



Whitepaper

Verantwoord en nuttig gebruik van generatieve AI

ChatGPT als sparringpartner

Inhoud

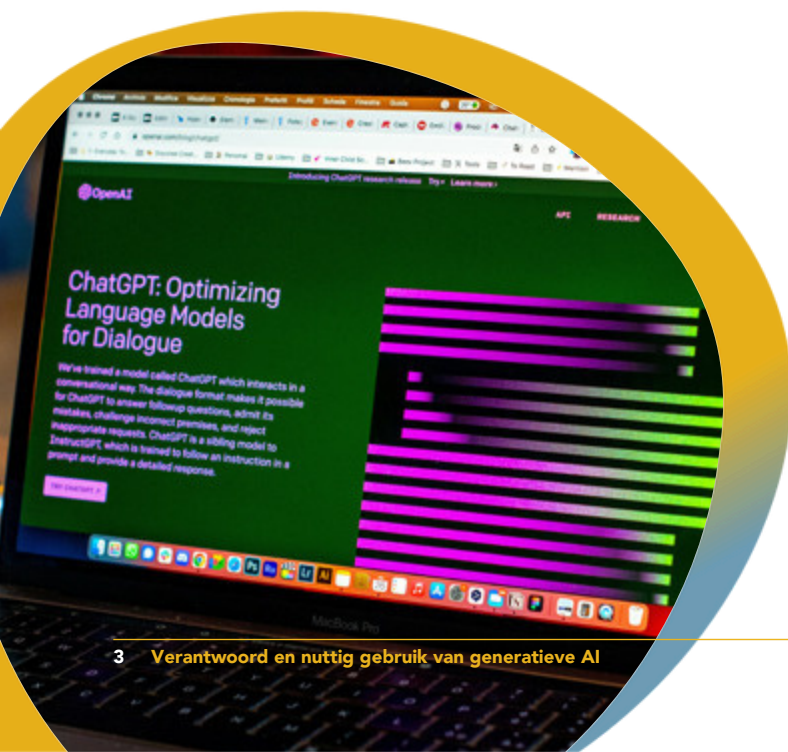
Introductie	3
Wat is ChatGPT en hoe gebruik je het?	4
Generative pre-trained transformer	4
Large Language Model	5
Wat zijn de ethische overwegingen bij het gebruik van ChatGPT?	6
Risico's	6
Maatschappelijke uitdagingen	7

Introductie

“Generatieve AI is een krachtig verlengstuk van het analytisch en scheppend vermogen van mensen gebleken”¹. In een kamerbrief van onder andere het ministerie van Koninkrijksrelaties en Digitalisering werd de Tweede Kamer begin 2024 geïnformeerd over de kansen en mogelijkheden van generatieve AI. De brede inzetbaarheid creëert kansen voor burgers, bedrijven, de overheid en samenleving. Tegelijkertijd werd de Kamer ook gewaarschuwd voor de risico’s, omdat publieke waarden en grondrechten zoals transparantie, privacy, autonomie en non-discriminatie in gevaar kunnen komen door onverantwoorde ontwikkeling of misbruik van generatieve AI². Eind 2023 werd in een Kamerbrief zelfs benoemd dat het gebruik van bijvoorbeeld ChatGPT door Rijksorganisaties in beginsel niet is toegestaan. De reden hiervan was dat niet-gecontracteerde generatieve AI-toepassingen (waaronder ChatGPT), over het algemeen niet aantoonbaar voldoen aan de geldende privacy- en auteursrechtelijke wetgeving³.

Niet geheel onverwacht dat generatieve AI bij veel overheidsorganisaties dan ook eerder als bedreiging dan als kans wordt gezien. AI verandert de voorwaarden waaronder we informatie krijgen en mogen gebruiken, en kan daarmee van invloed zijn op de democratie en onze rechtsstaat. Tegelijkertijd benadrukt de Nederlandse politiek ook het belang van aansluiten op de razendsnelle ontwikkelingen op het gebied van kunstmatige intelligentie. ChatGPT, maar ook andere vormen van generatieve AI bieden uitkomsten in productiviteit en effectiviteit. Ondanks dat het gebruik van ChatGPT voor rijksambtenaren dus sterk wordt afgeraden, blijkt uit een rondvraag van EenVandaag dat 74% van de ambtenaren aangeeft ChatGPT te gebruiken⁴. Dit zien wij ook in de dagelijkse praktijk bij onze opdrachtgevers. Toch weten veel medewerkers niet precies wat er wel en niet mag met tools als ChatGPT.

Een interessant dilemma dus; hoe kan je als maatschappelijke organisatie een veelbelovende en productieve tool zodanig inzetten dat je kan voldoen aan de geldende privacy en auteurswetgeving? We kunnen er niet omheen dat generatieve AI meer en meer onderdeel zal worden van onze toekomst. Verantwoord gebruik van ChatGPT met duidelijke randvoorwaarden biedt kansen. Laat generatieve AI niet tegen je werken, werk er juist mee samen. Maar hoe doe je dat? Hoe zet je ChatGPT verantwoord en nuttig in? In dit whitepaper leggen we je uit hoe ChatGPT werkt, wat de potentie is van deze vorm van AI, en wat het vraagt van de mens en maatschappelijke organisaties om samen te werken met deze nieuwe technologie.



- 1 Kamerbrief bij overheidsbrede visie generatieve AI
- 2 Kamerbrief bij overheidsbrede visie generatieve AI
- 3 Kamerbrief over voorlopig standpunt voor Rijksorganisaties bij het gebruik van generatieve AI
- 4 Meerderheid gemeenten gebruikt ChatGPT en bijna de helft weet niet wat hun medewerkers ermee doen



Wat is ChatGPT en hoe gebruik je het?

Generative pre-trained transformer

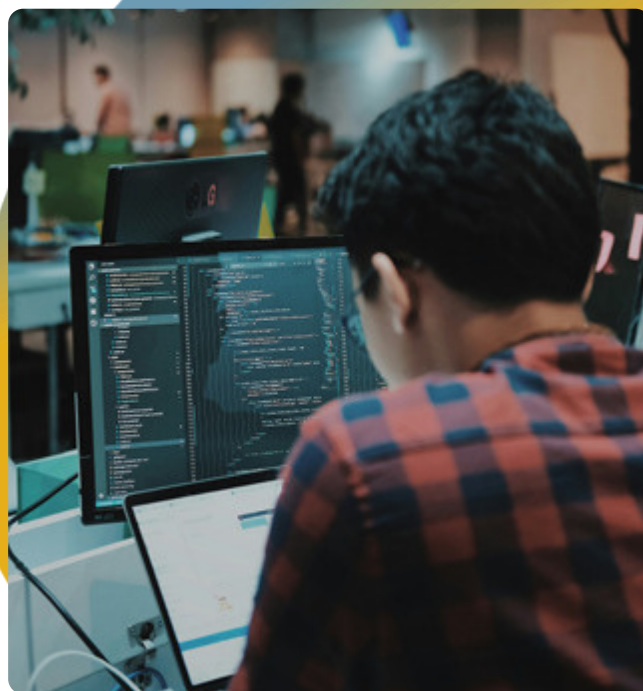
Om te beginnen is het handig om te begrijpen waar ChatGPT voor staat en wat het nou eigenlijk precies is. GPT staat voor generative pre-trained transformer. Het is een groot netwerk, dat is getraind op een zeer grote hoeveelheid data, om zo overtuigende en mensachtige taaluitkomsten te produceren. Een transformer is een geavanceerd type machine learning-model dat wordt gebruikt voor het verwerken en genereren van taal. Het model is generatief omdat het in staat is om nieuwe teksten te creëren die lijken op de voorbeelden waarmee het is getraind.

Pre-trained verwijst naar het feit dat het model wordt getraind aan de hand van een grote hoeveelheid data voorafgaand aan het specifieke doel waar het voor gebruikt wordt. De combinatie van deze drie elementen maakt het een krachtig instrument voor taalverwerking en tekstgeneratie⁵. Sinds de release van GPT-4 is ChatGPT naast het genereren van tekst ook in staat om afbeeldingen te maken.

⁵ Hoe werkt ChatGPT en wat kun je ermee?

Large Language Model

De technologie achter ChatGPT is deep learning: een methode die computers leert om zelfstandig te leren en beslissingen te nemen, zonder dat ze expliciet geprogrammeerd worden. In plaats van de computer precies te vertellen wat deze moet doen, laten we veel voorbeelden zien, zodat de computer zelf kan leren. ChatGPT werkt op basis van een Large Language Model (LLM). Dit betekent dat je een vraag (prompt) kunt stellen, waarna ChatGPT een tekst voor je genereert. Om deze tekst te kunnen genereren hoeft deze chatbot de prompt niet te begrijpen. De prompt geeft de chatbot een context, waarna deze door middel van kansberekeningen bepaalt welke woorden het meest logisch achter elkaar geplaatst kunnen worden om zo een zin te vormen. De tool creëert op deze manier nieuwe, unieke tekst, in plaats van bestaande teksten weer te geven zoals zoekmachines als Google doen. Duidelijk zijn in de prompts die je invoert is erg belangrijk. Hoe beter de prompt is geformuleerd, hoe beter je antwoord zal zijn. Je moet zo precies mogelijk uitleggen wat je van ChatGPT verwacht⁶.



⁶ ChatGPT in het onderwijs: Waar is ChatGPT toe in staat en wat zijn de beperkingen?



Wat zijn de ethische overwegingen bij het gebruik van ChatGPT?

Risico's

Een flinke dosis gezond verstand en bewustzijn van de ethische tekortkomingen is erg belangrijk bij het gebruik van ChatGPT. Een keerzijde van generatieve AI is namelijk dat er ook belangrijke risico's bestaan in de toepassing ervan. Een van deze risico's heeft te maken met privacy; bij het gebruik van de gratis versie van ChatGPT wordt alles wat je invoert gebruikt als trainingsmateriaal. De kans bestaat dan ook dat bijvoorbeeld persoonsgegevens opzettelijk of onopzettelijk aan ander publiek dan bedoeld kunnen worden doorgegeven. Dit vormt niet alleen een risico voor individuele gebruikers, maar ook voor bedrijven of overheidsinstanties. Wanneer ChatGPT wordt gebruikt als belangrijk onderdeel van de dagelijkse activiteiten waarbij belangrijke of zelfs vertrouwelijke informatie wordt ingevoerd, komt de gegevensbeveiliging in gevaar⁷. Wees dus extra alert en voer geen persoonsgegevens in, zeker wanneer je ChatGPT vanuit een maatschappelijke organisatie gebruikt. Heb je daar wat meer hulp bij nodig? Onze Academy biedt AI-trainingen aan die je helpen bij het veilig leren gebruiken van AI.

De Autoriteit Persoonsgegevens benoemt daarnaast dat de trainingsdata van generatieve AI mogelijk onrecht verkregen zijn. Deze trainingsdata kunnen persoonlijke gegevens of auteursrechtelijk beschermd materiaal bevatten. Omdat het niet mogelijk is om die persoonsgegevens vervolgens te laten verwijderen, voldoet ChatGPT niet aan de privacywet AVG⁸. Wees je daarom bij het gebruik van ChatGPT en het invoeren van prompts bewust van deze privacyrisico's, en handel daarnaar. Voer geen persoonsgegevens in en stel jezelf de vraag: 'Zou ik deze gegevens openbaar op het internet delen?'. Als hierop het antwoord nee is, deel het dan niet met ChatGPT.

⁷ ChatGPT - Gevaarlijk generatief

⁸ Fui-Hoon Nah, F., Zheng, R., Cai, J., Siau, K., & Chen, L. (2023). Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration. *Journal of Information Technology Case and Application Research*, 25(3), 277-304.

Ook belangrijk om in het achterhoofd te houden is dat de huidige versie van ChatGPT getraind is op data tot en met september 2023. Het kan daarom voorkomen dat wanneer je informatie opvraagt over een onderwerp na 2023, je geen of foutieve informatie krijgt⁹. In feite is ChatGPT immers niet ontwikkeld als een up-to-date zoekmachine, maar meer als een modern taalmiddel.

Het feit dat ChatGPT zeer mensachtige antwoorden kan geven, heeft helaas ook een keerzijde. Ook wanneer de chatbot het antwoord op een vraag niet weet, zal het een antwoord genereren. Het kan dus een verhaal verzinnen, zonder aan te geven dat het eigenlijk geen idee heeft¹⁰. Dit verschijnsel noemen we hallucinatie. Een belangrijk gevaar van hallucinatie is de verspreiding van desinformatie. In de toekomst zullen algoritmen verbeterd moeten worden om hallucinaties te voorkomen¹¹.

Maatschappelijke uitdagingen

Naast risico's brengt generatieve AI ook (maatschappelijke) uitdagingen met zich mee. Zoals we eerder in dit whitepaper uitlegden, wordt ChatGPT getraind op enorme hoeveelheden data. In sommige gevallen is deze data niet representatief voor de diversiteit van de samenleving. De output die ChatGPT geeft kan vervolgens vooroordelen bevatten en discriminerend zijn¹². Open AI, de ontwikkelaar van ChatGPT, zet menselijke controle en filters in. Deze moeten de meest duidelijke onwaarheden en ergste vooroordelen eruit halen. Maar ook hiermee blijft dit een belangrijke beperking van ChatGPT. Onderzoek heeft aangetoond dat mensen geneigd zijn deze vooroordelen en onwaarheden over te nemen¹³. Er is hier dus sprake van een feedbackloop: de AI wordt niet alleen gevoed door menselijke vooroordelen, AI voedt ook de mens met vooroordelen¹⁴. Voor veel van onze opdrachtgevers die werken aan het versterken van inclusiviteit en representatie is dit dus een belangrijk risico.



Daarnaast kan het gebruik van generatieve AI een versterkend effect hebben op de zogenaamde digitale kloof, bijvoorbeeld als het gaat om internetvaardigheden. Voor mensen die het moeilijk vinden om een tool als ChatGPT te gebruiken, wordt er met de snelle technologische ontwikkelingen een grotere afstand gecreëerd tot het gebruik van digitale tools. Deze afstand kan ook ontstaan voor gebruikers uit minderheidsculturen. Wanneer blijkt dat generatieve AI te weinig getraind is op data met kenmerken van hun culturen, kunnen zij taal- of culturele barrières ervaren wanneer zij ChatGPT gebruiken¹⁵.

9 ChatGPT in het onderwijs: Waar is ChatGPT toe in staat en wat zijn de beperkingen?

10 Gevaarlijke feedbackloop: mensen nemen vooroordelen over van AI (die alles van de mens heeft afgekeken)

11 Generative AI and ChatGPT Applications challenges and AI-human collaboration (1).pdf

12 Ethiek en ChatGPT

13 Vicente, L., & Matute, H. (2023). Humans inherit artificial intelligence biases. *Scientific Reports*, 13(1), 15737.

14 Gevaarlijke feedbackloop: mensen nemen vooroordelen over van AI (die alles van de mens heeft afgekeken)

15 Fui-Hoon Nah, F., Zheng, R., Cai, J., Siau, K., & Chen, L. (2023). Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration. *Journal of Information Technology Case and Application Research*, 25(3), 277-304.

Hoe werk je samen met ChatGPT?

De mogelijkheden van ChatGPT zijn oneindig. Binnen Van Dam Datapartners hanteren wij de volgende categorieën:

1. Informatief:

Stel een vraag, en ChatGPT geeft antwoord. Dit kun je vergelijken met een opdracht geven aan Google. Waar Google bronnen oplevert waar je vervolgens zelf je informatie uit moet halen, is ChatGPT efficiënter, en produceert het een geheel geschreven antwoord. Een belangrijk nadeel hiervan is dat de informatie niet gemakkelijk te verifiëren is op waarheden en bronnen.

2. Schrijfhulp:

Je kunt zelf stukken tekst invoeren en ChatGPT vragen deze samen te vatten, te parafraseren, te vertalen of te controleren op grammatica. Ook kun je bepaalde tekst door ChatGPT in een bepaald format laten gieten, zoals een sollicitatiebrief.

3. Tekst genereren:

Je kunt ChatGPT vragen om een stuk tekst te schrijven, bijvoorbeeld een beschouwing over een bepaald onderwerp. Je kunt daarnaast aangeven welke lengte of schrijfstijl de tekst moet hebben.



Zoals eerder benoemd dienen prompts als het primaire communicatiemiddel tussen de gebruiker en ChatGPT. Ze begeleiden als het ware de chatbot bij het genereren van een reactie die aansluit op wat je van ChatGPT wilt ontvangen. De kwaliteit van deze prompts is van invloed op de kwaliteit van de gegenereerde antwoorden, en daarmee is het begrijpen van de nuances in het gebruik van prompts erg belangrijk voor het creëren van effectieve en betekenisvolle interactie met ChatGPT. Het systematische proces van het zorgvuldig ontwerpen van prompts om waardevolle output te genereren, wordt prompt engineering genoemd¹⁶.

Belangrijk bij het maken van prompts is anticiperen op hoe ChatGPT de gegeven prompts zal interpreteren. Dit houdt in: rekening houden met factoren zoals de trainingsdata, vooroordelen, en andere beperkingen die de interpretatie en mogelijkheden van generatie kunnen beïnvloeden. Hieronder lees je een aantal factoren die van invloed zijn op de prompts:

- **Intentie:** denk na over het doel van de interactie, of het nu gaat om het ophalen van informatie, het genereren van inhoud, of juist een probleemoplossend antwoord.
- **Specificeer het domein:** wanneer je een vraag stelt aan ChatGPT die gaat over een specifiek domein, probeer dan om domeinspecifieke woordenschat of woorden passend bij de context van dat domein te gebruiken. Door extra context of voorbeelden te geven, kan het model helpen om meer nauwkeurige en relevante resultaten te geven.

¹⁶ Ekin, S. (2023). Prompt engineering for ChatGPT: a quick guide to techniques, tips, and best practices. Authorea Preprints.

- Duidelijkheid en specificiteit: zorg ervoor dat de prompt duidelijk en specifiek is. Zo verklein je de kans op dubbelzinnigheid of verwarring.
- Beperkingen: bedenk of er bepaalde beperkingen zijn die je ChatGPT kan meegeven, bijvoorbeeld gewenste lengte van het antwoord (een maximaal aantal woorden of zinnen), of een specifiek format. Wanneer je vraagt om een samenvatting van een bepaald stuk tekst, geef aan in hoeveel zinnen je deze samenvatting wenst te ontvangen¹⁷.

Hieronder sommen we de tips voor het maken van goede prompts voor je op:

- Gebruik korte, begrijpelijke zinnen. Gebruik daarbij woorden die gemakkelijk te begrijpen zijn. Vermijd herhalingen, of overbodige woorden zoals "heel" of "eigenlijk".
- Geef ChatGPT duidelijke en nauwkeurige instructies
- Geef ChatGPT zoveel mogelijk context mee, in korte zinnen.
- Wees alert op spelling en grammatica. Fouten die gebruikers invoeren kunnen leiden tot fouten in ChatGPT.

Wat vraagt het van de mens om samen te werken met generatieve AI?

Vergeleken met traditionele zoekmachines zoals Google, vraagt het gebruik van generatieve AI om een andere benadering, maar ook om een andere mate van bewustzijn. Waar Google je meerdere informatiebronnen geeft, genereert ChatGPT een specifiek antwoord voor de prompt die je invoert. Bij een Google-opdracht ben je als het ware zelf nog aan de beurt om te beoordelen welke bron je gebruikt en hoe betrouwbaar deze bron is. Hoewel ChatGPT het voordeel heeft efficiënter te zijn, is het de kunst om antwoorden overnemen zonder verificatie, te vermijden. Overmatige afhankelijkheid van generatieve AI kan dan ook invloed hebben op vaardigheden zoals creativiteit en kritisch of probleemoplossend denken¹⁸.



Zogenaemde AI-geletterdheid, maar ook datageletterdheid, is dan ook van cruciaal belang voor gebruikers. Vertrouw niet blindelings op de antwoorden. Vergelijk de antwoorden met je eigen ervaring en domeinkennis. Denk ook goed na over wat je met je vraag aan ChatGPT wilt bereiken. Wil je informatie of handelingsperspectief? Verwerk dit ook in je vraagstelling. Wees je als laatste bewust van de manier waarop ChatGPT reageert op jouw prompt. Welke aannames maakt ChatGPT en kloppen deze? Door de afwegingen te maken en bewust stil te staan bij het proces, maak je jezelf bewust van jouw rol in prompt engineering.

17 Ekin, S. (2023). Prompt engineering for ChatGPT: a quick guide to techniques, tips, and best practices. Authorea Preprints.

18 Fui-Hoon Nah, F., Zheng, R., Cai, J., Siau, K., & Chen, L. (2023). Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration. *Journal of Information Technology Case and Application Research*, 25(3), 277-304.

Voor gebruik bij maatschappelijke organisaties geldt dat er sprake moet zijn van een mensgerichte benadering. Houd rekening met ethiek en transparantie en wees je te allen tijden bewust van de vertrouwelijke informatie die je als maatschappelijke organisatie bezit. Ook wanneer je ChatGPT vraagt om een samenvatting te maken van een beleidsstuk met vertrouwelijke informatie. Zet ChatGPT of andere vormen van generatieve AI vooral in voor algemene tekstuele opdrachten die geen vertrouwelijke informatie of persoonsgegevens bevatten. Bijvoorbeeld bij het schrijven van een algemene vacaturetekst. Gebruik ChatGPT dan ook vooral als sparringspartner en vermijd het doen van onderzoek via ChatGPT.



Zelf kennismaken met AI: de DataDropping

Ook bij Van Dam Datapartners houden we de ontwikkelingen van AI nauwlettend in de gaten en stellen we onszelf als doel mensen bewust te maken van de mogelijkheden en onmogelijkheden van data. Wil jij hier met jouw organisatie ook aan werken? En daarnaast oefenen met de potentie én gevaren van een AI-bot? Maak dan kennis met onze DataDropping, de serious game die jouw organisatie op een leuke en laagdrempelige manier laat kennismaken met data.

In de DataDropping combineren we een serious game op basis van jouw eigen data met een versterkend trainingsprogramma afgestemd op jouw leerdoelen. Stap op de fiets en ga met onze speciale app op avontuur in jouw omgeving. Je beantwoordt uitdagende datavraagstukken over jouw organisatie en werkt samen met collega's om je innerlijke datakracht naar buiten te laten komen. Zo kun je op een laagdrempelige en leuke manier kennismaken met data en datagedreven werken. Meer over de DataDropping lees je [hier](#).

Verder praten over AI?

Wij staan je graag te woord! Neem contact op met expert Conny via

conny@vandamdatapartners.nl





VAN DAM
DATAPARTNERS

