Datum	10-04-2025
Behandeld door	Ambtelijk secretaris
E-mail	cpa@amsterdam.nl

Onderwerp: Delen onderzoeksbevindingen en risicomaatregelen, en advies krijgen over ChatAmsterdam en toekomstige doorontwikkelingen.

Advies van de Commissie Persoonsgegevens Amsterdam (CPA) over delen onderzoeksbevindingen en risicomaatregelen, en advies krijgen over ChatAmsterdam en toekomstige doorontwikkelingen. Het agendapunt is behandeld op de openbare vergadering van 10-04-2025.

Geachte mevrouw 5.1, 2, e

Tijdens de vergadering van de CPA van 10-04-2025 was aanwezig bij het agendapunt "Delen onderzoeksbevindingen en risicomaatregelen, en advies krijgen over ChatAmsterdam en toekomstige doorontwikkelingen", tezamen met 5.1, 2, e 5.1, 2, e 5.1, 2, e en 5.1, 2, e .

Bijgaand treft u aan het advies van de CPA.

A. Samenvatting advies

De Commissie geeft u de volgende **adviezen** ter overweging mee:

- maak expliciet wat het doel is van het gebruik van ChatAmsterdam en expliciteer voor welke vergelijking wordt gekozen: ChatAmsterdam die is gebaseerd op commerciële OpenAI die op infrastructuur van Microsoft wordt gedraaid versus het gebruik commerciële GenAI versus het gebruik van OpenWeights modellen (zoals Llama);
- als het gebruik van ChatAmsterdam op termijn kan resulteren in afname van vaardigheden en kennis van ambtenaren, dan is het vergroten van AI-geletterdheid niet of slechts een deel van de oplossing; een betere oplossing kan zijn om medewerkers te trainen in specifieke kennis en vaardigheden waarmee zij beter in staat zijn te controleren wat ChatAmsterdam produceert;
- stel voorafgaand aan een volgende evaluatie op basis van vooraf vastgestelde normen vast wanneer het gebruik van ChatAmsterdam geslaagd is. Lever vervolgens bij de uitrol maatwerk in het gebruik (door middel van specifieke doelstellingen) van ChatAmsterdam aan de verschillende afdelingen binnen de gemeente;
- stel vast en leg vast welke lijn wordt gevolgd voor de communicatie naar de burger over het gebruik van GenAI in het algemeen en ChatAmsterdam in het bijzonder;
- naast het door de gemeente geconstateerde risico op ongeoorloofde verwerking van (bijzondere) persoonsgegevens bij de invoer in ChatAmsterdam, is er ook een risico op ongeoorloofde verwerking van (bijzondere) persoonsgegevens bij de uitvoer van ChatAmsterdam;
- neem maatregelen om risico's als gevolg van bias, het gebrek aan accuraatheid van data, de verminderde betekenisvolle menselijke tussenkomst, het gebrek aan duurzaamheid bij het gebruik van AI en de afwezigheid van inzicht in de trainingsdata zoveel mogelijk te mitigeren.
- voor wat betreft de verwerkingsgrondslag gerechtvaardigde belangen van de verwerkingsverantwoordelijke adviseert de CPA de gemeente gemotiveerd te onderbouwen waarom zij zou vinden dat de inzet ChatAmsterdam valt onder de *ontwikkelingsfase* van een AI-model. Indien de gemeente dit onvoldoende kan onderbouwen, acht de CPA de kans klein

dat de gemeente succesvol een beroep kan doen op de grondslag gerechtvaardigd belang. In dat geval dient voor de toepassing van ChatAmsterdam voor de grondslag en een eventuele uitzondering op het verwerkingsverbod voor bijzondere persoonsgegevens te worden aangesloten bij de publieke taak waarvoor ChatAmsterdam wordt ingezet.

De Commissie:

- wenst te vernemen op welke wijze de gemeente Amsterdam gevolg heeft gegeven aan haar adviezen.

B. Aanleiding en omschrijving

Doel van het team ChatAmsterdam

Het primaire doel van Team ChatAmsterdam is het ontwikkelen en aanbieden van ChatAmsterdam als privacyvriendelijker en veiliger alternatief ten opzichte van commerciële aanbieders van AI-chatbots, zoals ChatGPT. Alle verwerkingen van persoonsgegevens binnen het proces dienen volgens de gemeente dit primaire doel. Met ChatAmsterdam wil het Team ChatAmsterdam leren hoe AI de gemeente kan helpen bij haar werkzaamheden, hoe zij het verantwoord gebruik van AI stimuleren, en de AI-geletterdheid binnen de organisatie bevorderen. Deze doelstelling wordt door Team ChatAmsterdam in samenwerking met P&O nader uitgewerkt.

Daarnaast hebben de gebruikers van ChatAmsterdam een grote verscheidenheid aan doelen waarvoor ze ChatAmsterdam willen gebruiken. Het is volgens de gemeente onmogelijk om alle doelen vooraf in beeld te brengen omdat het onmogelijk is om exact te voorspellen waarvoor medewerkers ChatAmsterdam gaan gebruiken.

Het team heeft in de onderzoeksfase de meest relevante gebruikersdoelen geïdentificeerd en opgenomen in de DPIA. De top 3 bestaat uit:

1. Tekstverwerking (opstellen, samenvatten en vertalen) – 47%
2. Informatie vergaren (voornamelijk interne informatie) – 18%
3. Overige (creatieve ideeën, data-analyse etc.) – 35%

Toelichting door Team ChatAmsterdam

Hieronder volgt een toelichting op de voorgenomen verwerking van ChatAmsterdam

Verwerkingsgrondslag

- Inloggegevens van gebruikers met een Amsterdam-account binnen de gemeente die inloggen bij ChatAmsterdam.

Het authentiseren van gebruikers en het opslaan van zijn de verwerkingen die gebaseerd kunnen worden op de grondslag algemeen belang (artikel 6, eerste lid sub e van de Algemene Verordening Gegevensbescherming (AVG). Artikelen 5, eerste lid sub f jo. 24, eerste lid jo 32, eerste lid van de AVG, schrijven voor dat verwerkingsverantwoordelijken passende technische en organisatorische maatregelen dienen te nemen om persoonsgegevens op passende wijze te beschermen.

- Persoonsgegevens die in de prompts door gebruikers worden ingevoerd (dat is nadrukkelijk niet de bedoeling).

Team ChatAmsterdam onderzoekt de mogelijkheid of er sprake is van een "gerechtvaardigd belang" bij het ontwikkelen van de ChatAmsterdam applicatie, zoals beschreven in artikel 6, eerste lid, sub f AVG. Dit artikel stelt dat verwerking van persoonsgegevens rechtmatig is als deze noodzakelijk is voor de behartiging van de gerechtvaardigde belangen van de verwerkingsverantwoordelijke of een derde, tenzij de belangen of de fundamentele rechten en vrijheden van de betrokkene zwaarder wegen. Dit is de eerste keer dat de toepassing van de grondslag "gerechtvaardigd belang" voor deze specifieke situatie wordt besproken.

Er zijn argumenten die zowel voor als tegen de proportionaliteit van de verwerking kunnen worden aangevoerd. Enerzijds is er sprake van een uitzonderlijk en incidenteel geval waarbij de gemeente een privacy-vriendelijker en veiliger alternatief voor commerciële AI-oplossingen beoogt te bieden. Dit belang wordt ondersteund door de EDPB-richtlijnen, die het gebruik van deze verwerkingsgrondslag in uitzonderlijke gevallen toestaan. Het ontwikkelen van een dienst zoals een chatbot om gebruikers te ondersteunen, zoals genoemd in de recente opinie 28/2024 van de European Data Protection Board (EDPB), kan onder omstandigheden als een gerechtvaardigd belang worden beschouwd.

Anderzijds zijn er zorgen over de aard en context van de verwerkingen, zoals de mogelijke verwerking van gevoelige persoonsgegevens en de schaal van de verwerking, hetgeen de impact op de betrokkenen vergroot. Bovendien kunnen betrokkenen de verwerkingen niet redelijkerwijs verwachten, wat de kans vergroot dat hun belangen zwaarder wegen in de belangenafweging.

Hoewel beide punten de toepassing van "gerechtvaardigd belang" ondersteunen, hebben we nog geen definitieve keuze gemaakt. Om een weloverwogen beslissing te maken, is advies van de CPA nodig. Dit advies kan helpen bij het bepalen of de grondslag "gerechtvaardigd belang" passend is voor de ontwikkeling van de ChatAmsterdam applicatie.

Verzoek aan de CPA

1. De CPA wordt gevraagd kennis te nemen van de bevindingen van het onderzoek naar het gebruik van ChatAmsterdam. Dit onderzoek is uitgevoerd met inachtneming van de suggesties van de CPA uit de vergaderingen van 29 februari 2024, 23 mei 2024 en 17 oktober 2024.
2. De CPA wordt om advies gevraagd over risico's die mogelijk over het hoofd zijn gezien, en waarmee rekening moet worden gehouden bij de toekomstige doorontwikkeling van ChatAmsterdam.
3. De CPA wordt gevraagd om advies te geven over de grondslag "gerechtvaardigd belang" (AVG, artikel 6, eerste lid, sub f. Na dit getoetst te hebben en alle mogelijke maatregelen te hebben genomen, blijft er een restrisico over. Hoe gaan we om met dit restrisico? Wat zijn de vervolgstappen die we hiervoor moeten nemen?

C. Advies

De CPA dankt de bij dit adviesverzoek betrokken medewerkers van de gemeente voor hun toelichting en de aangeleverde stukken.

Doel gebruik ChatAmsterdam

ChatAmsterdam is in 2024 drie keer besproken met de CPA. De eerste keer ging het om een pilot-onderzoek met een beperkte groep ambtenaren, de tweede keer om een onderzoek onder 1000 medewerkers van de gemeente. De derde keer in oktober is gekeken naar de bevindingen en adviezen. Tijdens de sessie van 10 april 2025 staat het gebruik van ChatAmsterdam door een pilot groep ambtenaren in november 2024 tot april 2025 centraal.

De gemeente geeft aan ChatAmsterdam in te willen zetten als privacyvriendelijker en veiliger alternatief voor commerciële aanbieders van AI-chatbots, zoals ChatGPT.

De CPA ziet het type AI dat de gemeente gebruikt – wel of niet als commerciëel te duiden – graag nader gepreciseerd. Dan komen ook de risico's van het gekozen type beter in beeld. ChatAmsterdam is gebaseerd op het commerciële OpenAI model, dat op infrastructuur van Microsoft wordt gedraad. Het commerciële perspectief is met de huidige opzet aldus niet geheel uitgesloten. ChatAmsterdam maakt immers gebruik van de API van Microsoft om gebruik te kunnen maken van het model van OpenAI. De prompts die ambtenaren invoeren worden verwerkt door een commerciële aanbieder, in dit geval Microsoft. OpenAI levert in dit geval enkel het model, en zal in de gegeven situatie waarschijnlijk geen persoonsgegevens verwerken. Als het gaat om het onderzoeken van privacyvriendelijker alternatieven zou ook het gebruik van Open Source modellen (zoals wellicht Llama) in de afweging meegenomen moeten worden.

Daarnaast is nog niet duidelijk wat de gemeente met 'veiliger' bedoelt. Gaat het hier om bijvoorbeeld de impact op gebruikers (de betrokken ambtenaren) en/of andere betrokkenen, of eerder om het risico op een inbreuk op auteursrechtelijk beschermd materiaal? De afweging om een ander model te gebruiken met het oog op een 'veiliger alternatief' vanwege betere prestaties kan bijvoorbeeld sneller inbreuk maken op eigendomsrechten. Daarmee rijst de vraag dan of het model dan echt veiliger is.

De CPA adviseert, kortom, om in de doelstelling van de gemeente een helder en completer onderscheid te maken tussen ChatAmsterdam en bestaande modellen. De CPA adviseert de gemeente daarnaast om scherp te stellen welk doel ChatAmsterdam heeft, en wat zij precies verstaat onder een veiligere toepassing: voor wie is de toepassing veiliger? De CPA adviseert in kaart te brengen tegen welke andere belangen die hogere veiligheid wordt afgewogen.

Rol ChatAmsterdam in de werkzaamheden van ambtenaren

De gemeente geeft aan dat ChatAmsterdam interne processen kan verbeteren en de efficiëntie van het werk vergroot. ChatAmsterdam heeft volgens de gemeente de potentie om de manier waarop de gemeente werkt te transformeren door tijd van ambtenaren te besparen, en de bespaarde tijd te besteden aan complexere taken.

De CPA wil benadrukken dat er een risico is dat ambtenaren door het gebruik van ChatAmsterdam afhankelijk worden van ChatAmsterdam, en hun taken niet meer zonder ChatAmsterdam kunnen uitvoeren. Hier heeft het uitgangspunt te gelden, dat de ambtenaar indachtig de verplichting van betekenisvolle menselijke tussenkomst altijd zelf in staat moet blijven om zelf de output van ChatAmsterdam te kunnen controleren. Ook het eenvoudiger maken van bijvoorbeeld ingewikkelde juridische teksten kan bovendien zeer complex zijn omdat de (vereenvoudigde) inhoud en de

betekenis van de output van dezelfde juridische precisie en kwaliteit dienen te zijn. Om dit te controleren zal betekenisvolle menselijke tussenkomst steeds nodig zijn.

Voor zover ChatAmsterdam wordt gebruikt om taken van ambtenaren over te nemen, kan de introductie van ChatAmsterdam op middellange termijn impact hebben op de vaardigheden en kennis van ambtenaren. Als het probleem is dat de vaardigheden en kennis van ambtenaren afnemen, dan is het vergroten van AI-geletterdheid niet of slechts een deel van de oplossing. Een betere oplossing kan zijn om medewerkers te trainen in specifieke kennis en vaardigheden waarmee zij in staat worden gesteld te controleren wat ChatAmsterdam produceert. Uitgangspunt hierbij kan zijn dat ChatAmsterdam weliswaar sneller zal werken dan een ambtenaar, maar dat wil niet zeggen dat de uitkomsten van ChatAmsterdam beter zijn dan die van de ambtenaar.

Evaluatie van ChatAmsterdam

De gemeente stelde dat zij middels intern kwantitatief en kwalitatief onderzoek het gebruik van ChatAmsterdam tijdens de pilot heeft onderzocht. De CPA heeft echter geen onderzoekscriteria gezien waaraan de uitkomsten van het onderzoek zijn geijkt. De CPA adviseert voor de volgende keer een zinvoller evaluatie te organiseren, waarbij vooraf duidelijk voorwaarden worden beschreven en vastgesteld om te bepalen of en wanneer ChatAmsterdam als applicatie geslaagd is. Hiervoor adviseert de CPA ook gebruik te maken van kennis die buiten de gemeente aanwezig is om te beoordelen in hoeverre ChatAmsterdam voldoet aan de door de gemeente gestelde eisen (zoals TADA-waarden), en aan de wettelijke vereisten zoals de AVG en de AI-Act.

De CPA adviseert de gemeente bij de uitrol van ChatAmsterdam per sector, en per directie en/of afdeling maatwerk te leveren. Dit is van belang in het licht van de doelbinding: als het doel van de gegevensverwerking vaag blijft, kan een doelbindingsprobleem onder artikel 5, eerste lid, sub b van de AVG ontstaan. Daarnaast kunnen in het kader van de AI-Act problemen ontstaan omdat het onbegrensd uitrollen van de applicatie al gauw in een hoog-risico categorie zou kunnen vallen. Dit kan worden verbeterd door te achterhalen en vervolgens zo specifiek mogelijk vast te stellen voor welke doeleinden en processen de diverse afdelingen ChatAmsterdam hebben gebruikt of gaan gebruiken. Zo zullen betrokken directies en afdelingen binnen de gemeente elk hun eigen doelstellingen en verwerkingswijzen voor ogen hebben waarvoor ChatAmsterdam gebruikt kan worden. Door directies zo specifiek mogelijk te laten duiden waarvoor ChatAmsterdam zal worden ingezet, en daarvoor maatwerk te leveren in de functionaliteiten en toepassingen, kunnen mogelijke risico's al vroeg in beeld komen, en waar nodig worden gemitigeerd.

Alternatieve modellen

De gemeente geeft aan dat de ChatAmsterdam-applicatie is ontworpen om het wisselen tussen modellen makkelijk te maken. Omdat de prompts op dit moment naar een API van Microsoft worden gestuurd en het model afkomstig is van een commerciële aanbieder (OpenAI), beveelt de CPA aan te onderzoeken of het gebruik van open source modellen mogelijk is. Dit onderzoek zou kunnen worden uitgevoerd door naar de output van verschillende modellen te kijken die is gebaseerd op de prompts die de ambtenaren hebben ingevoerd.

Communicatie met burger over gebruik Generatieve AI

De CPA adviseert de gemeente afstemming binnen de gemeente te organiseren over de communicatie met de burger over het gebruik van Generatieve AI in het algemeen en ChatAmsterdam in het bijzonder. Kiest de gemeente voor transparantie over het gebruik van ChatAmsterdam in het opstellen van brieven naar de burger, dan moet de gemeente een afweging

maken in welke mate en op welke moment de gemeente de burger informeert over de rol die generatieve AI heeft gespeeld in het proces.

Verwerking van bijzondere persoonsgegevens zonder uitzondering op het verwerkingsverbod

Volgens de door de gemeente aangeleverde beschrijving is er sprake van een verwerking van bijzondere persoonsgegevens ex art. 9, eerste lid, van de AVG. Deze verwerking vindt plaats als gebruikers bijzondere persoonsgegevens invoeren in de prompt. Er zijn maatregelen genomen om dit risico te mitigeren, maar de gemeente geeft aan dat ondanks de genomen maatregelen het risico op dergelijke verwerkingen zal blijven bestaan. Daarnaast ziet de CPA ook het risico dat er bij de uitvoer, dus bij het antwoord dat ChatAmsterdam geeft, alsnog bijzondere persoonsgegevens kunnen worden verwerkt en gedeeld. De CPA attendeert de gemeente erop dat er naast het geconstateerde risico van de verwerking van persoonsgegevens in de invoer, aldus ook een risico is op een ongeoorloofde verwerking van persoonsgegevens bij de uitvoer.

Met het gebruik van ChatAmsterdam bestaat aldus een reëel risico op het verwerken van bijzondere persoonsgegevens. Om het risico op ongeoorloofde verwerking van bijzondere persoonsgegevens verder te beperken en mitigerende maatregelen effectief te maken in het specifieke geval, is van belang dat de specifieke doeleinden waarvoor het model wordt gebruikt op de juiste wijze in kaart worden gebracht. Ook hier is van belang dat het risico op verwerking van bijzondere persoonsgegevens per directie en per sector kan verschillen. Per directie of sector kan dan ook worden beoordeeld of er mogelijk een uitzonderingsgrond op het verwerkingsverbod aanwezig is voor de specifieke gevallen.

Bewaartermijn van de prompts en monitoring

De gemeente geeft aan dat de prompts 30 dagen worden bewaard. De gemeente heeft gemotiveerd dat deze termijn nodig om is om verkeerd gebruik te kunnen traceren.

Vanwege het risico op onrechtmatige verwerking van (bijzondere) persoonsgegevens adviseert de CPA te onderzoeken of het gebruik van ChatAmsterdam continue kan worden gemonitord. Bij dit onderzoek dient te worden betrokken dat het monitoren als invasief kan worden gezien, omdat het prompts zijn die ambtenaren zelf invullen. Een mogelijke rechtvaardiging daarvoor kan zijn dat het de bedoeling is dat ambtenaren geen persoonsgegevens invoeren en het gegevens zijn die enkel te maken zouden moeten hebben met de taakuitvoering van de gemeente. Het is bovendien de enige manier om de filters die worden gebruikt, up-to-date te houden. Op deze manier kan meer structureel verkeerd gebruik van de applicatie snel worden opgespoord en kunnen filters en andere maatregelen adequaat worden bijgesteld.

Advies CPA over grondslag "gerechtvaardigd belang"

De gemeente gaf aan behoefte te hebben aan beantwoording van de vraag of de grondslag gerechtvaardigd belang van de verwerkingsverantwoordelijke ex art. 6, eerste lid, sub f AVG kan gelden als grondslag voor de verwerking van persoonsgegevens bij het ontwikkelen van de ChatAmsterdam applicatie. De gemeente stelde in haar aanvraag dat zij nog geen definitieve keuze heeft gemaakt met betrekking tot deze verwerkingsgrondslag. De CPA vraagt zich evenwel af op basis van welke verwerkingsgrondslag in de AVG de applicatie dan onlangs live is gegaan?

De AVG verbiedt het gebruik van de grondslag 'gerechtvaardigd belang' door overheidsinstanties in het kader van de uitvoering van hun taken. De vraag rijst of de verwerking van persoonsgegevens van de betrokken ambtenaren met het oog op het gebruik van de ChatAmsterdam applicatie valt onder de uitvoering van gemeentelijke taken. Indien dat het geval is dan valt het gebruik onder de

verwerkingsgrondslag en eventuele uitzondering op het verwerkingsverbod die zijn verbonden aan die taak. Dit pleit voor een aanpak en beoordeling per directie of sector. Indien de gemeente kan aantonen dat de persoonsgegevens niet wordt verwerkt voor de vervulling van haar taken, kan de zij overwegen de grondslag gerechtvaardigd belang te gebruiken voor het gebruik van ChatAmsterdam.

Echter, het gevolg hiervan is dat een beroep op een uitzondering op het verwerkingsverbod bij het verwerken van bijzondere persoonsgegevens wordt beperkt of uitgesloten. De uitzonderingen waar de gemeente zich op kan beroepen zijn verbonden aan de uitoefening van een publieke taak. Daarmee kan dan geen beroep meer worden gedaan op de grondslag gerechtvaardigd belang. Oftewel, als gekozen wordt voor de grondslag gerechtvaardigd belang, dan is er geen uitzondering mogelijk op eventuele verwerkingsverboden voor het verwerken van bijzondere persoonsgegevens.

Indien desondanks een beroep op het gerechtvaardigd belang wordt overwogen, is voor het juridische kader de opinie van de European Data Protection Board (EDPB) relevant. De EDPB maakt in zijn Opinie 28/2024 een onderscheid tussen het gebruik van deze rechtvaardigingsgrond voor de verwerking van persoonsgegevens in de *ontwikkelingsfase* van een AI-model ('development phase'), en het gebruik van deze grond voor de verwerking van persoonsgegevens in de fase waarin het AI-model daadwerkelijk wordt *gebruikt* ('deployment phase').[1] Bij het adviesverzoek kijkt de CPA specifiek naar het gebruik van de wettelijke grondslag met het oog op de verwerking van persoonsgegevens ten behoeve van het *gebruik* (de deployment phase).

Voor zover de gemeente kan aantonen dat de verwerking enkel plaatsvindt in het kader van de ontwikkelfase ten behoeve van de training van het AI-model, kan deze onder een gerechtvaardigd belang van de verwerkingsverantwoordelijke worden geschaard. In dat geval dient de gemeente (als verwerkingsverantwoordelijke) vervolgens deze stappen te nemen: de gemeente moet aantonen dat (i) de verwerking *noodzakelijk* is voor (ii) de behartiging van haar *gerechtvaardigd belang*; en dat (iii) de *belangen de grondrechten van de betrokkenen* (in dit geval persoonsgegevens van gemeenteambtenaren) die tot bescherming nopen, *niet* zwaarder wegen dan de belangen van de verwerkingsverantwoordelijke.

De CPA acht het niet waarschijnlijk dat de gemeente Amsterdam met de inzet van ChatAmsterdam valt onder de criteria voor de *ontwikkelingsfase* van een AI-model. De gemeente ontwikkelt in dit geval niet zelf een model; zij koopt toegang tot een API die haar output uit het model teruggeeft. Zij ontwikkelt het model dus niet zelf maar zet het in voor haar eigen doeleinden binnen de ChatAmsterdam-applicatie. De CPA adviseert de gemeente gemotiveerd te onderbouwen waarom zij alsnog zou vinden dat zij met ChatAmsterdam valt onder de *ontwikkelingsfase* van een AI-model.

Indien de gemeente dit onvoldoende kan onderbouwen, dan acht de CPA de kans klein dat de gemeente met succesvol een beroep kan doen op de grondslag gerechtvaardigd belang. In dat geval dient voor de toepassing van ChatAmsterdam voor de grondslag en een eventuele uitzondering op het verwerkingsverbod voor bijzondere persoonsgegevens te worden aangesloten bij de publieke taak waarvoor ChatAmsterdam wordt ingezet. Hierbij dient dan niet alleen rekening te worden gehouden met de ingevoerde data, maar ook met de opgeleverde resultaten van de prompt.

Tot slot

Als laatste punt brengt de Commissie in herinnering dat de rol van de CPA adviserend is, en niet instemmend. Het verzoek aan de Commissie om 'in te stemmen' of 'positief te adviseren' over het voornemen past niet bij de rol en taak van de CPA als adviesorgaan.

[1] EDPB, Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models (17 december 2024).

Voor het overige verwijst de CPA u naar bijgevoegde notulen van de vergadering over het betreffende agendapunt.

Hoogachtend,

Mevrouw H. Janssen
Voorzitter Commissie Persoonsgegevens Amsterdam

Uittreksel van het verslag van "delen onderzoeksbevindingen en risicomaatregelen, en advies krijgen over ChatAmsterdam en toekomstige doorontwikkelingen" agendapunt van de vergadering op 10 april 2025

ChatAmsterdam
Dit project is in 2024 al driemaal aan de orde geweest in de CPA. De eerste keer ging het om een pilot-onderzoek met een beperkte groep ambtenaren, de tweede keer om een onderzoek onder 1000 medewerkers van de gemeente. In oktober is gekeken naar de bevindingen en adviezen. Nu gaat het erom te bepalen welke kansen en risico's dit eigen Amsterdamse alternatief voor commerciële generatieve AI biedt. Alles is in eigen beheer ontwikkeld, rekening houdend met de TADA-waarden, autonomie en (cyber)veiligheid. De data worden na encryptie bewaard op eigen servers binnen Europa en de gemeente heeft controle op de promptfilters. Data gaan dus niet naar Open AI. Het streven is taalmodellen te gebruiken die passen bij de waarden die de gemeente hanteert en digitale onafhankelijkheid en transparantie versterken. Dit alles om ambtenaren in staat te stellen op een verantwoorde manier met generatieve AI te werken ter ondersteuning van hun werk en repetitieve taken. Via maatregelen als een privacyverklaring, gebruiksvoorwaarden, werkinstructies, bevordering van AI-geletterdheid, prompt- en contentfilters, toegangsbeveiliging, netwerkbeveiliging en grenzen aan de omvang van de invoer van stukken wil men de risico's zo veel mogelijk mitigeren.

Van november tot april is onderzocht hoe en waarvoor ambtenaren Amsterdam Chat gebruiken. Het gebruik is toegenomen en vindt vooral plaats voor tekstverwerking, informatieverzameling en creatieve toepassingen als data-analyse. In 4% van de gevallen hielden de filters een vraag tegen. En kreeg de aanvrager een waarschuwing. De indruk is dat men beter en efficiënter wekt en dat de waarborgen voor privacy en veiligheid worden gewaardeerd. De volgende stap is nu het vragen van toestemming aan de directies van diensten om dit verder uit te rollen en ontwikkelen en het dan in augustus beschikbaar te maken. Verder wil men gaan werken aan de ontwikkeling van een eigen Europese LLM (taalmodel).

De CPA is benieuwd waar het kantelpunt ligt om al dan niet verder te gaan met deze chatbot.

Dit lijkt een goed moment en daarom is men nu ook bij de CPA. De vraag is dan of doorgaan een goed idee is als men het afzet tegen wat er gebeurt als de gemeente geen stappen zet. De generatieve AI maakt een razendsnelle ontwikkeling door en de verwachting is dat ambtenaren AI hoe dan ook gaan gebruiken. Daarom is het van belang na te gaan hoe dat op een verantwoorde manier kan.

De CPA adviseert het dan ook te vergelijken met een handmatige manier van werken en met andere applicaties. Er bestaat ook het risico dat de ambtenaren te afhankelijk worden van IT. Het is nodig dat de vaardigheid om te controleren of een tekst klopt behouden blijft. Dat mag men niet aan machines overlaten. De ervaring wijst uit dat men meer op IT gaat leunen als de druk tot efficiency toeneemt.

Dat vindt men een goede vraag die de nodige ethische discussies zal oproepen. Op zich kan men nu al stellen dat de Amsterdam Chat veiliger is dan Chat GT. De ontwikkeling valt niet tegen te houden en dat brengt het risico met zich mee dat ambtenaren er niet goed mee weten om te gaan. Daarom is het ook belangrijk te werken aan IT-geletterdheid. Men moet zich realiseren dat het een hulpmiddel

is en dat het nodig blijft zelf na te denken en keuzes te maken. Het is geen middel om aanvragen af te handelen, maar wel een hulpmiddel om bijvoorbeeld ambtelijke taal beter begrijpelijk te maken. Daar moet dan nog wel een keer een controle door een ter zake kundige medewerker overheen om te controleren of te tekst inhoudelijk nog klopt. Dit systeem is niet bedoeld als vervanger, maar als versneller van het proces. Efficiency is niet het hoofddoel.

De CPA merkt op dat de techniek mogelijkheden biedt op kennis op te schalen die de medewerker op dat moment nog niet heeft. Uit het onderzoek blijkt dat 20% van de prompts gaat om het achterhalen van informatie en kennis. De chatbot zal dan geneigd zijn de gevraagde antwoorden te geven en de medewerker kan vervolgens niet beoordelen of die informatie welk klopt. Dat kan op den duur een negatief effect hebben op de vaardigheden die de ambtenaren nog zelf in huis hebben.

De projectmedewerkers zijn zich bewust van dit gevaar. Daarom is in het onderzoek gekeken waar medewerkers naar zoeken als ze de volledige vrijheid hebben. Ze zoeken vooral naar interne informatie over beleidsstukken. Het is een feit dat de techniek de mogelijkheid biedt om veel vlugger heel veel informatie te verzamelen. Daarom is het van belang het in de organisatie goed te laten doordringen dat het ongebreideld vergaren van informatie moet worden voorkomen. Deze technologie moet men alleen gebruiken voor kleine stukjes code. Het moet beheersbaar blijven en de organisatie moet ook voortdurend vragen blijven stellen over het eigen functioneren en nadenkt over mogelijke toekomstscenario's. Niemand kan alles controleren en men moet niet uit het oog verliezen dat ook mensen verkeerde afwegingen kunnen maken.

De CPA vraagt af het uitrollen na de pilot over de hele organisatie sluipenderwijs gaat, zoals dat vaker gebeurt bij de overheid, of dat er nog een kantelpunt voorzien wordt. Of er bijvoorbeeld een evaluatie voorzien is aan de hand van vooraf bepaalde criteria.

Het onderzoek was daarvoor deels ook bedoeld om zicht te krijgen op hoe medewerkers ermee omgaan. De resultaten vindt men bemoedigend genoeg om nu door te gaan. Verder werkt men binnen de gemeente samen met GRIP dat deelneemt aan diverse gemeentelijke samenwerkingsverbanden. Er zijn dus externe partijen die meekijken. Er is nog geen evaluatie vastgesteld. Maar het is zeker niet de bedoeling in een keer alle sluizen naar de organisatie open te zetten. Ten eerste moeten directies zich aanmelden voor deelname en zullen per directie andere maatregelen getroffen worden die aansluiten bij de daar gehanteerde werkprocessen. Het is daarom belangrijk via gesprekken met domeinexperts te achterhalen voor welke doeleinden en processen men de chat wil gaan gebruiken.

De CPA adviseert dan ook een zekere regie te nemen op de uitvoering van die processen om te voorkomen dat managers vooral gaan sturen op efficiency en besparing. Verder is de CPA benieuwd of ook transparant naar de burgers gemeld wordt dat de chatbox is ingezet voor het bewerken van een brief.

De verwachting is dat in haast elke applicatie op termijn een AI-component zal komen. In het project Heldere Taal wordt in een breder perspectief gekeken naar de begrijpelijkheid van gemeentelijke berichten. Daar kan dan meteen ook gekeken worden of de inzet van AI gemeld dient te worden. Een andere mogelijkheid zou het werken met een watermerk kunnen zijn.

De CPA vraagt of bepaalde tools van Microsoft voor deze chat met een opt-out zijn uitgeschakeld.

Dat is voor het cruciale deel het geval.

De CPA is benieuwd of inmiddels wel aangegeven kan worden hoeveel stroomverbruik er met Chat Amsterdam gemoeid is.

Dat blijkt nog steeds lastiger te bepalen dan men zou vermoeden vanwege de snelle ontwikkelingen. Er zijn wel gesprekken gaande om dit in beeld te krijgen. Het wordt waarschijnlijk als een van de variabelen meegenomen bij de keuze van een LMM.

De CPA adviseert bij de uitvoer goed te kijken naar binnengeslopen persoonsgegevens. Bij samenvattingen komt het bijvoorbeeld vaker voor dat er per ongeluk persoonsgegevens worden opgenomen.

De voorzitter concludeert dat de discussie nu binnen de gestelde tijd niet kan worden afgerond. Zij kondigt een advies aan waarin ook ingegaan zal worden op de openstaande vragen over risico's en gerechtvaardigd belang.