

Inzet AI-assistent voor gespreksverslaglegging wijkteams Amersfoort

Op 20 maart kwam de Klankbordgroep Digitale Ethiek (KDE) bijeen om zich te buigen over een adviesaanvraag van de gemeente Amersfoort omtrent de inzet van een AI-assistent bij het maken van verslagen van de gesprekken die leden van de wijkteams voeren met (potentiële) cliënten. De aanvraag bestaat uit drie deelvragen:

- Hoe zorgen we ervoor dat de inwoner voldoende vertrouwen heeft om toestemming te geven om het gesprek te laten verwerken door een AI-assistent?
- In hoeverre ziet de klankbordgroep mogelijke ongewenste effecten door het gebruik van deze AI-assistent?
- Welke onderzoeksvragen zou de klankbordgroep graag meegenomen zien worden in de evaluatieonderzoeken?

De adviesaanvraag werd schriftelijk onderbouwd en tijdens de vergadering mondeling toegelicht door twee betrokken medewerkers van de gemeente en Wijkteam Amersfoort.

Overwegingen

De KDE weegt zwaar mee dat het beoogde AI-instrument wordt ingezet op één van meest gevoelige taken van de gemeente: de ondersteuning van kwetsbare groepen bewoners die mogelijk gebruik (willen) maken van de WMO of de jeugdzorg. De inzet van algoritmes die betrekking hebben op deze groepen is niet onomstreden; het uit de hand lopen van de toeslagenaffaire wordt voor een deel geweten aan de te kritiekloze inzet van AI bij de toekenning van toeslagen. En het is bekend dat de modellen die AI gebruikt, niet vrij zijn van vooroordelen, maar in de regel juist een afspiegeling zijn van de vooroordelen van de programmeurs of al schuil gaan in de data waarmee de modellen zijn getraind.

Dat impliceert dat toekomstig gebruik van AI-instrumenten, juist als het kwetsbare groepen betreft, extra kritisch moet worden bejegend. Die kritische houding geldt niet alleen de inzet van het instrument zelf, maar ook het beheer ervan, het eigenaarschap, de evaluatie en de context.

In dit kader wil de KDE de volgende overwegingen delen met de gemeente:

1. De beoogde aanbieder van de AI-tool is een betrekkelijk jong bedrijf zonder een duidelijk track record; de reputatie ervan staat of valt met de reputatie van de eigenaren en directie van deze onderneming. Dat is kwetsbaar; mocht het bedrijf worden verkocht, dan kunnen nieuwe eigenaren met andere intenties hun producten en diensten blijven leveren. Daarom is het van belang dat:
 - a. De gemeente een kortlopend contract heeft afgesloten, dat indien gewenst al na één jaar kan worden afgebroken;

- b. In dit contract ook bepalingen worden opgenomen die betrekking hebben op overdracht van de aanbieder aan een nieuwe eigenaar of een andere directie.
2. De aanbieder maakt gebruik van modellen die niet in Europa zijn ontwikkeld en slaat data bovendien op bij Amerikaanse cloud-providers, ook al is dit een in tijd gelimiteerde opslag. In de huidige discussies over digitale strategische autonomie (zie daarvoor bijv. [deze nota](#) van het ministerie van Economische Zaken) ligt het voor de hand met de aanbieder het gesprek te voeren over de inzet van in Europa ontwikkelde modellen (bijv. het Franse Mistral) en Europese clouddiensten.
3. Gesprekspartners van de wijkteams (de doelgroep) moeten nooit het gevoel krijgen dat hun medewerking een voorwaarde is voor het verkrijgen van gemeentelijke hulp. De KDE vindt dat een risico; het argument voor het gebruik van het instrument is immers een behoorlijke tijdswinst bij de verwerking van de gesprekken; dus medewerkers kunnen, al is het onbewust, toch druk uitoefenen om akkoord te gaan. Training van medewerkers en constante monitoring op het gebruik van het AI-instrument is daarom van groot belang. Dit is des te relevanter omdat gebruik van het instrument ook de inzet van een opnameapparaat impliceert; een vernieuwing ten opzichte van de huidige situatie, nu gesprekken niet worden opgenomen. Ook dit kan voor kwetsbare mensen een extra obstakel zijn om zich open en eerlijk uit te laten.
4. KDE vindt daarom dat er een goed plan moet komen om de inzet te evalueren; daarbij moeten niet alleen kwantitatieve data zoals tijdsbesparing het uitgangspunt zijn, maar moet ook zo diepgaand mogelijk onderzoek worden gedaan naar de beleving van de cliënten (de doelgroep).
5. KDE merkt op dat de interface niet uitnodigt om kritisch te kijken naar de door de AI gegenereerde samenvattingen. Het is voor een medewerker te gemakkelijk om routinematig de samenvatting goed te keuren. In de interface zou gezocht moeten worden naar manieren om aandachtig doorlezen van de samenvatting af te dwingen.
6. De nieuwe werkwijze mag op geen enkele manier de kans van cliënten op succesvol bezwaar tegen besluiten van de gemeente op basis van deze gespreksverslagen beïnvloeden.
7. De ervaring leert dat aanbieders van dit type instrumenten de neiging hebben om deze steeds meer aan te kleden, waarbij de afhankelijkheid vergroot wordt, het risico op vendor lock-in stijgt, en de prijs ook steeds verder kan stijgen. KDE merkt daarom op dat dit advies uitsluitend de nu voorgestelde versie geldt.
8. Onduidelijk is hoe 'inclusief' het voorgestelde instrument is. Werkt het ook goed als de cliënt een spraakgebrek of een zwaar accent heeft? En wat betekent het voor een cliënt als het instrument om die reden niet te gebruiken zou zijn? KDE vindt het van belang dat een cliënt het niet als iets negatiefs interpreteert wanneer diens stem niet geschikt blijkt voor het gebruik van de AI tool.

Advies

KDE komt tot het advies dat het instrument ingezet kan worden, mits rekening wordt gehouden met bovenstaande overwegingen. Specifieke antwoorden op de deelvragen:

- Alhoewel advies over het winnen van vertrouwen van de bewoner niet tot de *scope* behoort van KDE, denkt de klankbordgroep dat transparantie over de afwegingen van de gemeente en de wijkteams, en zorgvuldige communicatie omtrent bovenstaande overwegingen, belangrijke factoren zijn om het vertrouwen van bewoners te winnen.
- De mogelijke ongewenste effecten zijn bij de overwegingen in kaart gebracht:
 - (onbewuste) discriminatie door trainingsdata, taalmodellen of houding medewerker
 - Scope-drift en vendor lock-in
 - Verlies van de menselijke blik, creatie van te grote afhankelijkheid van AI
- Ten behoeve van de evaluatie beveelt de KDE aan om te onderzoeken of er standaard benchmarks zijn op basis waarvan dit type instrumenten geëvalueerd kan worden. Vooral dient daarbij onderzocht te worden wat de daadwerkelijke effecten zijn van het gebruik van het instrument op de kwaliteit van de gespreksvoering en van het verslag. Noot: non-verbale communicatie en menselijke intuïtie moeten niet uit het oog worden verloren als het gaat om de interpretatie van gesproken taal; twee aspecten die voor AI-instrumenten nog moeilijk zijn te interpreteren.