

WAAR MOET DE GEMEENTE OP LETTEN BIJ HET GEBRUIKEN VAN GENERATIEVE AI?



Voor u ligt een advies van de ethische commissie voor data en technologie van de gemeente Groningen. Deze onafhankelijke commissie adviseert het college van burgemeester en wethouders gevraagd en ongevraagd over ethische vraagstukken rondom het gebruik van data en technologie.

Dit advies is een gevraagd advies, ingediend door wethouder Broeksma. Het gaat over de vraag waar de gemeente op moet letten bij het inzetten van generatieve AI. In dit verslag leest u over de aanleiding, onze aanpak, de casus zelf en de beantwoording van de vraag die in de casus centraal staat.

Aanleiding

Het college heeft, mede op verzoek van de raad, de commissie om advies gevraagd bij het inzetten van generatieve AI.¹ Binnen de organisatie wordt hierover ook al nagedacht. Zo wordt er een visie geschreven op het inzetten van generatieve AI. Hierbij ligt de focus op Microsoft's Copilot.³ Deze digitale assistent kan integraal onderdeel zijn van een nieuw Officepakket dat door de gemeente wordt uitgerold: M365.³

Aanpak

Allereerst is de ambtelijke ondersteuning van de commissie in gesprek gegaan met wethouder Broeksma om de precieze adviesvraag te achterhalen. Deze vraag bleek bewust zeer breed ("Hoe benutten we de wenselijk aspecten van generatieve AI en hoe houden wij het onwenselijke tegen?"), omdat de wethouder de commissie de ruimte wilde geven zelf de belangrijkste morele overwegingen te schetsen. Een afvaardiging van de commissie is hiermee aan de slag gegaan, schriftelijk aangevuld door de overige commissieleden. Hierbij is ervoor gekozen voort te bouwen op al verschenen rapportages en zo concreet mogelijk naar de situatie van de gemeente te kijken. Daarom zijn we ook in gesprek gegaan met een aantal medewerkers die al bezig is met een visie op generatieve AI, in het bijzonder Microsoft's Copilot.

1 Generatieve AI is software die patronen in grote hoeveelheden data kan ontdekken. Hierdoor is het in staat om te voorspellen wat, op basis van die patronen, waarschijnlijk de meest gewenste uitkomst is bij een vraag of instructie. In de praktijk betekent dit dat je in eenvoudige taal instructies kunt geven waarna de software tekst, plaatjes of zelfs video's kan maken, bewerken of beschrijven. Bekendste voorbeeld is ChatGPT, van bedrijf OpenAI. Daarnaast kan de software allerlei typen informatie in andere informatie omzetten: van tekst naar afbeelding of van video's naar tekst.

2 Copilot is een vorm van generatieve AI van Microsoft en dient als digitale assistent, zowel voor particulieren als voor organisaties ('Enterprise'). De gemeente gebruikt alleen de extra beveiligde Enterprise-editie. Hier zijn twee varianten van: een nu al voor alle medewerkers beschikbare chatbot Copilot Chat en de 'premium'-variant Copilot M365 die geïntegreerd kan worden in Office-applicaties.

3 Op dit moment zijn de verschillende Office-applicaties binnen de gemeente Groningen niet met elkaar verbonden. Met Office M365 is dit wel het geval. Voor de transitie naar M365 is een meerjarig programma opgericht met daaronder allerlei projecten.

Ons informatiepakket vóór aanvang van de casusbespreking bestond daarmee uit de volgende onderdelen:

- Casusinbrengformulier: feitelijk een samenvatting van alle gevoerde gesprekken en informatie
- Conceptvisiestuk generatieve AI: een eerste visiestuk vanuit de organisatie over generatieve AI
- Rapport 'Verkenning ChatGPT: overwegingen voor verantwoord gebruik⁴ van de ethische commissie van het Interprovinciaal Overleg

Bovenstaande aanpak heeft tot de volgende hoofdvraag geleid voor de commissie:

“Welke overwegingen wil de ethische commissie meegeven wanneer de gemeente Groningen, in het kader van haar taken en gelet op haar voorbeeldfunctie, generatieve AI zoals Microsoft’s Copilot inzet?”

Casus

De casus bestaat eigenlijk uit twee onderdelen. Enerzijds is er de brede vraag van het college over waar de gemeente op moet letten bij het inzetten van generatieve AI binnen de gemeentelijke organisatie. Hierbij heeft het college ons uitgenodigd zelf de belangrijkste punten aan te dragen. Anderzijds zijn er binnen de organisatie al concrete ontwikkelingen op dit gebied waarover we kunnen adviseren. We hebben beide onderdelen opgepakt.

Begin 2023 is generatieve AI voor het eerst onder de aandacht gebracht binnen de gemeentelijke organisatie. Dit was vooral vanwege de opkomst van ChatGPT van ontwikkelaar OpenAI eind 2022. Er werd uitgelegd wat generatieve AI is en wat je er zoal mee kunt. In de maanden erna zijn ook de juridische risico's van deze software gecommuniceerd, gevolgd door een praatplaat ChatGPT in de zomer van 2023. In deze praatplaat zijn kansen en risico's van generatieve AI, in het bijzonder ChatGPT, gepresenteerd, gevolgd door een aantal richtlijnen. Zo zijn medewerkers er op gewezen nooit persoonsgegevens of strategisch gevoelige informatie te delen met deze software, transparant te zijn over het gebruik ervan en altijd zelf de verantwoordelijkheid voor het resultaat te houden. Ook is er een summer school ChatGPT georganiseerd en heeft de gemeente in september een betaalde account bij OpenAI geopend waarvan geïnteresseerde medewerkers (uiteindelijk ongeveer 40) de inloggegevens kregen om te kunnen experimenteren met ChatGPT.

⁴ Zie [Verkenning ChatGPT, overwegingen voor verantwoord gebruik](#)

In november 2023 is deze strategie om twee redenen aangepast. Allereerst kregen alle medewerkers van de gemeente als onderdeel van het Microsoft-pakket toegang tot Copilot Chat (voorheen Bing Chat Enterprise). Dit is een chatbot vergelijkbaar met ChatGPT⁵, maar dan inclusief zoekmachine Bing. Omdat het een onderdeel van het Microsoft-pakket is, is het in lijn met de beveiligings- en privacyrichtlijnen van de gemeente die ze met Microsoft contractueel heeft vastgelegd waarmee het een beter alternatief is dan het experimentele ChatGPT-account van OpenAI. Ten tweede hanteert de Rijksoverheid vanaf december 2023 de richtlijn geen ongecontracteerde generatieve AI te gebruiken, waar ChatGPT vanwege diens voorwaarden onder valt.⁶ Daarom heeft de gemeente een sterke voorkeur voor Copilot Chat uitgesproken en is de experimentele account van OpenAI eind december weer gesloten. De organisatie heeft het gebruiken van ChatGPT of andere generatieve AI echter niet expliciet verboden.

Medewerkers hebben dus momenteel toegang tot Copilot Chat en de bestaande richtlijnen die zijn opgesteld voor ChatGPT gelden nog steeds. Daarnaast heb je een tweede variant: Copilot M365. Deze variant kan worden geïntegreerd in alle Office-applicaties binnen de 365-omgeving. Ze kan allerlei taken voor je uitvoeren – van het opstellen en redigeren van teksten in Word tot analyses in Excel tot het maken van presentaties in Powerpoint. Copilot M365 kan ook al het werk binnen Office-toepassingen integreren. Zo is het bijvoorbeeld mogelijk om in Powerpoint een presentatie te maken op basis van inhoud die zojuist is geschreven in Word, inclusief het meenemen van de feedback die is ontvangen via mails in Outlook over dit specifieke onderwerp.

Copilot M365 is een premium-onderdeel van het nieuwe M365-Officepakket dat nog geïmplementeerd moet worden en waar een project voor is gestart. Hier is al een projectleider voor aangesteld. Het project bestaat uit meerdere fases, waarbij de eerste fase gericht is op het op orde brengen van de basis. Aangezien Copilot zo hot is, is de implementatie hiervan naar voren getrokken in het project. Er wordt nog een adoptiespecialist ingehuurd. Deze persoon zal gaan meedenken over hoe M365 het beste geïntroduceerd moet worden. De adoptiespecialist zal ondersteuning bieden bij het overdenken van onderwerpen zoals wat er van een medewerker verwacht mag worden en de wijze waarop een medewerker met de applicaties dient om te gaan. In januari 2024 is er binnen de organisatie begonnen met het schrijven van een bredere visie op generatieve AI.

5 Onder de motorkap van Copilot Chat zit ChatGPT-4, omdat Microsoft grootaandeelhouder van OpenAI is. Ze gebruikt dus de technologie achter ChatGPT voor haar eigen, beveiligde, toepassingen.

6 Zie kamerbrief [Voorlopig standpunt voor Rijksorganisaties bij het gebruik van generatieve AI](#)

In algemene zin stelt de conceptvisie dat de Gemeente Groningen:

- een moderne gemeente met een moderne bedrijfsvoering is;
- kansen ziet in toepassing van generatieve AI, zeker met oog op de verwachte uitstroom door pensioengerechtigden de komende jaren;
- verantwoord en veilig gebruik van generatieve AI stimuleert;
- generatieve AI als ondersteuning van de medewerker ziet.

Hierbij wordt gekozen voor Microsoft's Copilot als standaard, zowel de nu al beschikbare Chat als de M365-variant. De bovenstaande overgang naar M365 wordt gezien als een kans om efficiënter met Office-applicaties te werken, waarin Copilot een belangrijke rol kan spelen. Daarnaast wordt Copilot als veilig gezien vanwege de contractuele afspraken met Microsoft. Verder worden alternatieve vormen van generatieve AI, zoals ChatGPT, niet expliciet verboden, maar wordt vooral het gebruik van Copilot gestimuleerd. Reden hiervoor is dat er mogelijk nog andere interessante generatieve AI-toepassingen komen die de gemeente niet bij voorbaat wil verbieden. Daarnaast stelt men dat verbieden alleen zin heeft als je er ook op controleert en dit wordt zowel lastig als onwenselijk geacht.

Advies

Ons advies bestaat uit de volgende onderdelen: algemene bevindingen, verhouding casus tot het afwegingskader, juridisch kader, beantwoording van de hoofdvraag en concrete acties.

Algemene bevindingen

Allereerst vinden we het positief dat de gemeente zich bewust is van het belang van het innemen van een standpunt. Daarnaast vinden we het goed dat de gemeente openstaat voor verandering en bezig is met kaders en het opstellen van een visie rondom de inzet van generatieve AI. Dat daarbij vooralsnog wordt ingezet op een gecontracteerde vorm van generatieve AI in de vorm van Copilot, met veel beveiligings- en privacywaarborgen, wordt door de commissie gewaardeerd. We vinden het verder positief dat er externe expertise wordt aangetrokken om de gemeente hierbij te ondersteunen. Wel heeft de commissie het idee dat de gemeente onbewust nog grotendeels in een verkennende fase zit, terwijl er nu al stappen worden gezet die pas horen bij een uitvoerende fase.

Afwegingskader

Voor ons is minder goed te volgen hoe de visie rondom inzet van generatieve AI zich verhoudt tot het Groningse ethische afwegingskader.⁷ In dit afwegingskader staan een achttal waarden, waarvan een aantal schuren met de manier waarop nu naar generatieve AI gekeken lijkt te worden:

Nuchterheid: de gemeente laat zich niet meeslepen door hypes, kijkt nuchter naar wat er nodig is en durft nieuwe oplossingen kritisch te bevragen.
De gemeente is daarbij niet bang om dingen anders te doen.

De commissie vraagt zich af: is er voldoende voorbij de hype van generatieve AI gekeken? En is de basis, zoals databeheer, wel voldoende op orde om het potentieel van generatieve AI volledig te benutten?

Controle op technologie: Groningers mogen erop vertrouwen dat de gemeente grip heeft op de gebruikte technologie.

We vragen ons af of dat hier het geval is. Hoe zit het met de omkeerbaarheid van werkprocessen wanneer Copilot wordt geïntegreerd? Hoe ziet een exit-strategie eruit wanneer Copilot toch niet bevalt of Microsoft de kosten verhoogt? Kunnen medewerkers voldoende begrijpen wat Copilot precies doet wanneer ze assisteert – en daarmee eigenaarschap over hun werk houden? En hoe wordt bepaald wanneer ze wel of niet zou moeten worden ingezet?

Rechtvaardigheid: Groningers mogen erop vertrouwen dat de gemeente kiest voor technische oplossingen die voor iedereen heilzaam zijn.

De commissie vraagt zich af of er voldoende zicht is op hoe het gebruiken van Copilot de interactie met inwoners kan beïnvloeden, bijvoorbeeld wanneer het wordt ingezet als vertaler tussen inwoner en overheid. En is Copilot voor alle medewerkers toegankelijk genoeg, of versneld het juist de scheiding tussen digitale 'cans' en 'cannots'?

Duurzaamheid: de gemeente kiest voor een oplossing die gunstig is voor de leefomgeving en het klimaat.

7 Zie voor het Gronings ethische afwegingskader [Bijlage-1-Invulling-ethische-commissie-en-data-ethiek-wensen-en-bedenkingen.pdf \(groningen.nl\)](#), behandeld door de gemeenteraad in december 2022.

De commissie wijst er op dat het enorm veel energie kost om generatieve AI te trainen. Ook is er veel stroom nodig voor het beantwoorden van een prompt, evenals water om de servers te koelen. Is voldoende duidelijk hoe groot de ecologische impact van het gebruiken van generatieve AI zoals Copilot is? Wil de gemeente daar al dan niet consequenties aan verbinden? En hoe voorkomt de gemeente dat generatieve AI veel méér wordt ingezet dan misschien noodzakelijk is?

Privacy: *Groningers worden niet aangetast in hun persoonlijke levenssfeer.*

Deze waarde lijkt met het gebruik van Copilot wel goed geborgd. Maar in hoeverre wordt de privacy van Groningers goed gewaarborgd wanneer alternatieven voor Copilot (zoals ChatGPT, Claude of Bard) alsnog gebruikt mogen worden door medewerkers?

Juridisch kader

Daarnaast zien we ook nog een aantal juridische zaken waarover moet worden nagedacht. Zo kan de output van Copilot strijdig zijn met auteursrecht, omdat Copilot getraind is op data die mogelijk onwettig verkregen is. Bovendien kan de medewerker niet weten of de gegenereerde tekst grotendeels overeenstemt met al bestaande auteursrechtelijk beschermde teksten van derden. Ook het niet expliciet verbieden van ongecontracteerde vormen van generatieve AI, zoals ChatGPT, kan leiden tot datalekken en/of strijdigheid met de AVG. Ten slotte is een DPIA⁸ wenselijk voorafgaand aan de invoering van Copilot, omdat we het niet ondenkbaar achten dat er in sommige werkprocessen altijd een risico bestaat dat (gevoelige) persoonsgegevens (onbedoeld) kunnen worden ingevoerd.

Beantwoording hoofdvraag

Op basis van bovenstaande wil de commissie de gemeente daarom de volgende overwegingen meegeven:

- **De commissie mist een duidelijk doel voor de inzet van generatieve AI, de noodzaak voor dat doel en of Copilot het beste middel is om dat doel te bereiken.** Het is goed denkbaar dat hier wel degelijk antwoorden op te vinden zijn in bestaande visies en waardenkaders van de gemeente, maar dat is voor ons op dit moment niet helder. Na bestudering van de huidige stand van zaken krijgt de commissie daarom de indruk dat de gemeente eigenlijk nog in een verkennende fase zit, terwijl er al wel stappen worden gezet die horen bij een meer uitvoerende fase.
- **Er is meer verbinding nodig tussen de organisatievisie, het Groningse afwegingskader en de visie op generatieve AI.** Zoals is besproken onder de kop afwegingskader is het ons niet helder hoe de visie rondom generatieve AI zich verhoudt tot de Groningse waarden.

⁸ Data Protection Impact Assessment: wat zijn de risico's van het invoeren van persoonsgegevens bij een verwerkingsproces?

- **Neem op organisatieniveau een duidelijk standpunt in over de inzet van generatieve AI.** Op dit moment lijkt de verantwoordelijkheid voor het inzetten van generatieve AI vooral bij de medewerkers te liggen. De organisatie blijft echter primair aansprakelijk wanneer er iets misgaat - ook indien medewerkers zelfstandig gebruik zouden maken van alternatieven voor Copilot. Daarnaast kan er ook een normerende werking uitgaan van verbieden van (alternatieven van) generatieve AI. Waar hou je de medewerkers aan als er niets expliciet vastligt? Hoe kun je ze dan verantwoordelijk houden? Ook de mate waarin de gemeente hierin al dan niet een voorbeeldrol buiten de organisatie wil innemen is onduidelijk.
- **Er is een exit-strategie nodig bij het inzetten van generatieve AI.** Dit geldt in zowel abstracte als praktische zin. In abstracte zin gaat het om de implementatie van allerlei vormen van generatieve AI in werkprocessen – kun je het straks ook nog zonder en zo niet, is dat erg? In praktische zin gaat het om verdere afhankelijkheid van Microsoft – kun je straks nog zonder Copilot en zo niet, is dat erg?

Concrete acties

Op basis van bovenstaande overwegingen adviseren we daarom dat:

- noodzaak en doel van de inzet van Copilot beter in kaart gebracht wordt. Dit staat weliswaar impliciet in de conceptvisie die de commissie heeft mogen inzien, maar het moet duidelijker worden gerelateerd aan de missie en de taken van de gemeente als organisatie;
- er een projectplan komt, inclusief business case, voordat Copilot M365 wordt geïmplementeerd;
- er kaders worden opgesteld waaraan medewerkers zich geacht worden te houden;
- er een register wordt opgesteld waarin wordt bijgehouden in welke werkprocessen welke rol generatieve AI gaat spelen en/of speelt;
- er wordt nagedacht over een exit-strategie met betrekking tot de inzet van generatieve AI in het algemeen en Copilot in het bijzonder;
- de risico's voor het inzetten van generatieve AI per functiegroep van de gemeente in kaart worden gebracht en vastgelegd worden;
- dat de risico's per functiegroep goed worden gemonitord;
- er een DPIA wordt uitgevoerd rondom het inzetten van generatieve AI om daarmee de belangrijkste risico's in kaart te brengen;
- de gemeente anticipeert op de aankomende Europese AI Act.