksten:

- Eerst hebben we een lijst gemaakt van tools die we al kenden. Zo heeft Gebruiker Centraal in maart van 2025 al eerder een sessie georganiseerd over dit type tools.
- Ook vanuit de community's van Gebruiker Centraal waren enkele tools al bekend. Gedurende het onderzoek hebben leden van de community's ons geattendeerd op andere tools.
- Deze lijst hebben we aangevuld met tools die het ministerie al in beeld had, onder andere via het onderzoek van Dialogic.
- Leden van de kerngroep en deelnemers aan de inputsessies hebben ons ook gewezen op tools.
- Daarna zijn we online gaan zoeken. We wilden weten welke andere tools er nog zijn. Tot eind maart 2026 hebben we naar tools gezocht. Elke tool die we hebben gevonden, hebben we in het overzicht gezet.

Let op: Het gebied van digitale tools ontwikkelt zich heel snel. Elke dag komen er nieuwe tools bij. Daarom is het onmogelijk om een volledig overzicht te geven van tools die er zijn. Door het toetsingskader dat in dit rapport wordt beschreven, kunnen ook tools die niet in het overzicht staan gewogen worden.

Wat deed de kerngroep?

We hebben een kerngroep samengesteld van communicatieprofessionals die met ons willen sparren over dit onderzoek. Voor deze kerngroep hebben een oproep gedaan aan verschillende community's. Daaruit kozen we 9 experts op het gebied van publieke diensten, begrijpelijkheid en digitale tools. Bij welke organisaties deze experts werken, lees je in Bijlage 6 – Gesprekspartners.

Wij kozen deze experts voor de kerngroep omdat ze:

- werken in de publieke sector
- ervaring hebben met communicatie naar eindgebruikers
- kennis en ervaring hebben met begrijpelijke taal.

Zij vormden samen de kerngroep. De kerngroep hielp ons met advies en zette ons aan tot actie. Zo konden we de kennis en ervaring van de leden van deze kerngroep gebruiken. Ook maakten we gebruik van hun netwerk.

We hebben bewust geen ontwikkelaars of leveranciers van tools in de kerngroep gezet. Zo bleven we onpartijdig.

In totaal hebben we 4 bijeenkomsten met de kerngroep gehouden: op 13 november, 18 december, 15 januari en 18 maart.

Bij belangrijke beslissingen hebben we de leden tussendoor op de hoogte gehouden. Bijvoorbeeld over de scopewijziging van het onderzoek. Dit leverde enkele persoonlijke gesprekken op met sommige leden.

Met welke experts hebben we gesproken?

Voor dit onderzoek hebben we met diverse experts gesproken. Bijvoorbeeld met experts op het gebied van:

- begrijpelijke communicatie en tools (wetenschappers en communicatieprofessionals)
- software-ontwikkeling
- testen
- AI
- aanbestedingen.

Uiteindelijk hebben we met experts van 7 verschillende organisaties gesproken. Welke organisaties dat zijn, lees je in Bijlage 6 – Gesprekspartners.

Wat hielden de inputsessies in?

We wilden graag weten wat de wensen en behoeften waren van de gebruikers van de tools voor begrijpelijke teksten. Om hierachter te komen, hebben we 3 online inputsessies georganiseerd: op 13, 15 en 21 januari.

Communicatieprofessionals die voor de overheid werken, werden via LinkedIn en de site van Gebruiker Centraal geattendeerd op deze inputsessies. Geïnteresseerden konden zich online opgeven. Uiteindelijk hebben 21 mensen dat gedaan. Bij welke organisaties zij werken, lees je in Bijlage 6 – Gesprekspartners.

We legden hen verschillende keuzes voor en gingen dieper in op een aantal onderwerpen.

Iedere inputsessie bestond uit 3 delen:

- Introductie met algemene vraag: gebruik je al een tool, zo ja welke is dat en vind je die tool prettig?
- Discussie op basis van 3 polls:
 - Poll 1: Wat heeft meer waarde: een tool die uitlegt hoe de tekst beter kan, of een tool die direct een herschreven tekst oplevert?
 - Poll 2: Wanneer moet de tool helpen?
 - Poll 3: Hoe diep moet de tool kunnen kijken?
- Verdiepende gesprekken in break-outs.

De uitkomsten van deze inputsessie vind je in [hoofdstuk 5](#) en geven richting aan het testen met gebruikers.

Hoe kwam het toetsingskader tot stand?

Kiezen uit 53 tools voor publieke dienstverlening is lastig. Welke tool past wel en welke niet? Om dat makkelijker en duidelijker te maken, hebben we een toetsingskader gemaakt. Het toetsingskader is in de basis gebaseerd op het afwegingskader in het onderzoek van Dialogic.¹ Deze basis hebben we aangescherpt specifiek voor digitale tools voor begrijpelijke teksten. Zo hebben we het aspect 'begrijpelijkheid' toegevoegd. En hebben we een aantal aspecten anders verwoord. Ook hebben we een aantal afwegingen toegevoegd of weggelaten.

Deze aanscherping hebben we gedaan op basis van:

- bevindingen van de kerngroep
- uitkomsten uit de inputsessies en
- interviews met experts.

Waarom is een deel van het onderzoek aangepast ten opzichte van het plan van aanpak?

Eigenlijk was het de bedoeling om ook bij de eindgebruikers te testen. Op basis van de tests wilden we 2 of 3 digitale tools aanbevelen aan het ministerie. Voor deze testen wilden we contact zoeken met enkele leveranciers.

Een ICTU-collega wees ons erop dat dit contact zoeken gevolgen kan hebben voor een mogelijke aanbesteding. De tools van de leverancier waarmee we contact zouden hebben gezocht, mogen dan niet meedoen in een aanbesteding. Deze woorden werden bevestigd door enkele juristen.

Hierna hebben we ervoor gekozen om de scope te wijzigen. We hebben geen contact meer gezocht met leveranciers en geen tools meer getest. In plaats daarvan hebben we het toetsingskader verder uitgewerkt en een plan geschreven

voor het testen van de tools. Deze vind je in dit rapport, samen met aanbevelingen voor het vervolg.

Noot: wie bedoelen we met 'gebruikers' en 'eindgebruikers'?

In dit rapport gebruiken we verschillende woorden voor mensen. Dat kan verwarrend zijn. Daarom leggen we uit wat we bedoelen.

Gebruikers = ambtenaren Als we het hebben over 'gebruikers', bedoelen we ambtenaren. Zij zullen de digitale tools gebruiken om de brieven begrijpelijker te maken.

Eindgebruikers = inwoners en ondernemers Als we het hebben over 'eindgebruikers', bedoelen we inwoners en ondernemers. Dit zijn de mensen die brieven van de overheid krijgen. Ze willen deze brieven goed kunnen begrijpen. Het is natuurlijk belangrijk om de kwaliteit van de brieven met hen te testen.

Kortom, bij: Gebruikers, lees: ambtenaren die begrijpelijkere brieven willen schrijven. Eindgebruikers, lees: inwoners/ondernemers die begrijpelijke brieven willen ontvangen.

3. Welke digitale tools voor begrijpelijke brieven zijn er?

In het digitale tijdperk verandert er veel. Dit gebeurt snel. Dit geldt ook voor de tools die ambtenaren helpen bij het schrijven van duidelijke teksten. Wij hebben onderzocht welke tools er zijn. Zie [hoofdstuk 2](#) voor hoe we dat hebben gedaan.

Hoe hebben we de tools gevonden?

Wij hebben 53 tools gevonden die kunnen helpen bij het begrijpelijker maken van teksten. Al deze tools kunnen hetzelfde: ze maken teksten eenvoudiger om te lezen. Maar ze verschillen op de volgende punten:

- **Herkomst:** sommige tools zijn Nederlands, andere komen uit het buitenland.
- **Functie:** sommige tools herschrijven je hele tekst. Andere geven alleen tips hoe je tekst beter kan, maar herschrijven de tekst zelf niet.
- **Gebruik:** sommige tools gebruik je via een website. Andere kun je downloaden als app of toevoegen aan je tekstprogramma.

Welke soorten tools zijn er?

Hieronder zie je een paar voorbeelden van tools die op onze lijst staan. Hiermee willen we laten zien welke soorten tools er zijn.

Let op: Dit zijn alleen voorbeelden. We raden deze tools niet aan. We hebben niet gekeken of deze tools veilig, betrouwbaar of goed zijn. We hebben alleen gekeken wat voor soorten tools er bestaan.

Een tool die herschrijft: Schrijf Simpel is een tool die teksten herschrijft naar eenvoudige taal. De makers van Lees Simpel hebben deze tool ontwikkeld. Lees Simpel helpt mensen die brieven of andere teksten ontvangen. Deze tool maakt moeilijke teksten begrijpelijker voor de lezer.

Schrijf Simpel is voor mensen die teksten schrijven. De tool helpt je om begrijpelijker te schrijven. Dit werkt als volgt: je voert jouw tekst in de tool in. Schrijf Simpel herschrijft dan de hele tekst naar B1-niveau. Bovendien laat de tool zien waarom iets is veranderd. Bijvoorbeeld omdat de zinnen te lang zijn of omdat er vaktaal is gebruikt.

Een tool die tips geeft: Klinkende Taal is ontwikkeld door Klinkende Taal B.V. en deels eigendom van Loo van Eck. Je kunt Klinkende Taal op 3 manieren gebruiken:

- als webversie op internet
- als plugin in je schrijfprogramma
- als quickscan voor een snelle check.

De tool controleert je tekst en geeft onder meer tips over:

- vaktermen, ofwel moeilijke woorden uit je vakgebied
- ingewikkelde woorden
- zinnen die te lang zijn.

Een tool die AI gebruikt: HIP (Helder, Intelligent en Productief) Communicatieassistent is een tool die AI gebruikt. PNA Group heeft deze tool gemaakt. Het Ministerie van BZK heeft dit mede gefinancierd, maar gaat niet over de inhoud. De tool is onderdeel van het ELSA-lab Terugdringen Armoede en Schulden.

De tool gebruikt een 'open source machine learning-model'. Dit is een vorm van AI. Door AI herschrijft de tool moeilijke teksten. Zo worden de teksten begrijpelijker voor inwoners en ondernemers.

Een tool die geen AI gebruikt: De tool 'Is het B1?' gebruikt geen kunstmatige intelligentie (AI). De tool is een handige website. Met deze tool controleer je of woorden B1 zijn of niet. De tool gaat uit van een woordenlijst van Loo van Eck, een communicatiebureau dat gespecialiseerd is in begrijpelijke communicatie (en naast 'Is het B1?' ook achter de tool 'Klinkende Taal' zit). Deze tool laat zien welke woorden B1 zijn en welke niet. Denk je dat een woord wel of niet B1 is? Dan kun je dat aangeven in de tool.

Een tool uit Nederland: LiNT is een Nederlandse tool die teksten controleert. De Universiteit Utrecht heeft deze tool gemaakt op basis van wetenschappelijk onderzoek. Onderzoekers hebben gekeken naar moeilijke en makkelijke teksten. Ze ontdekten welke kenmerken teksten moeilijk of makkelijk maken. LiNT kijkt naar deze kenmerken in jouw tekst. Daarna geeft de tool advies over hoe moeilijk jouw tekst is. Je voert jouw tekst in bij LiNT. De tool analyseert de tekst. Vervolgens krijg je een

advies over hoe moeilijk de tekst is. Daarna bepaal je zelf welke tips je overneemt om je tekst begrijpelijker te maken.

Een tool uit het buitenland: ChatGPT is een bekende AI chatbot uit Amerika. De tool is niet speciaal gemaakt om teksten te vereenvoudigen. Maar je kunt ChatGPT wel gebruiken voor dit doel. Dit doe je door de juiste instructie te geven. Met een goede prompt kan ChatGPT jouw teksten vereenvoudigen. In ChatGPT kun je een 'agent' maken. Dat is een speciale functie met een vast doel. Je maakt van tevoren een prompt voor dit doel. Bijvoorbeeld: een agent die teksten vereenvoudigt naar B1-niveau. Deze agent heet dan 'Simpel teksten maker B1'.

Een tool met alleen een website: Accessibility is een tool van Bartiméus. Je vindt deze tool op een website. De tool controleert hoe goed je tekst te lezen is. Je voert je tekst in op de website. Dan krijg je een score. Deze score laat zien hoe moeilijk of makkelijk je tekst is. De tool geeft scores van A1 tot C2, waarbij A1 is 'heel makkelijk te lezen' en C2 'heel moeilijk te lezen'. Zo weet je meteen of je tekst begrijpelijk is voor je lezers.

Een tool die je kan downloaden als plug-in: Tolkie Schrijfhulp is een tool die je in Microsoft Word kunt gebruiken. Je downloadt de tool als een plug-in. Daarna kun je de tool direct in Word gebruiken. De tool helpt je op 2 manieren:

- De tool herschrijft je tekst naar begrijpelijke taal.
- De tool geeft tips over hoe je je tekst begrijpelijker kunt maken.

De volledige lijst met beschikbare tools vind je in: Bijlage 1 – Lijst met tools. We hebben geprobeerd om alle tools op te nemen. Maar misschien hebben we er een paar gemist. Dit hebben we niet expres gedaan. Staat er een tool niet op de lijst? Laat het ons weten. Stuur een mail naar: dienstverlening@minbzk.nl.

4. Waar moeten de tools aan voldoen?

Bijlage 2 – Toetsingskader Wat is het toetsingskader?

Het toetsingskader gebruik je om digitale tools te beoordelen en te vergelijken. Het kader kijkt naar 7 aspecten: begrijpelijkheid, praktisch, technisch, financieel, organisatorisch, juridisch en ethisch. Tools die hoog scoren, sluiten het beste aan bij deze punten. Het is een eerste stap om te kiezen welke tools je verder wilt gebruiken. Maar we raden je wel aan deze tools eerst te testen.

Het toetsingskader dat we hebben samengesteld, vind je in [Bijlage 2 – Toetsingskader](#). Hoe we dit toetsingskader hebben samengesteld, lees je in [hoofdstuk 2](#).

Voor deze uitgangspunten gebruiken we de volgende definities:

- Begrijpelijkheid: dit aspect gaat over tekstkwaliteit. Is de tekst duidelijk, vriendelijk en geruststellend? Is de toon prettig? Is de informatie makkelijk te vinden en niet te veel of te weinig? Sluit de tekst aan bij de doelgroep?
- Praktisch: dit aspect gaat over hoe de tool is ingericht en hoe hij werkt.
- Technisch: dit aspect kijkt naar hoe de tool is ingericht. Is de tool veilig genoeg? Is het transparant?
- Financieel: dit aspect richt zich op de kosten en baten. Niet alleen wat de kosten en baten zijn voor een organisatie die de tool wil gebruiken, maar ook naar wat de kosten en baten voor de hele overheid zijn.
- Organisatorisch: dit aspect draait om hoe je de tool kunt inzetten binnen een organisatie en/of hoe je dat voor de hele overheid kunt regelen.
- Juridisch: dit aspect kijkt of de tool voldoet aan wetten, niet discrimineert en of de teksten juridisch kloppen.
- Ethisch: dit aspect controleert of de tool wel eerlijk genoeg is. Is de tool eerlijk ingericht? Is de tool duurzaam genoeg? Is de taal transparant? Heeft de mens de controle.

Hoe werkt het toetsingskader?

Tools worden per onderdeel beoordeeld en krijgen een score. Elk onderdeel heeft een weging. Sommige onderdelen wegen zwaarder dan andere.

Er zijn ook minimale eisen. Als een tool daar niet aan voldoet, valt die af. Dit noemen we 'knock-outs'.

Wij stellen hierbij de volgende minimale eisen voor:

- De tool zorgt ervoor dat de teksten begrijpelijk zijn.
- De data die de tool gebruikt, zijn eerlijk verkregen. Ofwel, de data waarop de tool is getraind is verkregen via licenties of is open source.
- De werkwijze is duidelijk, open en te controleren (geen black box).
- De techniek waarmee de tool werkt, is klaar voor grootschalig gebruik binnen de overheid.
- De gebruiker kan de tool controleren. De gebruiker beslist altijd uiteindelijk.
- De tool moet wettelijk toelaatbaar zijn, daarbij geldt ten minste:
 - De tool voldoet aan de AVG en andere privacyregels.
 - Werkt de tool met LLM/GenAI of regels? Dan moet hij voldoen aan de EU AI-verordening.

Hoe gebruik je het toetsingskader?

In het toetsingskader zijn de 7 aspecten uitgesplitst in verschillende afwegingen die een overheidsorganisatie moet maken als ze een tool wil implementeren. Bij elke afweging staat aangegeven hoe belangrijk deze is. Dit is de wegingsfactor. Op deze manier kan het toetsingskader gebruikt worden als check: voldoet de tool aan wat voor een overheidsorganisatie belangrijk is?

Wil een organisatie bepalen hoe een tool scoort, dan is daar wat onderzoek voor nodig. Niet van elke tool is bekend of deze voldoet aan de criteria. Soms moet de organisatie extra informatie vragen aan de maker. Is alle informatie bekend, dan kan de organisatie de tool beoordelen: van 1 tot 5. 1 = voldoet helemaal niet, 5 = voldoet volledig. Per criterium geldt een wegingsfactor, waarbij 1 = weegt licht, 2 = weegt gemiddeld en 3 = weegt zwaar. Er zijn een paar criteria die per organisatie iets kunnen verschillen qua weging. Bijvoorbeeld het criterium dat gaat over of de tool in de bestaande processen is in te richten of daarnaast zal komen. Hierin kunnen verschillende organisaties verschillende prioriteiten in hebben. En dus kunnen ze hier zelf een weging kiezen.

Een voorbeeld: bij het uitgangspunt 'Begrijpelijkheid' staan 4 criteria. Als een tool op het criterium 'De tool zorgt ervoor dat teksten makkelijk te begrijpen zijn' een 4 scoort en de wegingsfactor is 'zwaar' (3), dan is de gewogen score 12. Vervolgens kunnen ook de scores van de andere criteria bepaald worden. Als de scores van alle criteria bij elkaar worden opgeteld, is de score van dit uitgangspunt bekend. Vervolgens kan dit ook gedaan worden voor de andere uitgangspunten. Zo komt er een totaalscore uit.

Als een organisatie meerdere tools wil toetsen, dan moet dit voor elke tool herhaald worden. Zo wordt duidelijk hoe de tools ten opzichte van elkaar scoren. De tool met de hoogste totaalscore past het beste bij de organisatie.

Let op: dit beoordelen is een eerste stap. Het beoordeelt vooral niet-functionele eisen. Het zegt nog niets over de kwaliteit van de tool of wat gebruikers ervan vinden. Die stappen bespreken we in de volgende hoofdstukken.

Hoe test je tools op kwaliteit?

Het toetsingskader geeft handvatten voor niet-functionele eisen. Dat zijn randvoorwaarden waaraan een tool moet voldoen. Het zegt nog niets over de werkelijke kwaliteit van de tool. Overheidsorganisaties die de tool willen gebruiken, moeten daarom ook apart naar kwaliteit kijken.

Er zijn verschillende meetmethodes (metrieken) om kwaliteit te meten. Zo zijn er meetmethodes die meten:

- hoe nauwkeurig een model of algoritme is, ofwel hoe vaak een model of algoritme het goed heeft
- hoe betrouwbaar een model of algoritme is (geeft het model of algoritme hetzelfde antwoord op dezelfde vraag)
- een combinatie van deze twee onderdelen.

Voor deze meetmethodes bestaan automatische tools. Methodes die belangrijk zijn voor begrijpelijke taal zijn:

- BLEU-score

- BERTS-score.

We lichten deze methodes kort toe.

Met de BLEU-score meet je hoe goed de automatisch gemaakte tekst overeenkomt met de originele tekst. De score vergelijkt zinnen en kijkt naar woordovereenkomsten. Zo kun je controleren of de inhoud gelijk is gebleven. Wil je vereenvoudigde teksten evalueren, dan heb je niet genoeg aan de BLEU-score. Dit komt omdat deze meetmethode niet goed omgaat met verkorte zinnen. Gebruik de BLEU-score daarom samen met een andere methode.

De 'BERTS-score' vergelijkt de gegenereerde tekst met de originele tekst en meet in hoeverre deze overeenkomen. Waar de BLEU-score alleen kijkt naar exacte woordovereenkomsten, kijkt de BERTS-score ook naar de context van de woorden. Daardoor meet de BERT-score of de betekenis behouden is. Hierdoor is BERTS-score de belangrijkste methode voor tools die teksten begrijpelijk maken.¹

Als organisaties de methodes willen gebruiken om teksten te beoordelen die gemaakt zijn door de tools, dan is het belangrijk om gestructureerd te werken. Op deze manier kunnen ze de resultaten met elkaar vergelijken, en weten ze dat ze betrouwbaar zijn. In Bijlage 4 – Testplan staat hoe organisaties de beoordeling van teksten gestructureerd kunnen aanpakken.

5. Welke behoeftes hebben gebruikers voor deze tools?

Tijdens de 3 inputsessies hebben we gesproken met ambtenaren die interesse hebben in digitale tools voor begrijpelijke teksten (zie [hoofdstuk 2](#)). Hieronder staat wat er uit die sessies kwam.

Uitleg én herschrijving

De meeste medewerkers willen niet kiezen tussen:

- uitleg: de tool geeft suggesties voor de tekst
- automatisch herschrijven: de tool maakt de tekst direct begrijpelijker.

De meeste gebruikers willen beide. Aan de ene kant willen ze snel een kant-en-klare tekst, omdat ze niet veel tijd hebben. Aan de andere kant is het belangrijk voor hen om te begrijpen wat er verandert en waarom. Zonder uitleg voelt de tool als een black box, waardoor ze niet kunnen controleren waarom iets is aangepast. Zo leren ze er ook niets van. Uitleg helpt om beter te leren schrijven en om keuzes te kunnen uitleggen. Tegelijk moet die uitleg niet in de weg zitten. Daarom willen ze de optie hebben om de uitleg in of uit te schakelen dat ze zelf kunnen inschakelen. Of ze willen iets dat ze kunnen uitklappen.

¹ Alva-Manchego et al. (2020)

Vertrouwen en controle

Gebruikers vinden het belangrijk dat ze zelf blijven leren van dit soort tools. Ze willen niet alleen geholpen worden, maar ook 'opgevoed' worden in duidelijk schrijven. Vertrouwen speelt een rol: de teksten moeten kloppen, vooral bij juridische teksten. Een tool die alleen maar herschrijft zonder uitleg, voelt onveilig. Gebruikers zijn ook bang dat de schrijfstijl te saai wordt als alles automatisch wordt aangepast. Met uitleg en suggesties kunnen ze zelf meer controle houden over toon en stijl.

Snel een goede tekst

Tegelijk is gemak erg belangrijk. Wat uit de sessies duidelijk naar voren kwam, was dat de gebruikers een tool willen die makkelijk te gebruiken is. En die snel een goede tekst oplevert. Veel gebruikers hebben het druk en willen snel een goede tekst. Als de tool alleen uitleg geeft, gebruiken ze dat misschien niet. Een tool die direct herschrijft, kan helpen om sneller resultaat te krijgen. Vooral bij vaste formuleringen of jargon kan dat efficiënt zijn. Elke organisatie gebruikt andere woorden. De tool moet die woorden kennen. Een spelelement in de tool kan motiveren om beter te schrijven.

Tijdens of na het schrijven?

Over wanneer de tool hulp moet geven, lopen de meningen uiteen. Sommigen zien voordelen in hulp tijdens het schrijven, dat kan helpen om de structuur aan te brengen of sneller tot de kern te komen. Vooral bij standaardteksten of vaste formats kan dat tijd besparen. Maar anderen vinden dat juist storend. Zij willen eerst rustig kunnen schrijven, zonder 'meekijken'. Voor hen werkt een controle achteraf beter. Als middenweg zou de tool eventueel lichte signalen tijdens het schrijven kunnen geven, zoals het rode lijntje in Word dat verschijnt bij woorden die misschien niet goed gespeld zijn. Ook willen gebruikers controle over de tool tijdens het schrijfproces. Ze willen de mogelijkheid hebben om suggesties uit te zetten of alleen specifieke onderdelen te laten controleren.

Zelf kunnen kiezen

Gebruikers willen flexibiliteit in de hulp. Het gaat niet alleen om taal, maar vooral om de duidelijkheid en relevantie van de inhoud. De tool moet helpen om de boodschap helder te maken en overbodige informatie te schrappen. De gebruikers willen niet dat er veel details blijven staan of dat volledige wetsartikelen worden opgenomen. Maar tegelijk willen ze niet dat de tool automatisch bepaalt wat relevant is voor de lezer. Daarom willen gebruikers zelf kunnen kiezen welke aspecten de tool controleert. Bijvoorbeeld alleen zinsopbouw of juist ook structuur.

Gebruikers willen dus een tool die hen helpt om beter te schrijven, met uitleg én snelle oplossingen. Maar ze willen zelf blijven bepalen hoeveel hulp ze op dat moment nodig hebben.

6. Hoe test je tools met gebruikers?

Heeft een overheidsorganisatie de tools beoordeeld met het toetsingskader en de tekstkwaliteit gecheckt? Dan is het tijd om de tools te testen met gebruikers. Omdat we de behoeftes van gebruikers kennen, weten we precies waar je op moet letten. Dat vind je terug in het testplan.

Dit testplan is gemaakt om 6 tools te testen. De overheidsorganisatie kan hiervoor de 6 tools kiezen die het beste uit het toetsingskader en de kwaliteitstest kwamen. Ze kunnen met dit plan ook minder tools testen. Ze hoeven dan alleen het schema even aan te passen naar het aantal tools dat ze willen testen.

Hoe is de test opgezet? En waarom op deze manier?

Testen met mensen is altijd subjectief. Iedereen heeft een andere mening en kijkt anders naar de tools. Om toch zo objectief mogelijk te kunnen oordelen, hebben we een paar regels opgesteld. Die leggen we hier uit.

- **Elk groepje test elke tool één keer.** Zo krijg je van zoveel mogelijk mensen een mening over elke tool.
- **In elk groepje zitten mensen met verschillende functies door elkaar.** Mensen met bepaalde functies kunnen soms een voorkeur hebben. Denk bijvoorbeeld aan communicatieadviseurs en juristen. Ze vinden bepaalde dingen misschien belangrijker dan andere. Door verschillende functies in één groepje te stoppen, wordt dat effect minder.
- **De tools worden in zoveel mogelijk verschillende combinaties getest.** Als je tool 1 steeds vergelijkt met tool 2, lijkt tool 1 misschien heel goed. Maar als je hem vergelijkt met tool 3, valt hij misschien tegen. Dit heet 'paring bias'. Door de tools steeds in andere setjes te testen, voorkom je dat je een scheef beeld krijgt.
- **Elke tool wordt bij elke opdracht in totaal 2 keer getest.** Door elke tool 2 keer te testen, voorkom je dat een tool heel goed scoort, alleen maar omdat hij toevallig heel goed is in één specifieke opdracht. Als de tools alle opdrachten even vaak moeten doen, moeten ze overal goed in zijn om hoog te scoren.
- **De volgorde is steeds anders.** De volgorde waarin elk groepje de tools test, is zoveel mogelijk verschillend. Als een tool als 3e wordt getest, is er al voorkennis door de eerdere 2 tools. Als elk groepje die voorkennis heeft, beïnvloedt dit het oordeel. Dit heeft 'sequencing bias'. Door de volgorde willekeurig te maken, heb je daar geen last van.

In Bijlage 3 – Testen van staat hoe organisaties het testen kunnen aanpakken.

7. Hoe test je de teksten met eindgebruikers?

Het is belangrijk om de tools te testen zoals in hoofdstuk 6 staat. Zo ontdekken overheidsorganisaties of de tools makkelijk werken en of ze passen bij wat ze zoeken.

Maar er is iets nog belangrijkers: testen bij je doelgroep, bijvoorbeeld de inwoner of de ondernemer. Bij hen test je de herschreven teksten of de teksten die zijn verbeterd met tips van de tool. Alleen als de doelgroep deze tekst leest, weet een overheidsorganisatie zeker of ze deze begrijpt. En of de tekst duidelijker is geworden. Bovendien wordt meteen duidelijk wat er nog verbeterd moet worden.

Hoe een overheidsorganisatie dit kan aanpakken, staat in het stappenplan Test je tekst. Een korte uitleg van deze stappen vind je in Bijlage 5 – Stappenplan Test je tekst. Een uitgebreid stappenplan plus toelichting en tips staat op de [website van Gebruiker Centraal](#).

8. Hoe kies je de juiste tool?

Wil een overheidsorganisatie een tool kiezen die helpt bij het schrijven van duidelijke teksten? Dan zijn dit rapport en de bijlagen handig. Maar het is veel informatie, en dan is het voor de overheidsorganisaties lastig om te weten waar ze aan beginnen. Daarom zetten we hier de stappen op een rij die ze kunnen volgen.

- **Stap 1: De lijst met tools bekijken.**
De organisatie start met het bekijken van de lijst met beschikbare tools. (Let op: misschien zijn er al nieuwe tools bij gekomen die niet op deze lijst staan.)
- **Stap 2: Controleren op basisregels**
De organisatie controleert of de tools voldoen aan de minimale vereisten. Tools die niet voldoen, vallen meteen af.
- **Stap 3: Tools beoordelen met het toetsingskader**
De organisatie beoordeelt de overgebleven tools met het toetsingskader. Maximaal de 10 best scorende tools gaan door.
- **Stap 4: Tools testen met meetmethodes**
De organisatie test de overgebleven tools met de meetmethodes. Maximaal 6 tools gaan door naar de volgende stap.
- **Stap 5: Tools testen met gebruikers**
De organisatie test de overgebleven tools met gebruikers, volgens het testplan. De 3 beste tools gaan door naar de laatste stap.
- **Stap 6: Brieven testen met de doelgroep**
De organisatie laat inwoners en ondernemers de brieven lezen die zijn gemaakt met de overgebleven tools. Zo ziet de organisatie of de brieven goed begrepen worden.
- **Stap 7: De beste tool(s) kiezen**
De organisatie kiest de tool(s) die het beste uit de test komen.

Als een overheidsorganisatie deze stappen volgt, vallen de minder goede tools vanzelf af. Uiteindelijk houdt ze dan een paar tools over die goed werken voor duidelijke teksten.

9. Wat zouden we aanbevelen voor het vervolg?

We hebben nu een duidelijk overzicht van welke digitale tools er zijn voor begrijpelijke teksten. En we weten goed waar we op moeten letten bij het beoordelen, testen en kiezen van deze tools. Nu kunnen we kijken naar wat het Ministerie van BZK hiermee kan doen. We hebben de volgende aanbevelingen:

1. Zie digitale tools voor begrijpelijke teksten niet als wondermiddel

Digitale tools voor begrijpelijke teksten kunnen heel handig zijn voor ambtenaren om betere teksten te schrijven. Maar er is op dit gebied nog veel in ontwikkeling en de tools zijn nog niet perfect. Daarnaast geldt altijd: als je slechte teksten in een tool stopt, komen er slechte teksten uit. Dus als de tekst die je begrijpelijker wil maken al slecht is, maakt de tool dat niet beter.

Tools kunnen ambtenaren zeker helpen om begrijpelijker te schrijven. Maar ambtenaren moeten zelf nog steeds weten hoe ze een goede brief schrijven. Een tool vervangt geen opleiding, training, taalcoach of schrijfwijzer. Zie een digitale tool dus als extra hulpmiddel dat je breed kunt inzetten. Gebruik het samen met andere acties, zoals trainingen om medewerkers bewust te maken van begrijpelijke taal en schrijftrainingen.

2. Gebruik de adviezen uit dit rapport om slimme keuzes te maken

In dit rapport vind je handige adviezen om slimme keuzes te maken over digitale tools voor begrijpelijke teksten. Gebruik deze adviezen en deel ze met andere (overheids-)organisaties. Zodat zij ook de juiste keuzes kunnen maken. Zorg ervoor dat niet elke organisatie apart beslist welke tools goed zijn. En werk bij voorkeur met dezelfde regels en eisen. Probeer dit samen te doen met andere ministeries en gemeenten en provincies.

3. Onderzoek verder hoe BZK kan helpen om digitale tools voor begrijpelijke teksten breed in te zetten bij de overheid

Zoals we eerder zeiden, is er veel in ontwikkeling op dit gebied. Er is ook nog veel onduidelijk, terwijl mensen die bij de overheid werken wel behoefte hebben aan duidelijkheid. Ze komen met nieuwe vragen, zoals: 'welke tool mag en kan ik nu echt gebruiken?' En vaak zijn er nog geen antwoorden. Het is dus belangrijk om meer kennis op te bouwen over digitale tools en hoe je die breed bij de overheid kan inzetten. En om duidelijker te krijgen hoe deze tools kunnen bijdragen aan het idee van één overheid. Zodat elke inwoner en ondernemer bij elke overheidsorganisatie begrijpelijke teksten krijgt. We raden aan om dit vervolgonderzoek samen met overheidsorganisaties en samenwerkingsverbanden te doen.

Bronnen

- Alva-Manchego, F., Scarton, C. & Specia, L. (2020). "The (Un)suitability of Automatic Evaluation Metrics for Text Simplification". *Computational Linguistics*. https://doi.org/10.1162/COLI_a_00418
- Dialogic: Vankan, A. Verhagen, P., Boiten, M., Schuitemaker, N. & Brouwers, L. (2025). *Verkenning Inzet van AI voor dienstverlening aan burgers*. In opdracht van het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties.
- Hoekstra, M., Dom, L. & Veenstra, A.F.E. van (2024). "Quickscan AI in de publieke dienstverlening III". TNO. [Quickscan AI in de publieke dienstverlening III](#)
- Hoekstra, M., Bruijne, T., Schoenmakers, M. & Jones, P. (2025). "Overheidsbrede Monitor Generatieve AI: Inzicht in Generatieve AI inzet bij Nederlandse Overheidsorganisaties". TNO Vector. [Overheidsbrede Monitor December 2025](#)
- Rijksoverheid (2024). "Overheidsbrede visie generatieve AI". [Overheidsbrede visie Generatieve AI | Rapport | Rijksoverheid.nl](#)
- (2025). "Overheidsbrede Handreiking voor de verantwoorde inzet van generatieve AI". [Overheidsbrede handreiking voor de verantwoorde inzet van generatieve AI | Rapport | Rijksoverheid.nl](#)
- Sulem, E., Abend, O. & Rappoport, A. (2018). "BLEU is Not Suitable for the Evaluation of Text Simplification". *University of Jerusalem*.
- Verordening (EU) 2024/1689 van het Europees Parlement en de Raad van 13 juni 2024 tot vaststelling van geharmoniseerde regels betreffende artificiële intelligentie (AI verordening). *Publicatieblad van de Europese Unie*. L 2024/1689. [Verordening - EU - 2024/1689 - EN - EUR-Lex](#)

Bijlagen

Bijlage 1 – Lijst met tools (alfabetische volgorde)

#	Tool	Vorm	Publiek /Privaat	Open source?	Link	NL, EU of Niet-EU
1	Accessibility	Tool in website	Privaat	Onbekend	Accessibility	NL
2	AI rewriter to human	Website	Privaat	Onbekend	AI Rewriter	NL
3	AI-Migo	Onbekend	Publiek (RVO)	Onbekend	-	NL
4	AlleTaal	Tool	Publiek Gem. Amsterdam	Ja	AlleTaal	NL
5	Burger Lovable App	Website	Publiek (BZK, GreenPT, Digitaal Toegankelijk & Platform AI & Overheid)	Ja	Burger Lovable	NL
6	Capibaira	Widget, app of CMS	Privaat	Onbekend	Capibaira	NL
7	ChatGPT	Chatbot	Privaat	Nee	ChatGPT	Niet-EU
8	Claude	Chatbot	Privaat	Nee	Claude	Niet-EU
9	CoPilot	Chatbot	Privaat	Nee	CoPilot	Niet-EU
10	Copy AI	Tool	Privaat	Nee	Copy AI	Niet-EU
11	DeepL Write	Website	Privaat	Nee	DeepL	EU
12	Duck.AI	Chatbot	Privaat	Nee	Duck.AI	Niet-EU
13	Dutch Simplification	Via Hugging Face	KU Leuven	Ja	Dutch Simplification	EU
14	Ecosia	Chatbot	Privaat	Nee	Ecosia	EU
15	Eskritor AI Writer	Website	Privaat	Nee	Eskritor	Niet-EU
16	EU accessible tekst	Website + Extensie/Plug-in	Publiek (EU)	Ja	EU Accessible	EU
17	Euria	Chatbot	Privaat	Nee	Euria	Zwitserland
18	Eye-able	Plug-in	Privaat	Nee	Eye-able	EU

#	Tool	Vorm	Publiek /Privaat	Open source?	Link	NL, EU of Niet-EU
19	Gemini	Chatbot	Privaat	Nee	Gemini	Niet-EU
20	GreenPT	Chatbot	Privaat	Nee	GreenPT	NL
21	Grok	Chatbot	Privaat	Nee	Grok	Niet-EU
22	Hella	Agent in CoPilot	Publiek (Gem. Kampen)	Nee	CoPilot	-
23	HIP Communicatie assistent	Tool	Publiek-Privaat (PNA group i.o.v. BZK)	Ja	HIP	NL
24	Is het b1?	Website	Privaat	Nee	Is het b1?	NL
25	Jasper AI	Tool	Privaat	Nee	Jasper	Niet-EU
26	Klinkende taal	Webversie, plug-in, quickscan	Privaat	Nee	Klinkende Taal	NL
27	LAICA/GovChat-NL	Platform	Provincie Limburg	Ja	GovChatNL	NL
28	Language Tool	Website	Privaat	Nee	Language Tool	Niet-EU
29	Le Chat Mistral	Chatbot	Privaat	Nee	Le Chat	EU
30	Lesson Up	Tool	Privaat	Nee	Lesson Up	NL
31	LiNT	Website	Publiek (Universiteit Utrecht)	Ja	LiNT	NL
32	Lumo	Chatbot	Privaat	Ja	Lumo	EU
33	NotebookLM	Chatbot	Privaat	Nee	NotebookLM	Niet-EU
34	Postex	API, Add-on	Privaat	Nee	Postex	NL
35	Quillbot	Website + Extensie/Plug-in	Privaat	Nee	Quillbot	Niet-EU
36	ReadEasy.ai	Plug-in + Extensie	Privaat	Nee	ReadEasy.ai	NL
37	Rytr.me	Extensie	Privaat	Nee	Rytr	Niet-EU
38	SafeGPT	Tool en plug-in	Privaat	Nee	SafeGPT	NL

#	Tool	Vorm	Publiek /Privaat	Open source?	Link	NL, EU of Niet-EU
39	Schrijf Simpel	Website	Privaat	Nee	Schrijf Simpel	NL
40	Schrijfassistent Nederlands	Website	VRT, Standaard, KU Leuven	Nee	Schrijfassistent Nederlands	EU
41	Schrijfkader	Add-in	Privaat	Nee	Schrijfkader	NL
42	Scribbr	Website	Privaat	Nee	Scribbr	NL
43	Simpele teksten maker B1	Agent in ChatGPT	Privaat	Nee	Simpele teksten maker	Niet-EU
44	Slimtaal	Tool	Privaat	Nee	Slimtaal	NL
45	Taalscan	Tool	Publiek (Belastingdienst)	Onbekend	-	NL
46	Templafy	Tool	Privaat	Nee	Templafy	Niet-EU
47	Texamen	Tool	Privaat	Nee	Texamen	NL
48	Textmetrics	Plug-in	Privaat	Nee	Textmetrics	NL
49	Tolkie schrijfhulp	Word- of browserplug-in	Privaat	Nee	Tolkie	NL
50	Typetone AI	Plug-in	Privaat	Nee	Typetone	NL
51	Vlam.AI	?	Publiek (Rijksoverheid)	?	?	NL
52	Zenochat	Chatbot	Privaat	Nee	Zenochat	EU
53	Zoekeenvoudi woorden.nl	Website	Privaat	Nee	Zoek eenvoudige woorden	NL

Bijlage 2 – Toetsingskader

Afweging	Criteria	Score (1 – 5)	Weging (1, 2 of 3)	Gewogen score (Score * weging)
Begrijpelijkheid	1: De tool zorgt ervoor dat teksten makkelijk te begrijpen zijn De tool is speciaal ontworpen om teksten makkelijk te begrijpen. Hij gebruikt officiële en goedgekeurde regels, zoals B1, LiNT of ISO.		3	
	2: Inhoud en betekenis blijven kloppen Juridische en inhoudelijke woorden blijven kloppen. De betekenis verandert niet.		1	
	3: Toon van de tekst De toon van de tekst blijft hetzelfde.		1	
	4: Goed te vinden en overzichtelijk De tool maakt de tekst duidelijker zonder de opbouw te veranderen. Informatie blijft makkelijk te vinden.		1	
Juridisch	1: Gelijke behandeling De tool heeft geen instellingen of doelen die mensen uitsluiten of anders behandelen.		3	
	2: De tekst blijft kloppen volgens de wet De tekst blijft juist volgens de wet. Regels en termen die belangrijk zijn, blijven hetzelfde.		2	
	3: Bescherming van gegevens en privacy		3	

Afweging	Criteria	Score (1 – 5)	Weging (1, 2 of 3)	Gewogen score (Score * weging)
	<i>De tool voldoet aan privacyregels. Persoonlijke gegevens worden goed beschermd.</i>			
Technisch	<p><i>1: Veiligheid</i></p> <p><i>De tool verkleint de kans op datalekken en misbruik van buitenaf. Gegevens worden niet langer bewaard dan nodig.</i></p>		3	
	<p><i>2: Kwaliteit van de data</i></p> <p><i>De tool laat duidelijk zien welke data is gebruikt om hem te trainen. Die data zijn van hoog niveau.</i></p>		3	
	<p><i>3: Steeds hetzelfde resultaat</i></p> <p><i>De tool werkt zoveel mogelijk op dezelfde manier. Dezelfde tekst wordt zoveel mogelijk op dezelfde manier makkelijker gemaakt. De tool verzint zo weinig mogelijk fouten.</i></p>		2	
	<p><i>4: Bijgewerkt en actueel</i></p> <p><i>De tool wordt regelmatig verbeterd en bijgewerkt. Zo blijft hij veilig en goed werken.</i></p>		1	
	<p><i>5: Geschikt voor de hele overheid</i></p> <p><i>De tool kan uitgebreid worden zodat alle overheidsorganisaties hem kunnen gebruiken. Ook als veel mensen hem tegelijk gebruiken.</i></p>		Zelf invullen	

Afweging	Criteria	Score (1 – 5)	Weging (1, 2 of 3)	Gewogen score (Score * weging)
Functioneel	1: Toegankelijk en eerlijk ontwerp De tool is makkelijk te gebruiken voor ambtenaren. Hij behandelt iedereen gelijk en sluit niemand buiten.		3	
	2: Onjuiste informatie voorkomen De tool maakt zo min mogelijk verzinsels. Beter geen antwoord dan een fout antwoord of een foute herschrijving.		2	
	3: Inbouwen in werkprocessen Het beste is een tool die werkt in de programma's waarmee de gebruiker al werkt. Voor de gebruiker is het makkelijker als deze minder knip-en-plakwerk hoeft te doen en minder hoeft te wisselen tussen schermen.			Zelf invullen
	4: Organisaties kunnen de tool aanpassen Organisaties kunnen ervoor zorgen dat het resultaat past bij hun eigen wensen, bijvoorbeeld stijl of opmaak.			Zelf invullen
Organisatorisch	1: Zelf werken met de tool Kunnen organisaties de tool helemaal zelf gebruiken en regelen? Dus zonder de hulp van anderen nodig?			Zelf invullen

Afweging	Criteria	Score (1 – 5)	Weging (1, 2 of 3)	Gewogen score (Score * weging)
	2: De tool verder ontwikkelen samen met BZK Kunnen BZK en de leverancier samenwerken om de tool beter te maken?		Zelf invullen	
	3: Hulp van de leverancier voor alle overheidsorganisaties Hebben organisaties hulp nodig van de leverancier bij het gebruik van de tool? Werkt deze hulp ook als alle overheidsorganisaties de tool gebruiken?		Zelf invullen	
Ethisch	1: Duidelijkheid over hoe de tool werkt Is duidelijk of de tool AI gebruikt? Kun je zien waarom de tool teksten verandert? Zie je wat de tool precies heeft aangepast?		3	
	2: De mens beslist Een mens moet altijd de laatste keuze kunnen maken bij het gebruik van de tool.		3	
	3: Milieuvriendelijk gebruik Laat de tool zien hoeveel energie en water het gebruikt? Kan de tool milieuvriendelijker worden gemaakt?		2	
	4: Onderzoek door neutrale organisaties Andere organisaties kunnen de tool beoordelen. Deze beoordelingen worden openbaar gepubliceerd.		1	

Afweging	Criteria	Score (1 – 5)	Weging (1, 2 of 3)	Gewogen score (Score * weging)
Financieel	<p>1: Kosten en voordelen voor organisaties</p> <p>De kosten zijn niet te hoog vergeleken met wat de tool oplevert.</p>		2	
	<p>2: Kosten en voordelen voor de hele overheid</p> <p>Ook als alle overheidsorganisaties de tool gebruiken, zijn de kosten niet te hoog vergeleken met wat de tool oplevert.</p>		2	

Bijlage 3 – Testen van het model of algoritme

Als je een tool wilt gebruiken, is het belangrijk om ook naar de kwaliteit te kijken.

Er zijn verschillende meetmethodes (metrieken) om kwaliteit te meten. Zo zijn er meetmethodes die meten:

- hoe nauwkeurig een model of algoritme is, ofwel hoe vaak een model of algoritme het goed heeft
- hoe betrouwbaar een model of algoritme is (geeft het model of algoritme hetzelfde antwoord op dezelfde vraag)
- een combinatie van deze twee onderdelen.

Voor deze meetmethodes bestaan automatische tools. Methodes die belangrijk zijn voor begrijpelijke taal zijn:

- BLEU score
- BERTS-score.

We lichten deze methodes kort toe.

Met de BLEU-score meet je hoe goed de automatisch gemaakte tekst overeenkomt met de originele tekst. De score vergelijkt zinnen en kijkt naar woordovereenkomsten. Zo kun je controleren of de inhoud gelijk is gebleven. Wil je vereenvoudigde teksten evalueren, dan heb je niet genoeg aan de BLEU-score. Dit komt omdat deze meetmethode niet goed omgaat met verkorte zinnen. Gebruik de BLEU-score daarom samen met een andere methode.

De 'BERTS-score' vergelijkt de gegenereerde tekst met de originele tekst en meet in hoeverre deze overeenkomen. Waar de BLEU-score alleen kijkt naar exacte woordovereenkomsten, kijkt de BERTS-score ook naar de context van de woorden. Daardoor meet de BERT-score of de betekenis behouden is. Hierdoor is BERTS-score de belangrijkste methode voor tools die teksten begrijpelijk maken.²

Als je de methodes wilt gebruiken om teksten te beoordelen die gemaakt zijn door de tools, dan is het belangrijk dat je gestructureerd werkt. Op deze manier kun je de resultaten met elkaar vergelijken, en weet je dat ze betrouwbaar zijn.

Dit kun je als volgt aanpakken:

1. Kies een voorbeeldtekst waarvan je zeker weet waarom hij onbegrijpelijk is. Deze tekst is de referentietekst.
2. Laat de tool die tekst begrijpelijk maken.

² Alva-Manchego et al. (2020)

3. Vergelijk met de meetmethodes de begrijpelijk gemaakte tekst met de referentietekst.
4. Bereken de scores met een online calculator (voor elke metriek kun je er enkele vinden).
5. Als een tool goed scoort op beide methodes, heb je aangetoond dat de tool betrouwbaar is en goede kwaliteit levert.

Let op: als een tool volgens deze methodes goed scoort, zegt dit nog niet alles. Het blijft belangrijk om gebruikerstest te organiseren en de gegenereerde teksten kwalitatief te beoordelen.

Bijlage 4 – Testplan

Nadat een overheidsorganisatie de tools heeft beoordeeld aan de hand van het toetsingskader en de kwaliteit van de teksten heeft gecontroleerd, is het tijd om de tools te testen met gebruikers. Omdat we de behoeftes van gebruikers weten, weten we waar je op moet letten. Dat vind je terug in het testplan. Dit testplan is gemaakt om 6 tools te testen. De overheidsorganisatie kan hiervoor de 6 tools gebruiken die het beste door het toetsingskader en de kwaliteitstest gekomen zijn. Ze kunnen met dit testplan ook minder tools testen. Ze hoeven dan alleen het schema aan te passen naar het aantal tools dat ze willen testen. In bijlage 4: Testplan staat hoe ze dit testen kunnen aanpakken.

Organiseer (ten minste) twee sessies om de tools te gaan testen. Zet in om per sessie 6 tot 9 personen deel te laten nemen. Zorg hierbij dat de deelnemers bestaan uit een mix van communicatieadviseurs en/of taalcoaches, beleidsmedewerkers en/of uitvoeringsmedewerkers, en juridische medewerkers. Tijdens de testsessie doorlopen 3 groepen (elk van 2-3 personen) 3 opdrachten om de tools te testen.

De tools die elk groepje per opdracht test worden willekeurig verdeeld. Doordat je twee testsessies organiseert komt elke tool in totaal 2 keer aan bod, per opdracht. Ook varieert de samenstelling van de tools per opdracht. Elk groepje deelnemers test zo 1 keer elke tool.

Een aantal uitgangspunten voor het testen:

- Elk groepje test 1 keer elke tool;
- Elk groepje bestaat uit een mix van bovengenoemde functies;
- De tools komen in zoveel mogelijk verschillende combinaties aan bod;
- Elke tool komt in totaal 2 keer aan bod bij elke opdracht;
- De volgorde van tools die elk groepje aflegt is zoveel mogelijk verschillend.

In dit schema is dit overzichtelijk uitgebeeld:

Opdracht	Groep 1	Groep 2	Groep 3	Groep 4	Groep 5	Groep 6
1	Tools A + D	E + F	B + C	A + B	D + F	C + E
2	Tools B + E	A + C	D + F	C + D	A + E	B + F
3	Tools C + F	B + D	A + E	E + F	B + C	A + D

Om het behapbaar te houden stellen wij voor om in de eerste testsessie groepen 1 tot en met 3 te laten testen en in de tweede sessie groep 4 tot en met 6. De 6 tools die je wil testen kun je willekeurig koppelen aan een letter in het schema. Het gaat erom dat elke tool eerlijk getest wordt.

Testopdrachten

Voor het testen stellen wij de volgende drie opdrachten voor:

Opdracht 1, '**Gebruiksvriendelijkheid**'. Het groepje krijgt een voorbeeldbrief, deze gaan zij omzetten naar begrijpelijke taal met behulp van 2 tools. Het groepje let hierbij vooral op de volgende gebruiksaspecten:

- Aantal nodige handelingen;
- Werkt de tool zoals verwacht;
- Zit de tool op een logische plek;
- Vindt de groep de tool prettig in gebruik.

Opdracht 2, '**Resultaat**'. Het groepje krijgt 2 dezelfde brieven die zijn vereenvoudigd door 2 verschillende tools. Ze weten niet door welke tools de brieven zijn vereenvoudigd. Het groepje kijkt naar begrijpelijkheidsaspecten van de brief en hoe goed deze zijn verbeterd. Ze maken hierbij dus ook gebruik van de brief zoals hij bestond vóór de vereenvoudiging. Het groepje let op de volgende aspecten:

- Begrijpelijkheid van de woorden en zinnen;
- De structuur van de brief;
- De vindbaarheid van de (belangrijke) inhoud;
- De toon van de brief;
- De hoeveelheid aan informatie;
- Of de brief geruststellend overkomt.

Opdracht 3, '**Functionaliteiten**'. Het groepje gaat 'spelen' met 2 tools. Hierbij is het hoofddoel niet om daadwerkelijk een brief te vereenvoudigen, maar kijkt het groepje naar de functies van de tool. Hierbij letten ze op de volgende aspecten:

- Missen er bepaalde functies, die ze wel verwacht hadden;
- Zijn alle functies die erin zitten nodig;
- Zijn de functies eenvoudig en logisch toe te passen;
- Wat is het oordeel over de functies van de tools?

Opdracht 4, '**Nagesprek**'. Plenair bespreek je de tools en opdrachten na met alle deelnemers. Wat is opgevallen? Wat is het oordeel over de tools? Hiermee geef je ruimte aan kwalitatieve duiding van de testsessie.

Bijlage 5 – Stappenplan Test je tekst

Door de tools te testen weet je of de tools makkelijk werken en of ze passen bij wat je zoekt. Maar er is iets nog belangrijkers: testen bij je doelgroep, bijvoorbeeld de inwoner of de ondernemer. Bij hen test je de herschreven teksten of de teksten die je zijn verbeterd met tips van de tool. Alleen als de doelgroep deze tekst leest, weet een overheidsorganisatie zeker of ze deze begrijpt. En of de tekst duidelijker is geworden. Bovendien wordt meteen duidelijk wat er nog verbeterd moet worden.

We leggen hieronder kort uit hoe je dit aanpakt met het stappenplan Test je tekst. Hieronder geven we kort de stappen weer. Op de [website van Gebruiker Centraal](#) kun je dit stappenplan inclusief uitgebreide toelichting en tips vinden.

Introductie

Voor het testen van teksten met je doelgroep stellen we de hallway methode voor. Dit houdt in dat je gaat testen met mensen op locatie; op straat, in het buurthuis of bijvoorbeeld in de hal van de gemeente. Het stappenplan bevat een combinatie van 2 testmethoden: de plus- en minmethode, en het vraaggesprek. Je krijgt hiermee in korte tijd zicht op de sterke en zwakke stukken van de tekst. Het is een laagdrempelige methode en levert veel inzichten op.

Stap 1: Bereid je voor

De tekst is geschreven met een doel, zoals het aanzetten tot actie of informeren. Wat moeten lezers na het lezen van de tekst doen, kunnen of weten? Beschrijf dit zo specifiek mogelijk.

Bepaal ook wie je lezer is en wat dat betekent voor de plek waarop je wilt gaan testen. Waar bevindt je lezer zich?

Stel onderzoeksvragen en een lijst met vragen voor tijdens de test op. Bijvoorbeeld: nodigt de tekst uit tot lezen? Welke informatie uit de brief is de lezer bijgebleven?

Kies datum, tijd en locatie van je test en bereid voor wat je gaat vertellen aan de lezer. Verzamel alles wat je nodig hebt tijdens de test, zoals de tekst zelf, papier, pennen en bijvoorbeeld een klembord.

Stap 2: Voer de test uit

Spreek een persoon binnen jouw doelgroep aan. Gebruik de introductie die je in stap 1 hebt gemaakt. Start met een uitleg en laat de lezer daarna de tekst lezen. Vraag of diegene plussen (duidelijk) en minnen (onduidelijk, vragen) in de tekst aan wil geven. Bespreek de plussen en minnen en stel de vragen die je in stap 1 hebt geformuleerd.

Soms loopt een test anders dan je had gepland, bijvoorbeeld dat je vragen in een andere volgorde stelt. Dat is niet erg, zolang je antwoorden op je vragen krijgt.

Bedank de lezer.

Stap 3: Verwerk de feedback

Het is tijd om de opgehaalde feedback te verwerken. Om vervolgens tot conclusies te komen over de sterke en zwakke punten in je tekst. Je hoeft niet alle feedback over te nemen; je weegt af (bijvoorbeeld met een collega) wat je overneemt en wat niet. Kijk daarbij ook naar de doelen van je tekst die je in stap 1 formuleerde. Pas de tekst aan; herschrijf de tekst met de feedback die je mee wilt nemen. Je kunt ook die tekst weer testen als je dat wilt.

Bijlage 6 – Gesprekspartners

Tijdens dit traject hebben we met veel mensen gesproken. Wij hebben hun input gebruikt in dit rapport. Dat heeft belangrijke inzichten opgeleverd. Zo hebben we tijdens het hele traject maandelijks gesproken met een kerngroep die met ons meedacht. Ook hebben we experts benaderd en gesproken om zo hun input mee te nemen. Hieronder vind je een overzicht van de organisaties waar al deze gesprekspartners werken. Wegens privacyoverwegingen benoemen we hen niet bij naam.

Kerngroep

Gemeente Den Haag	Gemeente Haarlemmermeer
Gemeente Utrecht	Gemeente Waadhoeke
Gemeente Zwolle	Stichting Lezen en Schrijven
SVB	Vlaamse Overheid
Woonin	

Gesprekspartners

BZK	Hogeschool van Amsterdam
ICTU	Platform AI en Overheid
SVB	Taalunie
Universiteit Utrecht	

Deelnemers inputsessies

Belastingdienst	Bevolkingsonderzoek Nederland
Dienst Toeslagen	DPC
Gemeente Kampen	Gemeente Katwijk
Gemeente Oost Gelre	Gemeente Pijnacker-Nootdorp
Gemeente Rijswijk	Gemeente Roosendaal
Gemeente Sittard-Geleen	Gemeente Steenwijkerland
Gemeente Tilburg	Gemeente Valkenswaard
GGD Amsterdam	Inspectie Belastingen, Toeslagen en Douane
Kenniscentrum Sport en Bewegen	KVK
Logius	RVO
Waterschap Hunze en Aa's	

Bijlage 7 - Over de taal in dit rapport

Dit rapport gaat over digitale hulpmiddelen en begrijpelijk schrijven. Het leek me logisch om deze hulpmiddelen meteen te gebruiken voor het herschrijven van dit rapport. Nou ja, niet alle 53. Maar enkele. Daarbij had ik 2 voorwaarden:

- Ik moest de tool kunnen gebruiken vanaf mijn kantoor (sommige tools zijn afgeschermd).
- Ik hoefde er niet voor te betalen.

Ik koos tools zonder uitgebreid onderzoek, maar zorgde er wel voor dat er verschillende soorten aan bod kwamen. In hoofdstuk 3 staan 8 soorten tools. Ik heb 4 soorten gebruikt. In totaal werkte ik met 6 tools:

- een tool die herschrijft: Schrijf simpel
- een tool die geen AI gebruikt: Is het B1
- een tool uit het buitenland: Chat GPT, Claude en Gemini
- een tool met alleen een website: Accessibility

Ik heb deze tools niet onderzocht of getest. Ik heb ze gewoon gebruikt. En op basis daarvan mijn ervaringen opgeschreven. En die vielen tegen. Dat is ook de reden dat ik niet 1 maar 3 buitenlandse tools heb gebruikt.

Eigenlijk wilde ik alleen Chat GPT gebruiken. Maar al snel kreeg ik de melding 'Slow cowboy'. Ik had genoeg gedaan, en moest even wachten. Omdat de resultaten toch al tegenvielen, stapte ik over naar Claude. Daarna naar Gemini, omdat ik aan mijn dagelijkse limiet zat. Maar ook Claude en Gemini vielen tegen. Ze verschilden niet veel van Schrijf simpel. De grootste problemen waren:

- Niet alle teksten werden herschreven.
- Sommige moeilijke woorden bleven staan.
- Vaak werd te veel informatie verwijderd, ook informatie die belangrijk was.
- De resultaten waren niet altijd hetzelfde. Als ik dezelfde tekst 2 keer invoerde, kreeg ik 2 verschillende uitkomsten.
- Zinnen werden vaak verkeerd herschreven. Daardoor veranderde de betekenis.
- Steeds werden dezelfde fouten gemaakt.
- Soms werd er iets verzonnen dat er niet in stond.

Vaak moest ik de herschreven tekst daarna nog zelf aanpassen. Daarvoor gebruikte ik www.ishetb1.nl. Maar daar had ik ook niet altijd iets aan. Je kunt maar een beperkt aantal woorden invoeren. Bij veel woorden krijg je de melding 'Dit staat niet in onze woordenlijst'. Voor jargon en ambtelijke woorden helpt de site weinig. En als een woord wel in de lijst staat maar geen B1 is, heb je er nog niet altijd iets aan. Je krijgt soms zoveel opties dat het lastig is om de juiste te kiezen. En soms krijg je bij een zelfstandig naamwoord ook werkwoorden als optie. Dan moet je zelf maar uitzoeken hoe dat in je tekst past.

Ook gebruikte ik de tool Accessibility. Deze tool laat zien welk taalniveau een tekst heeft. Dat leverde 2 inzichten op:

- Een taalniveau zegt niet alles. Dan weet je nog niet hoe goed bijvoorbeeld de tekst aansluit op de doelgroep. En of de boodschap overkomt.
- Eén tool is vaak niet genoeg. Accessibility vertelt je het taalniveau. Maar hoe pas je het aan? Daarvoor heb je een andere tool nodig. Maar welke?

Eigenlijk is er een tool nodig die helpt om beter te schrijven, met uitleg en snelle oplossingen. Dat vind ik niet alleen — dat vinden ook de gebruikers (hoofdstuk 5). En die tool moet ook nog nauwkeurig en betrouwbaar zijn. Het liefst allebei.

Het zal een uitdaging zijn om zo'n tool te vinden. Ik denk dat de BERTS-score daarbij kan helpen. Deze meetmethode kijkt naar woordovereenkomsten en of de betekenis bewaard blijft (zie hoofdstuk 4). Op basis van die score kan een tekst worden aangepast.

Tools worden beter. Dat weet ik zeker. Maar of een herschreven tekst klopt, begrijpelijk is én de juiste betekenis heeft — dat moet altijd gecontroleerd worden. En dat blijft mensenwerk. Eerlijk gezegd vind ik die constatering helemaal niet erg.

Joost Mangnus