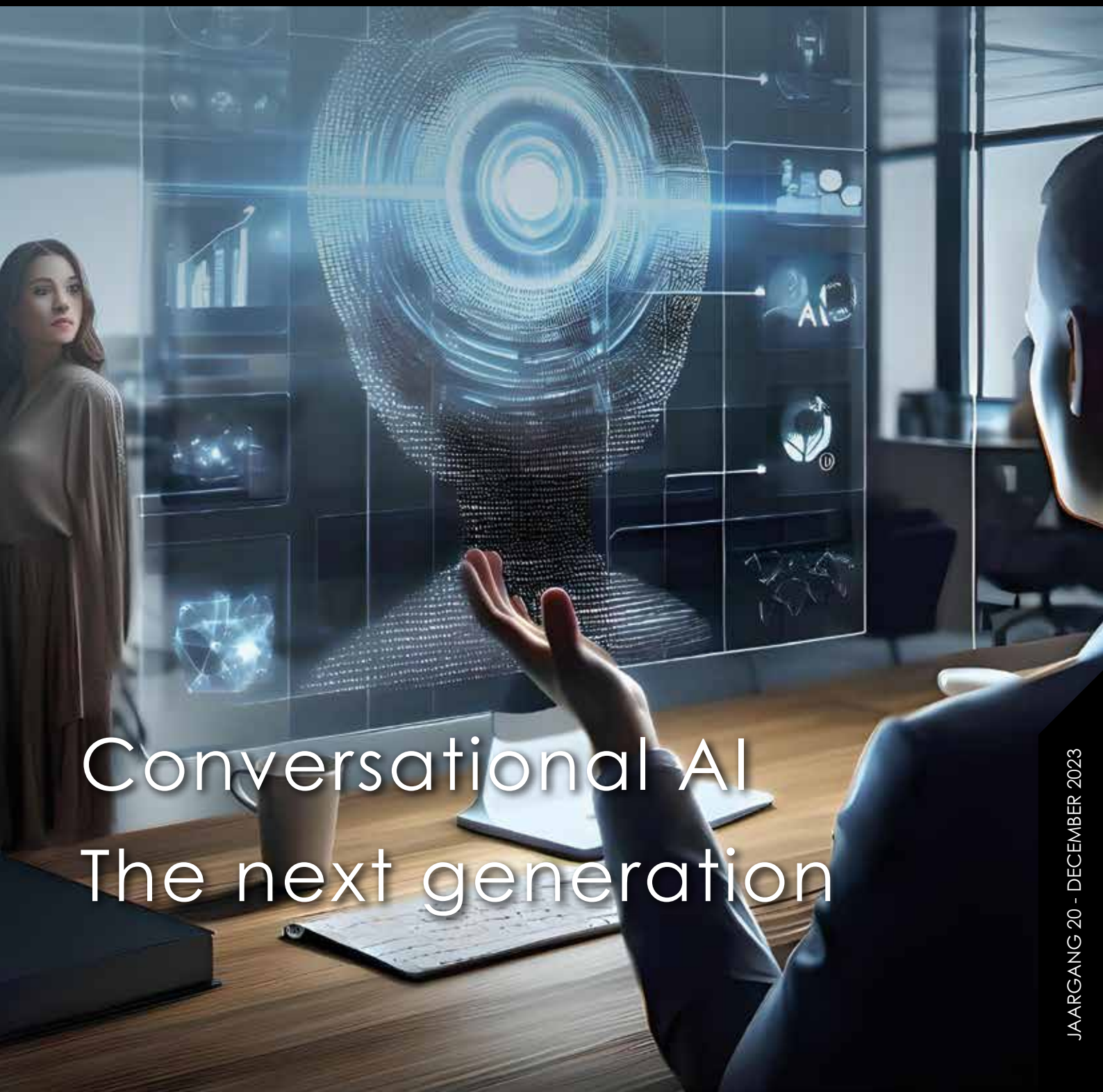
DIXIT

20 JAAR DIXIT!

TIJDSCHRIFT OVER TAAL- EN SPRAAKTECHNOLOGIE



Conversational AI
The next generation

JAARGANG 20 - DECEMBER 2023

INHOUD

Ruud Verhoeven, developer in code & design en vormgever van de DIXIT heeft concept en beeld van de voorkant met behulp van AI geproduceerd en bewerkt.

VOORWOORD

Beste lezer,

Met gepaste trots presenteer ik deze speciale editie van DIXIT, waarmee we tevens het twintigjarig jubileum van ons gewaardeerde tijdschrift vieren. Twintig jaar geleden zette DIXIT zich in om de ontwikkelingen op het gebied van taal- en spraaktechnologie in Nederland op de voet te volgen en mede vorm te geven. Anno 2023 kunnen we met voldoening terugkijken op een periode van enorme innovatie, waarin natural language processing en conversational AI enorm zijn gevorderd.

Toch vraag ik me af of de angstwekkende voorspelling van Musk, "AI zal banen vernietigen", mijn toekomst zal bepalen. Zal mijn voorwoord nog relevant zijn in een wereld gedomineerd door taalmodellen? Of is deze tekst zelf al geschreven door zo'n slimme machine?

Aan de ene kant hebben deze modellen inderdaad laten zien tot verrassend coherente en menselijk klinkende teksten in staat te zijn. Maar aan de andere kant is een voorwoord meer dan alleen correct Nederlands.

Het gaat ook om het overbrengen van een persoonlijk gevoel en een specifieke boodschap. Kunnen taalmodellen dat even goed als een mens? Ik betwijfel het sterk. Een echt goed voorwoord - één dat de lezer raakt, nieuwsgierig maakt, tot nadenken stemt - vereist een menselijke creativiteit en sensitiviteit die AI voorlopig ontbeert.

Bovendien draait een voorwoord om meer dan de tekst alleen. Het schept een band tussen auteur en lezer. Als lezer wil je een gevoel krijgen bij de persoon achter de woorden. Dat 'menselijke element' is lastig te vangen voor een computermodel.

Hoewel de technologie snel vordert, verwacht ik niet dat menselijke auteurs binnen afzienbare tijd overbodig zullen worden, in weerwil van sombere voorspellingen. Integendeel, juist vanwege de opmars van AI zal het persoonlijke, menselijke aspect van teksten als voorwoorden alleen maar belangrijker worden.

Wijze woorden of ijdele hoop? Het is wellicht nog te vroeg om daarover te oordelen. Daarom blijft DIXIT het voortouw nemen in het volgen van de nieuwste ontwikkelingen op het gebied van taal- en spraaktechnologie.

Ik wens u veel leesplezier met deze speciale jubileumeditie!

Fabrice Nauze,
voorzitter Stichting NOTaS

DIXIT: Tijdschrift over toegepaste taal- en spraaktechnologie – 20e jaargang, editie Conversational AI. The next generation. DIXIT is een uitgave van Stichting NOTaS, Toernooiveld 100, 6525 EC Nijmegen. Tel. 024-3512108– E-mail info@notas.nl – www.notas.nl **Redactieadres:** Stichting NOTaS, Toernooiveld 300, 6525 EC Nijmegen **Redactie:** Arjan van Hessen: a.j.vanhessen@utwente.nl; Henk van den Heuvel: h.vandenheuvel@let.ru.nl; Marieke den Os (secretariaat NOTaS): info@notas.nl. **Hoofdredactie themanummer:** Carla Verwijmeren: carla@y.digital **Advertenties:** Stichting NOTaS – info@notas.nl, 024-3512108 Abonnementen: Voor een gratis abonnement kunt u zich wenden tot een van de NOTaS-deelnemers **Druk:** BredaPress **Verantwoording:** DIXIT is een uitgave van Stichting NOTaS. Overname van de artikelen is alleen toegestaan met bronvermelding en na toestemming van Stichting NOTaS. Stichting NOTaS en de bij deze uitgave betrokken redactie en medewerkers aanvaarden geen aansprakelijkheid voor mogelijke gevolgen die zouden kunnen voortvloeien uit het gebruik van de in deze uitgave opgenomen informatie.

WE ZIJN OP WEG, MAAR HET IS NOG LANG NIET GOED GENOEG

Welkom bij de 20ste editie van DIXIT, die deze keer volledig gewijd is aan de nieuwste ontwikkelingen op het gebied van Conversational AI. De vaste lezers onder u denken wellicht 'hadden we twee jaar geleden niet ook een editie over Conversational AI?' En dan heeft u het juist! Maar de ontwikkelingen zijn in het afgelopen jaar zo hard gegaan, dat we er niet omheen kunnen om het thema dit jaar opnieuw centraal te stellen.

Door:
Carla Verwijmeren,
CCO - Y.digital

Het is natuurlijk een enorme eer om voor deze bijzondere jubileumeditie de gastredacteur te mogen zijn. Het is weer een prachtig inhoudelijk magazine geworden, waarbij we er met het hele redactieteam naar hebben gestreefd om de meest relevante perspectieven samen te brengen. Van ervaringen van consumenten en de laatste wetenschappelijke inzichten, tot de meest actuele technologische ontwikkelingen en bijbehorende juridische en ethische kaders.

Het voorrecht van de gastredacteur is dat ik hier ook mijn persoonlijke beschouwingen met u kan delen. Hierbij heb ik mij laten inspireren door de laatste reclamecampagne van zorgverzekeraar OHRA die zich sterk afzet tegen chatbots. "Artificial intelligence speelt een steeds grotere rol in ons dagelijks leven, maar daardoor krijgen we ook steeds vaker te maken met de eindeloze, frustrerende gesprekken met zogenaamd slimme chatbots," aldus de verzekeraar. En dat is helaas heel vaak herkenbaar.

Nederlandse consumenten niet tevreden over de kwaliteit van chatbots

Uit de publicatie van de laatste Nationale Voice Monitor – een onderzoek onder meer dan 1.000 Nederlandse consumenten – blijkt dat maar liefst 90% van de respondenten ontevreden is over de kwaliteit van chatbots. Tegelijkertijd blijkt uit de kwalitatieve interviews dat organisaties de technologie niet meer kunnen wegdenken uit hun klantcontactstrategie. En dit is pure noodzaak, want service op een acceptabel niveau houden in tijden van vergrijzing en een sterk krimpende arbeidsmarkt vraagt om niet-conventionele oplossingen. De introductie van Large Language Models en toepassingen als ChatGPT luiden hiermee een nieuw tijdperk in.

Adoptie van nieuwe technologie gaat in sneltreinvaart

In februari 2023 publiceerde de Zwitserse Bank UBS dat ChatGPT de snelst groeiende app ooit is. Volgens analisten van de bank haalde de app in een recordsnelheid een gebruikersaantal van meer dan 100 miljoen, waarmee het partijen als TikTok en Instagram ver achter zich liet. Microsoft was als investeerder al betrokken bij de ontwikkelingen binnen OpenAI, de partij achter ChatGPT. De tech-gigant maakte in januari 2023 bekend de investeringen verder op te voeren met een meerjarige investering van vele miljarden. Dit alles leidde tot grote zorgen bij Google en Meta, die inmiddels ook hun eigen modellen (Bard respectievelijk LlaMa) hebben geïntroduceerd. En

vandaag de dag is het aantal nieuwe modellen op de markt nauwelijks meer bij te benen.

Belangrijke vertragende factoren voor adoptie door publieke en private sector

Hoe veelbelovend de nieuwe technologie ook is, er spelen steeds meer ethische vraagstukken bij de inzet van de Large Language Models, waardoor ze niet zomaar als end-to-end oplossingen voor chatbots kunnen worden ingezet. Een bekend probleem betreft hallucinaties, ofwel output die semantisch en syntactisch correct is maar los staat van de werkelijkheid. Andere zorgpunten zijn transparantie, de bescherming van de privacy, data security en de ecologische footprint.

Daarnaast roept de kwaliteit van de data waarmee de modellen zijn getraind veel vragen op. Want bestaande vooroordelen in de dataset en een onevenredige representatie van doelgroepen versterken de kans op discriminatie. Zo hebben we bijvoorbeeld gezien dat wanneer het gaat over de rol van manager, sommige LLM's al snel refereren aan mannen.

Adoptie door criminelen neemt exponentieel toe

Tegelijkertijd spelen er ook andere grote dilemma's rondom het gebruik van de modellen. Hoe kun je misbruik door criminelen - bijvoorbeeld het voeren van nep gesprekken met gekloonde stemmen - voorkomen? Dit soort contacten worden met inzet van de technologie steeds geloofwaardiger, waardoor ze nauwelijks meer van echt zijn te onderscheiden. Een alarmerend initiatief op dit vlak is het op cybercriminele fora veelbesproken WormGPT. Een variant die voortkomt uit het in 2021 gelanceerde, open source language model GPT-J. Maar dan zonder de ethische beperkingen die de grote tech partijen inmiddels gelukkig wel steeds beter toepassen.

Wet- en regelgeving

Er ligt uiteindelijk een grote verantwoordelijkheid bij de partijen die de modellen ontwikkelen. Microsoft, Google, Anthropic en Meta zetten in op zelfregulering met de oprichting van het Frontier Model Forum. Maar zelfregulering in een industrie waar de commerciële belangen (GPT 3.5 Turbo kost $ 0,002 en GPT 4 al $ 0,06 per 1.000 tokens!) en maatschappelijke risico's zo groot zijn, vraagt ook om externe regulering en controle. Begin 2024 wordt naar verwachting de allereerste uitgebreide AI-wet ter wereld aangenomen, de Europese AI Act. Deze wet beperkt het gebruik van AI, algoritmes en machine learning en beoogt de risico's voor mens en maatschappij te minimaliseren. Na goedkeuring van deze wet komt er waarschijnlijk een overgangsperiode van twee jaar, voordat de wet volledig van kracht wordt.

Transformatie van Conversational AI

Er is, zoals gezegd, een revolutie gaande in het speelveld van Conversational AI. Maar met bovenstaande ontwikkelingen in het achterhoofd, is er ook nog veel onduidelijk. Bovendien zien we dat organisaties die net miljoenen euro's hebben geïnvesteerd in de conventionele Conversational AI deze niet zomaar overboord gooien. Dit biedt een enorme kans voor organisaties die tot op heden de technologie maar beperkt hebben ingezet, want zij kunnen zonder legacy de stap maken naar de nieuwe wereld. En ik durf hierbij de voorspelling te doen dat ook OHRA deze beweging uiteindelijk gaat maken. Wellicht primair met een strategie die gericht is op de interne ondersteuning van de medewerkers (want hiervoor biedt de nieuwe technologie ook heel veel mogelijkheden), maar het kan niet uitblijven dat ook customer service er in de toekomst aan moet geloven. De campagne ten spijt…

Tot slot

Ik las ter inspiratie voor mijn inleiding de titel van de allereerste DIXIT - uit 2003 - nog eens terug: 'Spraaktechnologie, de puberale jaren voorbij,' met als subtitel 'Van herkenning naar interpretatie.' Ik twijfel er niet aan dat we destijds al baanbrekende stappen hebben gezet, maar tegelijkertijd is het op dit moment nog altijd niet goed genoeg om een meerderheid van consumenten tevreden te stellen. Zo lang slechts 10% van de consumenten tevreden is over de assistentie van Conversational AI-systemen, is het aan alle betrokkenen - wetenschap, leveranciers én publieke en private organisaties - om kritisch te blijven en gedreven om het iedere dag weer beter te doen. Ik hoop dat deze DIXIT u hiervoor veel handvatten biedt! •

CONVERSATIONAL AI: LLMS ARE NOT ALL YOU NEED

Het in 2017 verschenen paper 'Attention is all you need'[1] heeft, op zijn zachtst gezegd, veel reuring teweeg gebracht in de wereld van de taaltechnologie. De baanbrekende *Transformer* Deep-Learning modellen die daarin beschreven werden, hebben als bekendste verschijningsvorm de Large Language Models (LLM's). Deze modellen kunnen enorme hoeveelheden taalgegevens verwerken en daaruit nieuwe informatie distilleren. Zij herkennen complexe taalpatronen en kunnen deze generaliseren, waardoor ze in staat zijn om nieuwe taaluitingen te voorspellen. Maar zijn ze ook op zichzelf voldoende voor een Conversational AI systeem dat aan de verwachtingen van gebruikers voldoet?

Door:
Alexandra Redmann, AI Engineer en
Ian FitzPatrick, CTO
bij Y.digital

Op basis van LLM technologie lanceerde tech-startup OpenAI een serie GPT-modellen waaronder het inmiddels alom bekende ChatGPT. Het voeren van een menselijk lijkende chatconversatie met een AI was daarmee voor veel mensen een feit. Het lijkt haast triviaal om dan ook te stellen dat LLM's krachtige tools zijn voor Conversational AI.

Het belang van context

Wanneer gebruikers interacteren met Conversational AI systemen, of het nu chat of spraak is, vinden ze het cruciaal om begrepen te worden.[2] Ze verwachten dat het systeem onmiddellijk en correct interpreteert wat er wordt gezegd en op de juiste manier reageert. Om aan deze verwachtingen te voldoen, moet het Conversational AI systeem in ieder geval voorzien worden van de juiste informatie om het gespreksdoel (van de conversatiepartner) te kunnen bereiken.

Het is bijvoorbeeld essentieel om rekening te houden met de context van het gesprek, dat wil zeggen alle informatie die misschien niet direct benoemd wordt, maar wel bekend is over de conversatiepartner en de bredere situatie. Het Conversational AI systeem moet rekening houden met kennis van de wereld en feitenkennis om correcte en volledige responsen te kunnen geven. En het moet conclusies kunnen trekken en nieuwe inzichten kunnen afleiden uit de combinatie van wat er in het gesprek gezegd is en overige beschikbare achtergrondkennis. Naast dit alles moet het in staat zijn om natuurlijke taal te begrijpen en taal te genereren die vloeiend, grammaticaal correct en consistent is in zijn tone-of-voice. AI met al is het een ambitieuze lijst met eisen die we stellen aan een succesvol Conversational AI systeem. We kunnen onszelf de vraag stellen: "Welke middelen heeft een LLM om zelfstandig aan deze eisen te voldoen?"

Context meegeven aan een LLM

Interacties met LLM's verlopen via *prompts*: stukjes tekst met instructies en context die de statistische taalgeneratie van de LLM de gewenste kant opstuurt. Deze ogenschijnlijk simpele techniek kan veelzijdige uitkomsten teweeg brengen.[3] Een van de meest innovatieve prompt methoden is het zogenoemde Reason+Act (ReAct) paradigma, een manier van prompts construeren waarmee je de LLM probeert te verleiden tot 'redeneren'[4] en vervolgens 'handelen'.[5] Een veelgebruikte toepassing daarvan

[1] https://arxiv.org/abs/1706.03762
[2] https://pages.y.digital/nationale-voice-monitor
[3] Zie bijdrage 'Large Language Models: Geen silver bullet', Dixit 2022 (https://notas.nl/dixits) voor voorbeelden
[4] LLMs zijn statistische taalgeneratoren en redeneren dus niet in de menselijke zin
Zie verder https://aclanthology.org/2023.findings-acl.67.pdf en https://dl.acm.org/doi/10.1145/3442188.3445922
[5] https://arxiv.org/abs/2210.03629

is Retrieval Augmented Generation (RAG), een techniek waarbij de LLM 'redeneert' over hoe het een gebruikersvraag kan beantwoorden en een zoekvraag (voor bijvoorbeeld een zoekmachine) genereert. Als tussenstap voert een stukje computercode de zoekvraag daadwerkelijk uit en koppelt de resultaten terug aan de LLM die daarmee (hopelijk) alle relevante informatie heeft om de vraag te beantwoorden (zie Box 1 voor een illustratie). Omdat de LLM met deze techniek aangevuld wordt met een stukje computercode betekent ook dat elk computer toegankelijk systeem in principe hiermee toegankelijk gemaakt kan worden voor de LLM, dus ook bijvoorbeeld een Customer Relationship Management (CRM) systeem waarin klantgegevens kunnen worden opgehaald voor extra context over de gesprekspartner van de Conversational AI.

Dat alle voor de conversatie benodigde informatie (bijvoorbeeld de voorgaande gesprekscontext) en instructies via de prompt moet verlopen, levert ook een belangrijke beperking. Het aantal *tokens* (woorddelen) dat een LLM tegelijk kan verwerken (de *context window*) heeft een harde bovenlimiet. Voor ChatGPT ligt die bijvoorbeeld op 4096 tokens, maar er zijn ook voorbeelden van LLM's met een veel grotere token limiet zoals Claude van AI-startup Anthropic, die naar verluidt 100.000 tokens tegelijkertijd aankan.[6] Hoe deze token limieten de komende tijd zullen veranderen is moeilijk te voorspellen, maar feit is wel dat ze uiteindelijk gehouden zijn om binnen beschikbare rekenkracht (met name GPUs) in datacentra te draaien, en dat met grotere schaal, meer rekenkracht (en dus hogere gebruikskosten), energie-[7] en waterverbruik[8], of trade-offs met responstijden gemoeid zullen zijn.

Context in het menselijk taalverwerkingssysteem

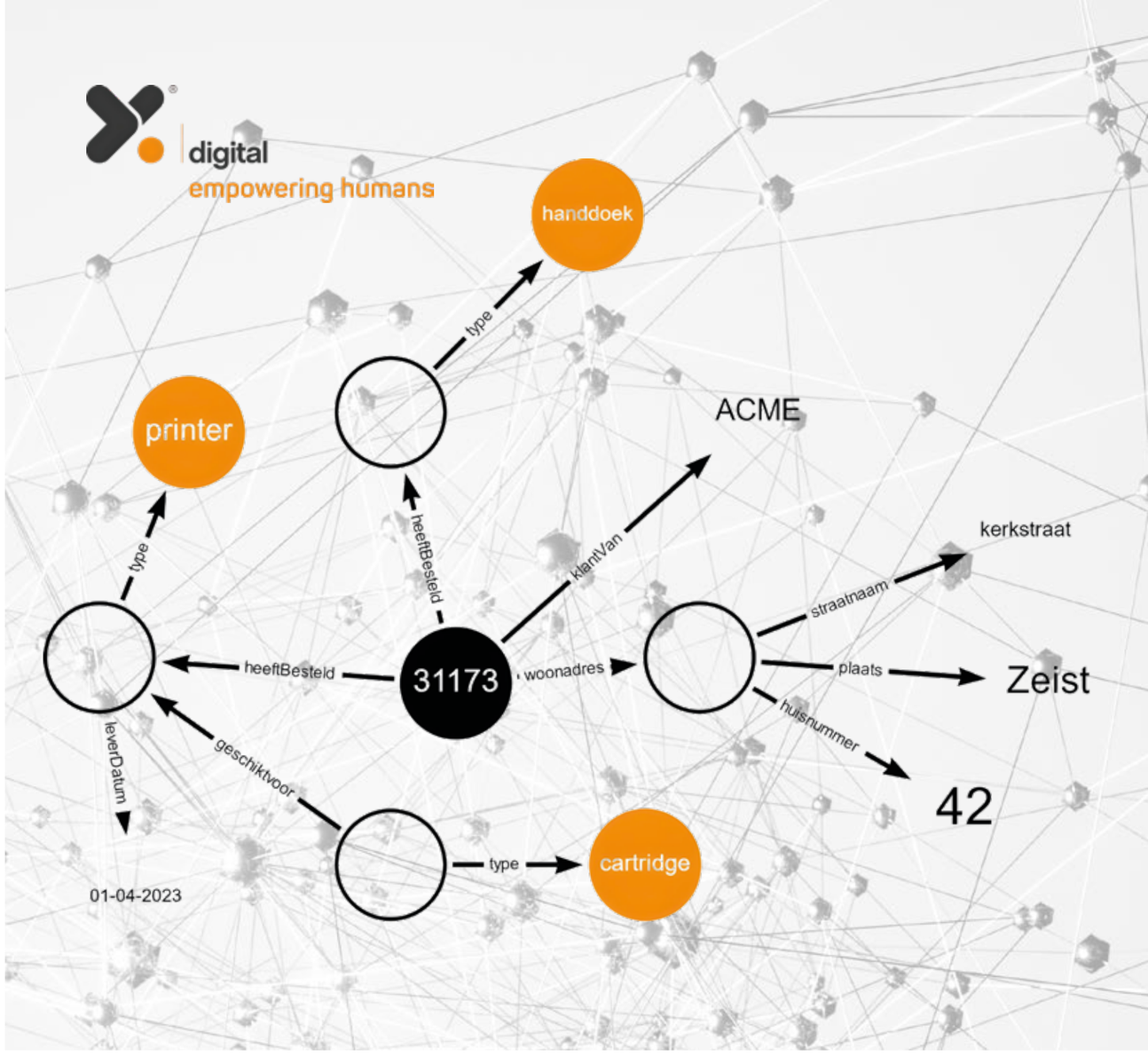
Het ReAct paradigma heeft ook de eigenaardigheid dat dit het taalgeneratiesysteem (de LLM) in de regiestoel zet: de LLM 'redeneert' en regisseert vervolgens de zoekmachine. Hoe anders is dit in het menselijk taalverwerkingssysteem waarin taalgeneratie en redeneren functioneel gescheiden lijken te zijn (voor een andere kijk hierop zie 'The Language of Thought' door J. Fodor[9]). Ons brein kan daarnaast vrijwel direct beschikken over veelzijdige kennis (biografisch, semantisch, episodisch, pragmatisch) die gedurende jaren is opgebouwd door middel van interactie met een wereld die een LLM slechts via de beperkte lens van het geschreven woord (of beter gezegd: een numerieke representatie daarvan) kan zien. De cognitieve neurowetenschappen hebben ons in de afgelopen 100+ jaar laten zien dat ons taal-en-cognitiesysteem geen monoliet is, maar een complexe interactie tussen modulaire systemen met ieder een eigen specialisme. Actuele context wordt bijvoorbeeld bijgehouden in ons werkgeheugen, maar we kunnen ook een beroep doen op ons lange termijn geheugen voor informatie over gebeurtenissen uit het verleden of voor feitenkennis uit ons semantisch geheugen.

Dit alles doet de vraag rijzen: "Overvragen we de LLM niet door het zoveel verschillende cognitieve petten op te zetten?". Of specifieker: "Is het realistisch om te verwachten dat we de LLM van voldoende context kunnen voorzien binnen de beperkingen van zijn *context window*?". Als we van onze Conversational AI steeds meer menselijkheid en nauwkeurigheid gaan verlangen, lijkt dit onwaarschijnlijk en zal hoe dan ook de rek van de *context window* steeds verder op de proef worden gesteld. Er is kort gezegd behoefte aan een manier om de LLM te ontlasten in het verwerken van heel veel context, door het te voorzien van gerichte *kennis*.

Context middels Knowledge Graphs

Een veelgebruikt middel om kennis weer te geven is de zogenaamde *Knowledge Graph*: een soort netwerk dat concepten, hun eigenschappen en de relaties daartussen op een gestructureerde manier modelleert. In een *Knowledge Graph* kan zowel generieke als domein specifieke kennis worden weergegeven (zie Box 2). Deze vorm van kennisrepresentatie leent zich uitstekend voor (automatisch) redeneren en om nieuwe kennis af te leiden op basis van de beschikbare informatie.[10] Het model kan, bovendien, voortdurend worden uitgebreid of aangepast als er nieuwe gegevens beschikbaar komen.

Er zijn verschillende manieren denkbaar om *Knowledge Graphs* en LLM's te laten samenwerken.[11] Zo kunnen LLM's (maar ook andere AI technieken) worden gebruikt om de constructie van de *Knowledge Graphs* te vergemakkelijken en kan kennis uit *Knowlege Graphs* (statisch en/of beredeneerd) bijvoorbeeld aan een LLM prompt worden meegegeven, om zo op een compacte manier de LLM te voorzien van context. Een dergelijke aanpak heeft het voordeel dat het de AI op weg kan helpen door menselijke kennis in het systeem te vervatten. Daarmee wordt kennisextractie (en abstractie) niet overgelaten aan de grillen van de statistische taalgenerator en wordt het risico op hallucinaties ingeperkt.



Box 1: Voorbeeld van een Knowledge Graph met kennis van de bestelgeschiedenis van een klant van een fictieve webshop. Op basis hiervan kan kennis afgeleid worden (bijvoorbeeld producten zonder een leverdatum zijn nog niet geleverd) waarmee vervolgens een LLM een specifieke instructie gegeven kan worden, bijvoorbeeld: "Vraag de klant of zij contact opneemt over de handdoek, die nog niet geleverd is."

Conclusies en aanbevelingen

Veel Conversational AI toepassingen zetten de LLM onterecht op de regiestoel en verlangen veel meer van een *Generatief* systeem dan waar het voor bedoeld is. Hoe verleidelijk (en eenvoudig) het ook is om de LLM tot het centrum van onze Conversational AI te verheffen, noodzakelijk is het niet. Klassieke (en bewezen) AI en Machine Learning technieken kunnen (vaak kosten effectiever dan een LLM) functioneel de rol vervullen van menselijke cognitiemodules. Regie over deze modules kan daarnaast belegd worden bij door de mens geschreven algoritmen waardoor de LLM wordt vrijgespeeld om zijn *raison d'être* uit te voeren: taalgeneratie. •

Mogelijk interessante literatuur

- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big?. In Proceedings of the 2021 ACM conference on fairness, accountability,

[6] https://synthedia.substack.com/p/anthropics-llm-claude-now-has-a-75k
[7] https://hbr.org/2023/07/how-to-make-generative-ai-greener
[8] https://www.businessinsider.com/chatgpt-generative-ai-water-use-environmental-impact-study-2023-4
[9] https://plato.stanford.edu/entries/language-thought/
[10] https://arxiv.org/pdf/2212.05767.pdf
[11] Voor een overzicht zie https://arxiv.org/pdf/2306.08302.pdf.

and transparency, pp. 610-623. https://doi.org/10.1145/3442188.3445922.

- DirectResearch (2022). De Nationale Voice Monitor. https://www.directresearch.nl/blogs/de-nationale-voice-monitor-2022/.
- FitzPatrick, I. (2022). Large language models: Geen silver bullet. In Dixit. Tijdschrift over toegepaste taal- en spraaktechnologie, 19, 1, pp. 28-29.
- Fodor, J. A. (1975). The language of thought (Vol. 5). Harvard university press.
- Gendron, W. (2023). ChatGPT needs to 'drink' a water bottle's worth of fresh water for every 20 to 50 questions you ask, researchers say. https://www.businessinsider.com/chatgpt-generative-ai-water-use-environmental-impact-study-2023-4.
- Kinsella, B. (2023). Anthropic's LLM Claude Now Has a 75K Word Context Window. Consider What That Means. https://synthedia.substack.com/p/anthropics-llm-claude-now-has-a-75k.
- Kumar, A. & Davenport, T. (2023). How to Make Generative AI Greener. https://hbr.org/2023/07/how-to-make-generative-ai-greener.
- Pan, S., Luo, L., Wang, Y., Chen, C., Wang, J., & Wu, X. (2023). Unifying Large Language Models and Knowledge Graphs: A Roadmap. arXiv preprint arXiv:2306.08302. https://doi.org/10.48550/arXiv.2306.08302.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł. & Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30. https://doi.org/10.48550/arXiv.1706.03762.
- Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., & Cao, Y. (2022). React: Synergizing reasoning and acting in language models. arXiv preprint arXiv:2210.03629. https://doi.org/10.48550/arXiv.2210.03629.

PROMPT

"You run in a loop of **Thought**, **Action**, **PAUSE**, **Observation**.

At the end of the loop you output an **Answer**.
Use Thought to describe your thoughts about the question you have been asked.
Use Action to run one of the actions available to you – then return PAUSE.
Observation will be the result of running those actions.

Your available actions are:

wikipedia:
e.g. wikipedia: Django
Returns a summary from searching Wikipedia

[...]

Example Session

Question: What is the capital of France?
Thought: I should look up France on Wikipedia
Action: wikipedia: France
PAUSE

You will be called again with this:

Observation: France is a country. The capital is Paris.

You then output:

Answer. The capital of France is Paris"

Box 2. Prompt om ReAct te implementeren, inclusief een voorbeeld voor een sessie met een gebruiker en de gewenste reactie door het LLM (naar voorbeeld van https://simonwillison.net/2023/Aug/3/weird-world-of-LLMs/). Het model wordt gevraagd om te "redeneren" over een probleem en krijgt de opdracht om een geschikte actie te kiezen die kan worden uitgevoerd door aanvullende code. De actie, zoals zoeken op wikipedia, wordt uitgevoerd en het resultaat wordt door het model gebruikt om het uiteindelijke antwoord te construeren.

GROTE TAALMODELLEN:
JURIDISCH NOG DEELS ONVERKEND TERREIN

Nu Grote Taalmodellen (meestal Large Language Models of LLM's genoemd) en hun toepassingen gemeengoed lijken te worden, is het bijna onvermijdelijk dat deze ook juridische impact zullen hebben. In dit artikel geven we een doorkijk naar de juridische vraagstukken die spelen of beginnen te spelen en wat er op wetgevingsterrein te verwachten valt op de korte termijn.

Door:
Walter van Holst en Mariette Lokin, Hooghiemstra & Partners

Vanuit een juridisch perspectief zijn de vragen rond LLM's op te splitsen in een drietal rechtmatigheidsvragen:

1. Rechtmatigheid van het trainingsproces;
2. Rechtmatigheid van de toepassing van het LLM;
3. Rechtmatigheid van de uitvoer van het LLM.

Ad 1. Rechtmatigheid van het trainingsproces

Hier wreekt zich dat in het recht verschillende, onafhankelijk van elkaar bestaande concepten van rechten op informatie bestaan. Zo zijn er eigendomsachtige rechten zoals het auteursrecht en het databankenrecht, maar ook het recht op gegevensbescherming. Het laatstgenoemde recht is nadrukkelijk geen eigendomsachtig recht, maar een mensenrecht dat in het spel komt als de trainingsdata informatie bevat die gerelateerd is aan natuurlijke personen (mensen van vlees en bloed).

Een van de grote auteursrechtelijke vragen die boven de markt hangt, is in hoeverre het 'trainen' van LLMs een auteursrechtelijk relevante verveelvoudiging is en of die al dan niet onder een uitzondering valt. In Europa kennen wij in het auteursrecht (wat per lidstaat verschilt, maar wel in hoge mate gelijksoortig ingericht is) veelal een gesloten stelsel: verveelvoudiging van een auteursrechtelijk beschermd werk is niet toegestaan behoudens wettelijke uitzonderingen en beperkingen op het auteursrecht, en uiteraard behoudens toestemming van de rechthebbende (de auteur of een rechtsopvolger daarvan). In de Verenigde Staten kent men een open stelsel, met het concept van *fair use* voor gebruik, waarbij het vragen van toestemming aan de rechthebbenden maatschappelijk disproportioneel zou zijn. Die laatste benadering laat meer ruimte voor technologische vernieuwing zonder wetgevingsactiviteiten, maar vergt meer rechtsvorming in de rechtspraak.

Voor gebruik van *artificial intelligence* (AI) en *machine learning* (ML) in een wetenschappelijke context zijn in Europa uitzonderingen geschapen op het auteursrecht. Toch zal enige frictie onvermijdelijk zijn en is de vraag of trainingsdata rechtmatig worden gebruikt een bron van onzekerheid.

Fair use

De redeneerlijn van de vooral Amerikaanse partijen die LLM's trainen en die nu in de belangstelling staan, is gebaseerd op *fair use*, gecombineerd met een concept dat wij in Europa ook kennen. Dat is het concept van zodanige transformatie van het werk dat het oorspronkelijke karakter ervan verloren is gegaan. De *fair use* doctrine gaat uit van het idee dat het auteursrecht niet in de weg mag staan van exploitatievormen door anderen dan rechthebbenden, waar toestemming vragen praktisch gezien onhaalbaar is of zelfs onwenselijk is. Toegepast op het trainen van LLM's betekent dit dat door de veelheid van werken hier een argument ligt en/of dat hier geen sprake is van het eenvoudigweg kennis nemen van een werk. De vervolgstap, wie recht-

hebbende is van het resultaat van de training, is een openstaande vraag, waarbij het nog de vraag is nog of LLM's transformatief genoeg zijn om binnen de Amerikaanse doctrines op dit terrein te passen of niet.

Auteursrecht
Voor wie geen beroep kan of durft te doen op de uitzonderingen voor wetenschappelijk onderzoek is het goed te weten dat in de Europese context ruwweg de volgende categorieën van auteursrechtelijke bescherming bestaan:

- Publiek domein (rechtenvrij): het gaat hier om materiaal waarvan de beschermingsduur van het auteursrecht is verstreken. Hoofdregel is zeventig jaar te rekenen na het overlijdensjaar van de langstlevende auteur;
- Creative Commons (CC) of soortgelijke licenties: met name de CC-BY- en CC0-licenties vereisen geen aanvullende licenties. Dit kan anders liggen bij de CC-ND (geen afgeleide werken) en CC-NC (geen commercieel gebruik) varianten. CC0 is met name prettig omdat deze ook een expliciete databanken rechterlijke licentie omvat, iets wat bij de overige Creative Commons licenties niet altijd het geval is.
- Overige: alle rechten zijn doorgaans voorbehouden. Over deze categorie bestaat nog de grootste onzekerheid. Het is tegelijkertijd de grootste categorie, omdat dit nu eenmaal het uitgangspunt van het auteursrecht (en ook het databankenrecht) is.

Dit alles is slechts de auteursrechtelijke dimensie van het trainingsproces van LLM's. Veel trainingsdata bevatten gegevens die direct of indirect tot mensen herleidbaar zijn, zogenaamde persoonsgegevens. Dan komt ook de Algemene Verordening Gegevensbescherming (AVG) in beeld. De AVG vereist voor de verwerking van persoonsgegevens een verwerkingsgrond. Voor het trainen van LLM's zal dit veelal het zogenoemde gerechtvaardigd belang zijn, wat zeker in een wetenschappelijke context goed verdedigbaar is. Voor het resulterende LLM is het vooral de vraag of personen uit de trainingsset terug te herleiden zijn, dit is een meer feitelijke vraag, zij het met een juridische consequentie. Opvallend is daarbij dat de wetenschappelijke inzichten op dit terrein nog aan verandering onderhevig zijn. Een kanttekening bij de grondslag van gerechtvaardigd belang is overigens dat zodra het om de zogenoemde bijzondere categorieën van persoonsgegevens gaat (bijvoorbeeld over gezondheid, religie, etniciteit, politieke voorkeuren, seksuele geaardheid) de AVG aanvullende zorgvuldigheidseisen opwerpt, want daarvoor geldt in beginsel een verwerkingsverbod. Opnieuw, in een wetenschappelijke context zijn hier uitzonderingen mogelijk, daarbuiten wordt het snel lastiger zonder specifieke wet- en regelgeving.

Ad 2. Rechtmatigheid van het gebruik van een LLM
Voor het gebruik van een LLM spelen vanuit een auteursrechtelijk perspectief weinig andere vragen dan bij het trainen van een LLM, met de kanttekening dat een getraind LLM naar alle waarschijnlijkheid zo ver af zal staan van de trainingsdata dat de auteursrechten op de trainingsset mogelijk geen rol meer spelen. De juridische onzekerheid hierover zou daarom lager uit kunnen vallen.

Supervisie
Vanuit AVG-oogpunt daarentegen wordt het gebruik van een LLM al snel complexer dan de training. Bij gebruik zal het getrainde model namelijk veelal toegepast worden op andere persoonsgegevens dan in de trainingsset zaten, en ontstaat de vraag of sprake is van geautomatiseerde besluitvorming met een aanmerkelijk effect op de betrokkene. Het eerste aspect van deze vraag (is er sprake van geautomatiseerde besluitvorming) is vooral in het spel als sprake is van inzet van een LLM als 'autopiloot', dus zonder menselijke 'co-piloot' die supervisie houdt. Het tweede aspect (is er sprake van aanmerkelijk effect) betreft de mogelijke impact op het leven van het subject van de besluitvorming. Die kan al aanmerkelijk zijn bij inzet van een LLM voor bijvoorbeeld plagiaatdetectie in afstudeerscripties, maar zeker bij inzet voor het selecteren van CV's van kandidaten in een sollicitatieproces. Als van aanmerkelijk effect sprake is, dan bestaat onder de AVG het recht van betrokkene om niet onderworpen te worden aan geautomatiseerde besluitvorming (behoudens enkele uitzonderingen). Voor juristen ontstaan interessante vragen over de grenzen van het inzagerecht van betrokkenen, omdat dat zich uitstrekt tot de "onderliggende logica, alsmede het belang en de verwachte gevolgen voor de betrokkene" van de inzet van geautomatiseerde besluitvorming (zie ook artikel 15 lid 1 onder h AVG).

Papegaaien
Op dit laatste punt hebben LLM's vooralsnog een grote zwakte: ze worden ook wel 'stochastische papegaaien' genoemd omdat het veelal slecht uit te leggen is waarom ze tot bepaalde resultaten zijn gekomen. Sterker nog, zij hebben een neiging om stellige uitspraken te doen die niet noodzakelijkerwijze feitelijk juist zijn, wat ook wel het 'hallucineren' of 'fabuleren' van LLM's genoemd wordt. De AVG legt bij geautomatiseerde besluitvorming de lat hoger dan bij menselijke besluitvorming, hoewel die evenzeer arbitraire trekken kan hebben en waarbij in de praktijk aan menselijke beslissers ook niet altijd ruimte voor 'gezond verstand' gegeven wordt.

Verder is het onder de AVG zo dat nieuwe technologieën pas op grote schaal toegepast kunnen worden op persoonsgegevens als een *data protection impact assessment* (DPIA) heeft plaatsgevonden. Een belangrijk onderdeel hiervan is een risicobeoordeling en de verplichting in het geval (ook nadat maatregelen om de risico's te mitigeren zijn genomen) grote restrisico's overblijven, de toezichthouder te raadplegen. Deze laatste kan gedurende zo'n voorafgaande raadpleging de voorgenomen verwerking van persoonsgegevens verbieden.

Ad 3. Rechtmatigheid van de uitvoer van een LLM
Bij inzet van de uitvoer (het resultaat) van een LLM is de context van die inzet van belang. Zo heeft de inzet van de uitvoer van een LLM in bijvoorbeeld nieuwsartikelen en met name het 'hallucineren' of 'fabuleren' van feiten al situaties opgeleverd waarbij mensen onterecht van seksuele intimidatie beschuldigd zijn. Nog los van de AVG geldt dat in de meeste jurisdicties ook regels over smaad en laster bestaan. Onder de AVG bestaat een verplichting zorg te dragen voor de juistheid van de verwerkte persoonsgegevens, wat opnieuw lastig te waarborgen is met LLM's.

Nabije toekomst
Bovenop de hiervoor genoemde juridische kaders is in de EU een wetgevingstraject gaande om de inzet van AI aan waarborgen te onderwerpen in de AI Act. Over deze verordening (die net als de AVG rechtstreekse werking zal krijgen in Nederland) is nog geen politieke overeenstemming over bereikt, maar de contouren staan wel vast. Er worden eisen gesteld aan de risicobeheersing door partijen die AI-systemen (dus ook LLM's) op de markt brengen, maar ook aan de gebruikers. Met name de real-time verwerking van biometrische gegevens die waargenomen zijn in publieke ruimten door overheden (bijvoorbeeld via camera's) wordt aan banden gelegd, vanwege het hoge risico dat hier voor betrokkenen aan is verbonden. Aan systemen met een lager risico worden minimaal transparantievereisten gesteld. Dit *compliance*-risico wordt zoals gezegd bij zowel leveranciers als gebruikers neergelegd, maar voor zogenaamde hoog-risico-AI wordt dit expliciet bij de leverancier gelegd. Separaat is er nog de AI Liability Act, die aansprakelijkheid voor buitencontractuele schade door de exploitatie van AI regelt. Hierin wordt voor een certificeringsmodel gekozen zoals we dat kennen van de CE-keurmerken. Het is realistisch om aan te nemen dat deze regels rond 2026 in werking zullen zijn. Te verwachten valt dat dit de bescherming tegen ongewenste neveneffecten van de inzet van LLM's verder zal kunnen beteugelen, maar of ze de ontwikkeling van LLM's kunnen bijhouden, blijft de vraag. •

HOE GAAT GENERATIVE AI DE TTS-INDUSTRIE ONTWRICHTEN?

Door:
Jaebok Kim,
Research Lead,
ReadSpeaker

Generatieve AI is een soort kunstmatige intelligentie die nieuwe inhoud kan creëren, zoals tekst, afbeeldingen, audio of andere media. Hiervoor leren generatieve modellen[1] de patronen en structuur van hun trainingsgegevens en genereren vervolgens nieuwe gegevens. Wel zijn ze gevoelig voor de input en zullen ze vergelijkbare kenmerken hebben als deze input.[2] Generatieve AI verschilt van traditionele AI doordat het niet alleen gegevens analyseert of classificeert, maar ook nieuwe gegevens produceert die voorheen waarschijnlijk niet bestonden.[3]

Enkele voorbeelden zijn:

- **Tekstgeneratie:** Generatieve AI kan teksten schrijven over verschillende onderwerpen, zoals nieuwsartikelen, verhalen, gedichten en samenvattingen. ChatGPT is bijvoorbeeld een generatief AI-model dat gesprekken kan voeren met gebruikers.[4]
- **Afbeeldingen genereren:** Generatieve AI kan (redelijk) realistische afbeeldingen maken van bijvoorbeeld gezichten, objecten of landschappen. StyleGAN is bijvoorbeeld een generatief AI-model dat afbeeldingen van menselijke gezichten van hoge kwaliteit kan genereren.[5]
- **Code genereren:** Generatieve AI kan code schrijven voor verschillende programmeertalen, zoals Python, Java, Delphi of C#. Codex is bijvoorbeeld een generatief AI-model dat code kan genereren uit beschrijvingen in natuurlijke taal.
- **Audio genereren:** Generatieve AI kan realistische geluiden of muziek maken, zoals spraaksynthese, stemklonen of muziekcomposities. WaveNet[6] is bijvoorbeeld een generatief AI-model dat natuurlijk klinkende spraak kan genereren.

In dit artikel richten we ons op het genereren van audio, in het bijzonder spraaksynthese ofwel tekst-naar-spraak (TTS). De generatieve AI lijkt de tekst-naar-spraak industrie te ontwrichten doordat zij het op een eenvoudige manier mogelijk maakt om natuurgetrouwe, natuurlijk klinkende en aangepaste of gepersonaliseerde stemmen te creëren.

Ten eerste zullen we zien hoe traditionele AI-methoden worden gebruikt op het gebied van TTS. Ten tweede zullen we zien hoe generatieve AI beter geschikt is voor TTS in vergelijking met traditionele methoden, maar ook wat de beperkingen zijn. Ten derde introduceren we AI-pipelines die ReadSpeaker heeft ontwikkeld om TTS op maat te produceren en bespreken we toekomstige onderzoeksrichtingen.

Het gebruik van AI in TTS

Voor de opkomst van generatieve AI werd traditionele AI al veel gebruikt op het gebied van TTS. Voor een beter begrip zal ik de pipelines van TTS doorlopen, die doorgaans uit de volgende drie stappen bestaat:

- **Tekst voorbewerken**
  Dit omvat het coderen van de ingevoerde tekst in een opeenvolging van 'klinkende' eenheden zoals tekens of fonemen. Deze stap kan ook het normaliseren van de tekst inhouden, zoals het uitschrijven van afkortingen en het corrigeren van spelling. Veel van deze elementen zijn sequentiële classificatieproblemen waarbij vaak traditionele algoritmen worden gebruikt.
- **Voorspelling van akoestische parameters**
  Dit bestaat uit het genereren van een reeks akoestische parameters uit de gecodeerde tekst-symbolen. Akoestische parameters worden vaak aangestuurd vanuit een spectrogram dat de intensiteit van boventonen (formanten) weergeeft die in de loop van de tijd veranderen. Het moet in staat zijn om niet alleen de fonetische kenmerken van spraak vast te leggen om de linguïstische inhoud te behouden, maar ook de toonhoogte en het stemtimbre van de spreker. Mel-schaal spectrogram (kortweg melspectrogram) is een van de meest gebruikte parameter types, omdat het op efficiënte wijze de gevoeligheid van het menselijke auditieve systeem nabootst. Om de sequentie te voorspellen, wordt een neuraal netwerkmodel gebruikt om de koppeling tussen de gecodeerde tekstsymbolen en de akoestische parameters te leren. Traditioneel zijn voor deze stap niet per se generatieve modellen nodig, omdat men dacht dat het voldoende was om één-op-één koppelingen tussen de gecodeerde tekst-symbolen en de akoestische parameters vast te leggen. Generatieve modellen zijn echter als recente trend aantrekkelijk omdat ze meer natuurlijke variaties kunnen modelleren.
- **Vocoding**
  De eerder voorspelde opeenvolging van akoestische parameters is niet echt hoorbaar. Een golfvorm is een signaal dat de amplitude van geluid in de tijd weergeeft en dat door mensen kan worden afgespeeld en beluisterd. De laatste stap bestaat dus uit het genereren van spraakgolven of golfvormen uit de voorspelde opeenvolging van akoestische parameters en deze stap wordt vaak 'vocoding' genoemd. De uitdaging is dat een relatie tussen akoestische parameters en golfvormen niet altijd één-op-één is, maar één-op-veel, omdat akoestische parameters gedeeltelijke of abstracte representaties van golfvormen zijn. Daarom past generatieve AI perfect bij deze taak omdat het nieuwe en zinnige voorbeelden kan selecteren.

Wat kunnen we beter wel of niet doen met generatieve AI?

Vocoding is een goed voorbeeld van het gebruik van generatieve AI in TTS; maar het zou meer moeten zijn dan dat. Generatieve AI-modellen kunnen de kenmerken van menselijke spraak leren, zoals intonatie, verbuiging, emotie en accent, en spraak genereren die deze kenmerken nabootst. Een van de belangrijke en interessante eigenschappen van generatieve AI is dat het model nieuwe stemmen kan genereren. Stel dat we honderd sprekers in een trainingsdataset hebben en een generatief AI-model trainen. Dan kunnen we misschien een stem genereren van een nieuwe spreker die niet bestaat in de dataset, en dat doen op een controleerbare manier (bijvoorbeeld door parameters te configureren om de stem te karakteriseren).

Sterker nog, we kunnen ons de integratie voorstellen van een prompt om kenmerken te beschrijven; laten we zeggen "een mannelijke stem van middelbare leeftijd met een Amerikaans accent die besluitvaardig klinkt". Dit is duidelijk een flinke uitdaging van het (nabije) verleden, toen generatieve AI nog niet zo ver ontwikkeld was.

Tot nu toe klinkt generatieve AI-gebaseerde TTS veelbelovend, maar we moeten niet te enthousiast zijn. Er zijn nog steeds veel problemen waardoor veel TTS-aanbieders aarzelen om generatieve AI in te zetten voor hun bedrijf. Sommige problemen zijn bijna opgelost, maar andere zijn nog ver verwijderd van praktische oplossingen.

Generatieve TTS is bijvoorbeeld inherent instabiel. De modellen worden verondersteld nieuwe stemvormen te genereren, maar ze zijn gevoelig, er zou een conflict kunnen optreden tussen nieuwheid en begrijpelijkheid. Met andere woorden, nieuwheid geeft aan hoe ver de nieuwe stemvormen afliggen van de standaardantwoorden die worden gemeten aan de hand van begrijpelijkheid. Hoe kunnen we nieuwe voorbeelden maken zonder de verstaanbaarheid aan te tasten?

De mate van vrijheid kan worden gewogen afhankelijk van de toepassingen. Voor het leren van een taal hebben we bijvoorbeeld een hoge nauwkeurigheid van articulatie (of uitspraak) nodig. Aan de andere kant, voor amusementsdoeleinden (bijvoorbeeld spelletjes) kan een gemiddelde nauwkeurigheid goed genoeg zijn.

Tot slot, maar praktisch gezien behoorlijk belangrijk, zijn veel generatieve modellen rekenkundig duur en vereisen ze een enorme hoeveelheid trainingsgegevens. Dit is vooral een issue voor kleine aanbieders, die veel meer moeite moeten doen om de noodzakelijke data te krijgen dan de grote tech bedrijven. Voor een betere toepassing van generatieve TTS zouden deze problemen in de nabije toekomst moeten worden opgelost.

Hoe/waar gebruikt ReadSpeaker AI?

Als pionier op het gebied van TTS-technologie heeft ReadSpeaker op grote schaal AI-gebaseerde methoden toegepast in de pipelines van productontwikkeling. Bijvoorbeeld, op AI gebaseerde tools labelen automatisch verzamelde opnamen en meten bovendien automatisch de kwaliteit van TTS. Echter, op basis van meer dan 20 jaar ervaring in de industrie, weten we dat AI niet al onze pipelines die menselijke interventies nodig hebben, volledig kan vervangen. ReadSpeaker slaagt er nu in om de beste menselijke kennis te integreren in AI-systemen om zo TTS op maat te produceren voor onze klanten.

En toekomstig onderzoek?

Overal wordt tegenwoordig gesproken over de beloften van generatieve AI. Zeker, het is waarschijnlijk een van de meest bijzondere ontwikkelingen in ons leven. Maar, we mogen de data achter deze ontwikkelingen niet vergeten. Zoals eerder vermeld, vereist generatieve TTS in het algemeen een enorme hoeveelheid trainingsmateriaal. Waar komt dat vandaan? Van ons, mensen! Spraak is een van de meest unieke eigenschappen die ons speciaal maakt, en naarmate we meer spraak leveren aan AI, stellen we die in staat om meer realistische of mensachtige spraak te synthetiseren. Betekent dit dat spraak minder uniek wordt? Hoe beschermen we onze uniekheid als mensen niet langer onderscheid (kunnen) maken tussen menselijke en kunstmatige spraak? Terwijl we uitkijken naar wat generatieve TTS ons in de nabije toekomst zal brengen, moeten we nu al nadenken over de gevolgen die het zou kunnen hebben. •

ReadSpeaker
The Voice of the Web!

[1] https://en.wikipedia.org/wiki/Generative_artificial_intelligence
[2] https://research.aimultiple.com/generative-ai/
[3] https://www.techrepublic.com/article/what-is-generative-ai/
[4] https://www.gartner.com/en/topics/generative-ai
[5] https://en.wikipedia.org/wiki/StyleGAN
[6] https://en.wikipedia.org/wiki/WaveNet

DE BLISSBOT: GEPERSONALISEERDE GESPREKKEN IN HET NEDERLANDS OVER DAGELIJKS LEVEN EN WELBEVINDEN

In het project BLISS (Behaviour-based Language-Interactive Speaking Systems), dat werd gefinancierd via het NWO-programma DATA2PERSON, werkten onderzoekers van het Centre for Language and Speech Technology (CLST) van de Radboud Universiteit samen met collega's van Human Media Interaction (HMI) van de Universiteit Twente, het bedrijf Games for Health (GfH) en het bedrijf ReadSpeaker. Zij werkten aan de ontwikkeling van een chatbot die met mensen in het Nederlands kan spreken over hun dagelijks leven en welzijn. In de eerste fase van het project was de beoogde doelgroep ouderen in zorginstellingen.

Door:
Iris Hendrickx, Wieke Harmsen, Catia Cucchiarini, Helmer Strik, CLST, RU, Nijmegen

De BLISSbot kan gepersonaliseerde gesprekken voeren over geliefde activiteiten, passies, interesses en sociale participatie. Het idee is dat deze gesprekken kennis oplevert die mensen inzicht geeft in hun eigen leven. De BLISSbot heeft een geheugen en kan daarom tijdens deze gepersonaliseerde gesprekken terugkomen op onderwerpen die eerder zijn besproken met dezelfde persoon. Verder kunnen gebruikers met de BLISSbot communiceren door te praten in plaats van te typen. Dit maakt de BLISSbot toegankelijker, vooral voor mensen die typen lastig vinden.

Een belangrijk aspect van dit project is dat er gewerkt is aan de ontwikkeling van Nederlandstalige taal- en spraaktechnologie die in de chatbot kan worden geïncorporeerd en die nu ook voor andere toepassingen kan worden gebruikt.

Corona

Een groot probleem bij de uitvoering van het BLISS-project was dat de experimenten uitgevoerd moesten worden tijdens de Covid-pandemie. Omdat het toen onmogelijk was om zorginstellingen te bezoeken, moest de hele lokale architectuur op een laptop overgeheveld worden naar een online architectuur en moest er gezocht worden naar een alternatieve doelgroep die wel bereikbaar was voor experimenten.

Een lijn van onderzoek in BLISS was gericht op het verbeteren van de automatische spraakherkenning (ASH). Het is bekend dat de kwaliteit van de spraakherkenning minder goed is voor spraak van ouderen. Dit heeft onder andere te maken met de afwijkende spraak van ouderen en de beperkte beschikbaarheid van spraakdata waarop ASH-systemen getraind kunnen worden. In het kader van BLISS zijn dan ook spraakdata verzameld en beschikbaar gesteld die nu ook door anderen kunnen worden gebruikt. De *BLISS Spoken Dialogue Dataset*[1] bestaat uit Nederlandse spraakopnamen van volwassen deelnemers die spreken met het BLISS-dialoogsysteem (v1) over alledaagse bezigheden en activiteiten waar ze plezier aan beleven. De data bevat 55 opnamen met een gemiddelde duur van 2 minuten en 34 seconden. Daarnaast is ook de BLISS Dialogue Summaries Dataset publiek beschikbaar.[2] Deze dataset bestaat uit samenvattingen van dialogen met de BLISSbot.

De spraakherkenningscomponent

De automatische spraakherkenningcomponent van de BLISSbot bestond uit een Kaldi-model, waarbij het akoestische model is getraind op het Corpus Gesproken Nederlands, en het taalmodel op het Twente Nieuws Corpus. We hebben geexperimenteerd met het creëren van BLISS-taalmodellen door de taalmodellen uit te breiden met dialoogdata. Dit leidde echter niet tot betere herkenning van de spraak. Aan het einde van het project hebben we spraakopnames laten herkennen met behulp van het end-to-end spraakherkenningsmodel Whisper. Dit leidde over het algemeen tot betere herkenningsresultaten.

De stem van de BLISSbot

Om de stem van de chatbot te optimaliseren is onderzoek gedaan naar het effect van twee tekst-naar-spraak-modules van Readspeaker: de neurale stem versus een meer klassieke variant. De neurale stem zou beter zijn in het uitdrukken van prosodie en daardoor meer emotie laten doorklinken. De verwachtingen waren ook dat een neurale-netwerk-gebaseerde stem ervoor zou zorgen dat het gesprek meer aanvoelt als een gesprek met een vriend in plaats van een gesprek met een assistent. Experimenten met echte gebruikers toonden echter aan dat de kwaliteit van de stem geen effect had op de manier waarop het gesprek met de chatbot werd gepercipieerd of gewaardeerd door de gebruikers.

Informatie vertellen over jezelf en praten over lastige onderwerpen

In het kader van het BLISS-project hebben we ook verschillende methoden onderzocht om mensen aan te moedigen zich open te stellen en informatie te delen met de BLISSbot (self-disclosure). We hebben onderzocht of een BLISSbot die meer over zichzelf vertelt ervoor zorgt dat gebruikers ook meer over zichzelf gaan vertellen. Verder hebben we onderzocht hoe een chatbot gebruikt kan worden om te praten over onderwerpen waarover moeilijk te praten (of zelfs maar na te denken) is. Daarom zijn we vorig jaar een samenwerking aangegaan met een ander NWO-project op de Faculteit Letteren van de Radboud Universiteit: Beyond A Fear Of Death. Binnen dit project wordt onderzocht of films met een zinvolle boodschap (in tegenstelling tot op plezier gericht entertainment) een gesprek over de dood kunnen faciliteren. Mensen hebben de neiging om niet over de dood na te denken en hun gedachten daarover te onderdrukken, ondanks het feit dat de dood een onlosmakelijk onderdeel van ons leven is. Omdat chatbots een anoniem, geduldig en niet-oordelend platform bieden voor discussie, waren we geïnteresseerd in het combineren van de doelstellingen van de twee projecten. We onderzochten in hoeverre we chatbots kunnen inzetten om mensen te laten nadenken en te laten praten over de dood. Onze resultaten lieten zien dat mensen er zeker wel voor openstaan om met een chatbot te praten over een dergelijk gevoelig onderwerp.

De BLISSbot in de praktijk

De BLISSbot werd verder beschikbaar gesteld aan bedrijven in het kader van het NWO Data-2Person top-up project. In dit project werd de BLISSbot gebruikt om een in de zorg gehanteerde vragenlijst in te vullen. In plaats van een vragenlijst op papier, die de cliënt zelfstandig invult, worden de vragen nu door de BLISSbot gesteld en kan de cliënt deze mondeling beantwoorden. Dit heeft het voordeel dat de vragenlijsten laagdrempeliger zijn in gebruik en dat er in de toekomst tijdens het invullen van de vragenlijst ook gelijk vragen gesteld kunnen worden en vragen verduidelijkt kunnen worden.

En hoe nu verder?

Het BLISS-project is inmiddels afgerond, maar een sprekende chatbot kan voor allerlei zorgtoepassingen nuttig zijn, zoals het vroegtijdig opsporen van ziektes die zich in spraak manifesteren zoals Parkinson, dementie, depressie en chronische obstructieve longziekte (COPD). Dit laatste wordt nu onderzocht in een nieuw project Ontwikkeling van een digitale COPD begeleider voor een betere levensstijl (DACIL) | DACIL (ru.nl)[3], waarin CLST participeert samen met vele andere partners zoals Universiteit Maastricht, Vrije Universiteit Amsterdam, in samenwerking met Lung Foundation, Lung Alliance, NVALT, Universiteit Twente, PHAROS, Kenniscentrum Digitale Zorg, Zorginstituut, Zana Technologies GmbH, KI:elements en Monitair BV. Het project DACIL heeft als doel mensen met COPD te ondersteunen in het omgaan met hun aandoening door een digitale coach te ontwikkelen. Deze kan hun stemgeluid en ademhaling beoordelen, een behandeladvies geven als hun toestand slechter wordt en ook leefstijladviezen geven om de COPD te verbeteren. •

Publicaties

Naast de datasets heeft het BLISS-project vele wetenschappelijke publicaties opgeleverd waarvan het proefschrift van Jelte van Waterschoot een van de belangrijkste is: van Waterschoot, J. B. (2 December 2021). Personal and Personalized Conversations: Designing Agents Who Want To Connect With You. PhD Thesis, University of Twente. https://doi.org/10.3990/1.9789036552691

Ook werd een workshop georganiseerd rond het thema van gepersonaliseerde chatbots waaraan wetenschappers en bedrijven hebben deelgenomen: UMAP2021 workshop: Personalized Intelligent Conversational Agents (PICA), June 25, 2021, Utrecht, The Netherlands.

Dankwoord

We danken de collega's in het BLISS-project: Jelte van Waterschoot, Marïet Theune, Rob Tieben, Esther Klabbers, Marcel de Korte, Loes van Bemmel en de studenten die hebben meegeholpen waaronder Troy Koster, Esmee van Zon, Charlotte van Beek, Lissy Severins, Karlijn Spijkers en Nikki Evers.

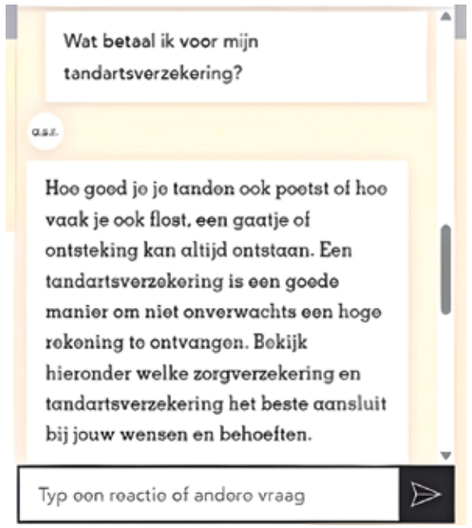
[1] http://hdl.handle.net/10032/tm-a2-r5
[2] http://hdl.handle.net/10032/tm-a2-v3
[3] https://www.ru.nl/onderzoek/onderzoeksprojecten/ontwikkeling-van-een-digitale-copd-begeleider-voor-een-betere-levensstijl-dacil

DE KLANT VERWACHT MEER VAN DIGITALE ASSISTENTEN

Er is nog heel wat werk aan de winkel voor organisaties die zich bezighouden met taal- en spraaktechnologie, zo blijkt uit de Nationale Voice Monitor. Dat is een onderzoek - geïnitieerd door Y.digital en Direct Research - onder circa 1.000 Nederlandstalige consumenten (18+) naar het gebruik van chatbots en spraakassistenten.

Door:
Carla Verwijmeren, CCO bij Y.digital

De grootste uitdaging voor spraakassistenten ligt op het gebied van het spraakmodel. Voor een goede speech-to-text functionaliteit moeten ook dialecten, accenten, streektalen en stemmen van ouderen verwerkt kunnen worden zodat gebruikers beter begrepen worden. Daarnaast is taalbegrip van zowel chatbots als spraakassistenten een onderwerp dat veel meer aandacht vergt. Vaak wordt gedacht dat het alleen mis gaat bij complexe vragen, maar helaas worden ook eenvoudige vragen in de Nederlandse standaardtaal nog wel eens niet goed beantwoord.



Voorbeeld van een eenvoudige vraag die niet goed wordt begrepen.

Consumenten willen beter begrepen worden door chatbots en spraakassistenten

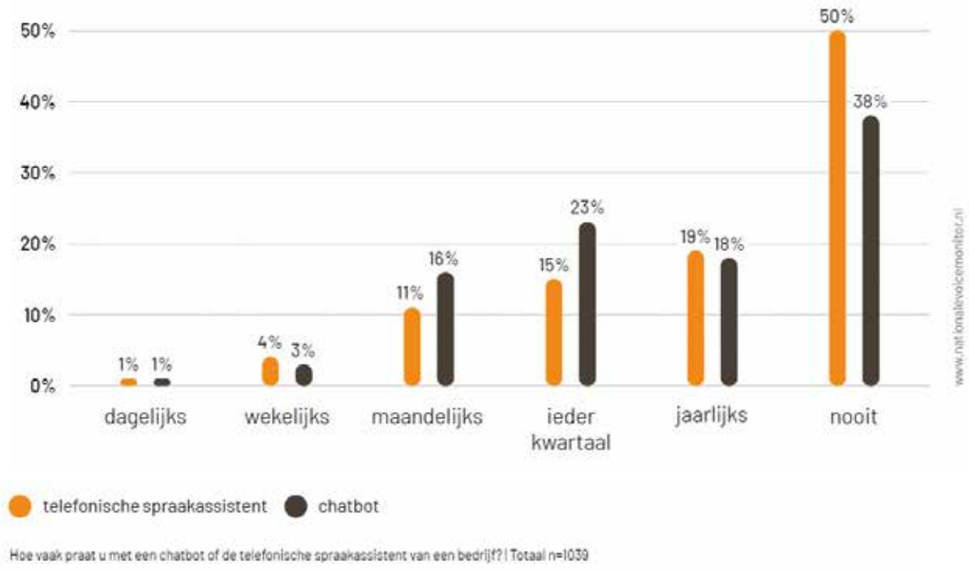
Steeds vaker komt de Nederlandse consument in aanraking met digitale assistenten, zoals een chatbot op een website of een spraakassistent in een contactcenter. Hoewel consumenten open staan voor geautomatiseerd klantcontact, voelt 90% van de Nederlandse consumenten zich hierbij niet goed begrepen. Het is de vraag of de verdere opkomst van de Large Language Models (LLMs) en Knowledge Graphs (KG) hierin een kentering teweeg kan brengen. De verwachtingen van de gebruiker worden er in ieder geval niet minder op.

Geautomatiseerd klantcontact kan grootste ergernissen helpen wegnemen

In contact met organisaties heeft inmiddels 62% van de consumenten ervaring opgedaan met chatbots en 50% met spraakassistenten. De twee grootste ergernissen van de Nederlandse consument bij telefonisch contact met een organisatie zijn lange wachttijden (64%) en uitgebreide keuzemenu's (58%). Digitale assistenten worden door zowel organisaties als consumenten gezien als mogelijke oplossing voor het tegengaan van deze ergernissen. Consumenten verwachten dat spraakassistenten in contactcenters een oplossing kunnen zijn voor lange wachttijden (50%), lange keuzemenu's (31%) en medewerkers met te weinig verstand van zaken (27%).

Consument wil kunnen overstappen van geautomatiseerd contact naar een medewerker

Digitale assistenten werken volgens de consument goed voor eenvoudige vragen en toepassingen. Zo is 77% van de consumenten enthousiast over doorverbinden, 60% over hulp bij klantidentificatie en 48% over het afhandelen van eenvoudige vragen. Maar er is nog weinig vertrouwen in de technologie als het gaat om complexere vragen of langere gesprekken. Zo'n 10% van de Nederlandse consumenten heeft op dit moment



Hoe vaak praat u met een chatbot of de telefonische spraakassistent van een bedrijf? | Totaal n=1039

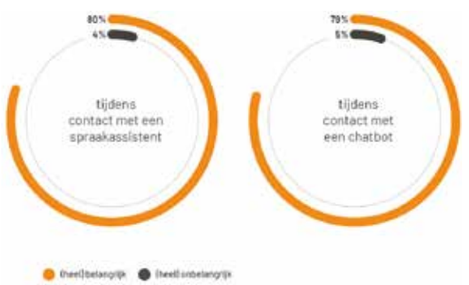
evenveel vertrouwen in een mens als in een digitale assistent. Van de Nederlandse consumenten vindt circa 80% het dan ook belangrijk om over te kunnen stappen van een chatbot of spraakassistent naar een medewerker als dat nodig is.

Managen van klantverwachtingen is heel belangrijk

Uit het onderzoek blijkt verder dat circa 13% van de consumenten een positieve houding heeft ten opzichte van chatbots en spraakassistenten. De ontwikkelingen op dit gebied gaan razendsnel. Intelligente technologische ontwikkelingen zullen in de toekomst een steeds grotere rol spelen binnen klantcontact. Organisaties moeten wel bouwen aan het vertrouwen van de consument. Uit het onderzoek blijkt dat het van belang is om de basis van het klantcontact goed op orde te hebben en duidelijk te zijn over wat een chatbot of spraakassistent wel of (nog) niet kan.

Consumenten willen een stem die duidelijk is en niet robotisch klinkt

Circa 33% van de Nederlandse consumenten vindt het belangrijk dat een organisatie een vas-



Hoe belangrijk is het voor u dat u, tijdens een gesprek met een digitale assistent, de mogelijkheid heeft om aan te geven dat u een medewerker wilt spreken? | n=1039
Stel dat u contact heeft met een bedrijf, en in aanraking komt met een chatbot of telefonische spraakassistent. In hoeverre bent u het eens of oneens met onderstaande stelling? | n= 1039

te stem gebruikt voor een spraakassistent. Het kunnen inspelen op emoties door een spraakassistent vindt 27% van de consumenten belangrijk. Op de vraag of de spraakassistent hetzelfde dialect of accent moet spreken als de consument, geeft 57% van de consumenten aan dit niet belangrijk te vinden. Ook vindt 67% van de consumenten het niet belangrijk dat de stem van een spraakassistent van hetzelfde geslacht is als dat van de gebruiker. Wel hechten consumenten er waarde aan dat de stem van een spraakassistent rustig en duidelijk klinkt en niet robotisch is.

Helft Nederlandse consumenten deelt liever geen persoonlijke informatie met digitale assistent

Consumenten delen hun persoonlijke data liever met een mens dan met een digitale assistent, zo blijkt uit het onderzoek. Een op de drie consumenten maakt zich zorgen over de privacy bij het delen van persoonlijke informatie. Vooral het BSN nummer (77%) en betaalgegevens (74%) zijn data die mensen liever niet delen. Tegelijkertijd investeren organisaties veel om te voldoen aan alle privacy- en data security eisen, en overige

wet- en regelgeving bij de ontwikkeling van chatbots en spraakassistenten. Uit diepte-interviews met zes organisaties blijkt dat dit bij alle organisaties een hoge prioriteit heeft. Hierbij zijn ze vooral op zoek naar een goede balans tussen privacy enerzijds en – de door de consument gewenste - gepersonaliseerde service anderzijds.

Hoe komen we naar betere spraakmodellen voor de Nederlandse taal?

Vrijwel alle private en publieke organisaties in de Nederlandse markt hebben behoefte aan een betere kwaliteit van spraakmodellen. Om hiertoe te komen, is een hele andere aanpak nodig, initiatieven zijn nu nog te veel versnipperd. Er is geen enkele organisatie in de Nederlandse markt die over kwalitatief en kwantitatief voldoende audiodata beschikt om spraakmodellen te kunnen trainen op alle accenten, dialecten, streektalen en doelgroepen. Bovendien voldoen beschikbare data vaak ook niet aan wet- en regelgeving om te mogen gebruiken voor zowel de publieke als private sector. Om krachten te bundelen is een nieuwe Stichting opgericht: de Nederlandstalige Spraakcoalitie. Wil je hier meer over weten? Check dan: www.spraakcoalitie.nl.

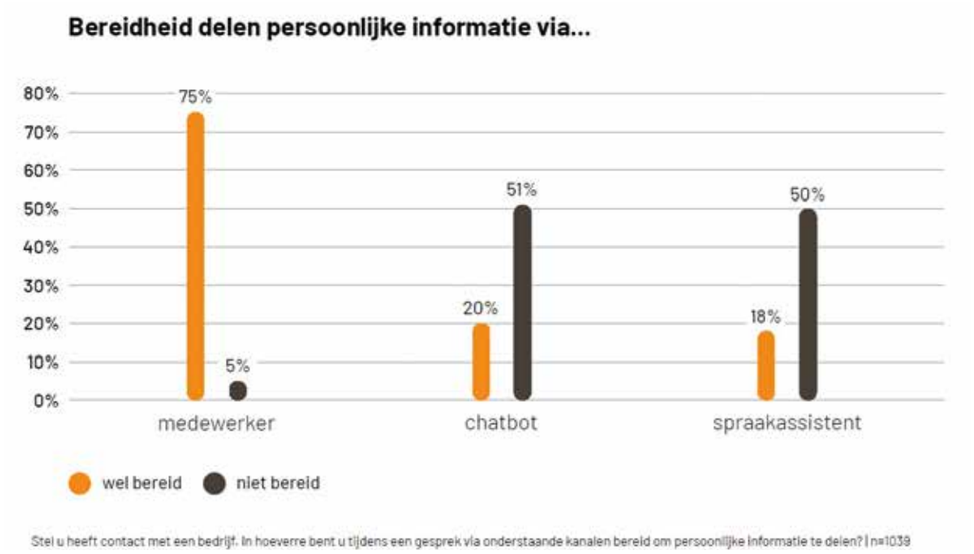
Hoe gaan we naar een verbeterd taalbegrip?

De verdere adoptie van Large Language Models in combinatie met toepassingen als ChatGPT zorgen voor nog hogere verwachtingen bij de Nederlandse consument als het gaat om het begrip van chatbots en spraakassistenten. Tegelijkertijd weten we dat deze modellen geen daadwerkelijk begrip hebben, noch toegang hebben tot databronnen die organisatie-specifiek zijn. Tot slot zijn het vanuit oogpunt van data security en privacy ook geen end-to-end oplossingen. Het wordt daarmee de kunst om de taalvaardigheid van deze modellen te koppelen aan domeinspecifieke kennis binnen organisaties. In deze DIXIT beschrijft Ian Fitzpatrick van Y.digital zijn visie hierop. Interessant voor iedere organisatie die hierin verder wil!

Over de Nationale Voice Monitor

De Nationale Voice Monitor is een onderzoek naar het gebruik van chatbots en spraakassistenten in de Nederlandse markt. Het onderzoekspanel bestond uit 1.039 respondenten, dat is representatief voor de Nederlandse bevolking (18+ jaar). De resultaten werden verrijkt met de praktijkinzichten vanuit zes Nederlandse organisaties en een expertpanel vanuit de partners bij het onderzoek. Het onderzoek werd uitgevoerd van januari 2023 tot en met maart 2023 door onderzoeksbureau DirectResearch en AI-bedrijf Y.digital. Dit gebeurde in partnership met Achmea, de Klantenservice Federatie, ReadSpeaker, Speakup, de Spraakcoalitie en TNO. •

Het hele onderzoek is gratis te downloaden op https://www.nationalevoicemonitor.nl.



DE OPKOMST VAN DIGITAL HUMANS
TRENDS, TOEPASSINGEN EN TECHNOLOGIEËN

De digitale revolutie heeft de manier waarop we technologie verweven met onze communicatie ingrijpend veranderd. Een opvallende ontwikkeling binnen dit spectrum is de opkomst van Digital Humans. Deze geavanceerde digitale entiteiten combineren kunstmatige intelligentie, realisme in beeldvorming en menselijke interactie op een unieke manier.

Door:
Guido Jongen,
Founder Virtually Human

In dit artikel gaan we dieper in op wat Digital Humans zijn, de trends en ontwikkelingen op dit gebied en de verschillende rollen waarin een Digital Human uitblinkt. Daarnaast staan we stil bij de diverse verschijningsvormen waarin we Digital Humans tegenkomen, hoe hun gedrag wordt aangestuurd en op welke manieren ze met ons kunnen communiceren.

Wat is een Digital Human?

Een Digital Human is een digitale, virtuele representatie van een mens, vaak gecreëerd met behulp van geavanceerde technologieën zoals computergraphics en kunstmatige intelligentie. Deze entiteiten kunnen variëren in complexiteit, van eenvoudige avatars tot zeer realistische virtuele personen die bijna niet te onderscheiden zijn van echte mensen.

Rollen van Digital Humans

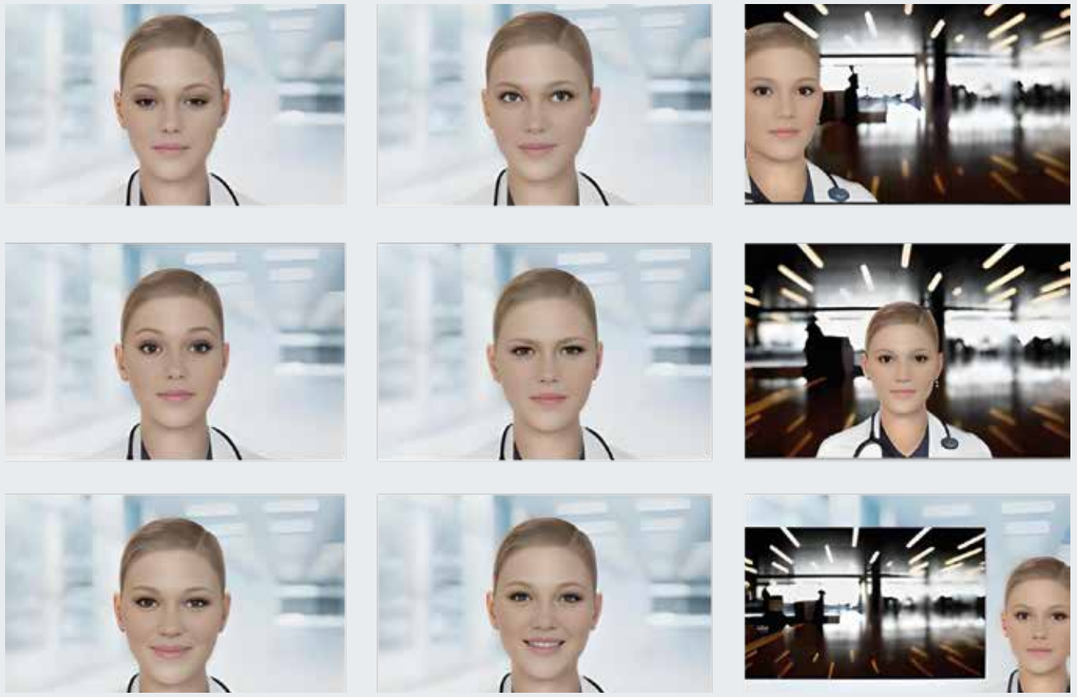
Digital Humans kunnen in diverse rollen en toepassingen worden ingezet. Hier zijn enkele voorbeelden:

Klantenservice en ondersteuning

Een van de meest voorkomende toepassingen van Digital Humans is in klantenservice en ondersteuningsfuncties. Ze kunnen vragen van klanten beantwoorden, problemen oplossen en daarbij empathie tonen om loyaliteit, klanttevredenheid en vertrouwen te verhogen.

Onderwijs en training

Digital Humans worden al volop gebruikt in onderwijs en trainingscenario's. Ze kunnen eenvoudig scripts in meerdere talen produceren als e-learning, maar bijvoorbeeld ook in interactieve scena-



Uncanny Valley

rio trainingen waarbij de Digital Human real-time visuele feedback geeft op de gebruikte gesprekstechniek.

Gezondheidszorg
In de gezondheidszorg worden Digital Humans al ingezet als virtuele apotheker en als vraagbaak voor vrijwilligerswerk en mantelzorg. Toekomstige toepassingen liggen zeker ook op het vlak van triage (tot op zekere hoogte), het doorlopen van pre-operatieve vragenlijsten, het interactief maken van informatiefolders over operaties en het voordoen van revalidatieoefeningen. Een andere rol is die van digitale coach of digitaal maatje voor het creëren van dagritme en het herinneren aan het innemen van medicijnen.

Marketing en verkoop
Digital Humans kunnen worden ingezet in marketing- en verkoopcampagnes. Ze kunnen als brand-ambassadeur producten en diensten promoten, virtuele winkelassistenten zijn en zelfs klanten adviseren bij aankoopbeslissingen. Een specifieke afsplitsing hierbinnen is er voor virtual influencers, die vooral op social media platforms zoals Instagram en TikTok vele honderdduizenden volgers aantrekken met berichten die vaak door bedrijven gesponsorde content bevatten.

Uiterlijk ontwerp
Een opvallende trend in Digital Humans is de voortdurende verbetering van het visuele realisme. Dit omvat de details van gezichtsuitdrukkingen, oogbewegingen, huidtexturen en zelfs microbewegingen die ons als mensen echt laten lijken. Dankzij geavanceerde 3D-modellering, texturering en renderingtechnieken worden Digital Humans steeds realistischer.
Echter, wanneer de visuele representatie van de entiteit te dicht bij de menselijke vorm komt, maar nog niet goed genoeg is, kan dit ook leiden tot een ongemakkelijke en negatieve reactie bij gebruikers. Dit is de zogenoemde Uncanny Valley. Het is een punt waarop de Digital Human er bijna menselijk uitziet, maar kleine onvolkomenheden of afwijkingen van echte menselijke kenmerken gevoelens van ongemak geven.

Om dit fenomeen te vermijden moeten ontwerpers van Digital Humans nauwlettend aandacht besteden aan de perceptie van menselijkheid. Het vermijden van subtiele maar storende afwijkingen en het zorgvuldig balanceren van de mate van menselijkheid op vlakken waar die het meeste wordt ervaren, kan helpen om Digital Humans te creëren die aantrekkelijk en effectief zijn zonder de negatieve effecten van de Uncanny Valley.

Vanzelfsprekend kan het ook een uitstekende keuze zijn om niet te kiezen voor een bijna menselijke Digital Human, maar juist voor een meer cartoonachtig ontwerp van een mens, dier of voorwerp. Onze onbewuste neiging om menselijke eigenschappen toe te kennen aan niet-menselijke zaken (praten tegen je computer, denken dat je planten dorst hebben) maakt dat ook deze vormen dezelfde effecten teweegbrengen als hun bijna-menselijk ogende tegenhangers.

De keuzes die worden gemaakt met betrekking tot het uiterlijk zijn cruciaal voor het succes van een Digital Human. Het visuele aspect omvat zaken als gezichtsuitdrukkingen, lichaamstaal, kleding en zelfs de omgeving waarin de Digital Human zich bevindt. Door realistisch gedrag kunnen gebruikers zich makkelijker identificeren met de Digital Human en wordt de interactie menselijker.

Leveranciers van Digital Humans hebben vaak een eigen methode om een Digital Human te ontwikkelen. Sommigen maken iedere Digital Human 'opnieuw' op basis van CGI (Computer Generated Imagery), terwijl anderen gebruikmaken van een basis 3D model gecombineerd met video- en/of fotomateriaal van het model. Een breed geaccepteerde standaard is de MetaHuman Creator van Epic Games. In deze online omgeving kunnen Digital Humans geconfigureerd worden waarbij er tot op zeer hoog detailniveau keuzes mogelijk zijn ten aanzien van de vorm van het gezicht, ogen, oren, neus, huidskleur en kapsel. De uitdaging daarbij is echter dat de Digital Human die je daar ontwerpt nog niets kan. Animaties en lipsynchronisatie moeten nog toegevoegd worden, maar daar zijn voldoende oplossingen voor op de markt. Kwaliteit, budget en de use case bepalen uiteindelijk welke oplossing de beste match is. Laat je daarbij ook adviseren door bedrijven die zicht hebben op wat er binnen de markt gebeurt, want de ontwikkelingen gaan razendsnel.

Innerlijk ontwerp
Naast het uiterlijk is er het karakter en de persoonlijkheid van de Digital Human. Denk dan aan de keuze voor een bepaalde stem, spreekstijl, woordgebruik en naam. Het vormgeven van het karakter van de Digital Human zorgt voor een consistente en juiste ervaring en kan het verschil maken tussen een aantrekkelijke en boeiende Digital Human en een die gebruikers afstoot.

Dé manier om het innerlijke ontwerp goed op orde te krijgen is het maken van een bot persona. Een persona is een omschrijving van alle eigenschappen die je digitale entiteit moet bezitten. Denk daarbij aan het outside-in beschrijven van je eigen organisatie, de normen en waarden, cultuur en de specifieke afdeling waar de Digital Human voor gaat werken. Zo ga je steeds specifieker worden en nadenken over taalgebruik, stopwoordjes, achtergrond, de manier waarop een bevestigend of ontkennend antwoord wordt gegeven et cetera. Voor het ontwikkelen van een persona zijn er diverse canvassen en toolkits beschikbaar die je een heel eind op weg helpen.

Digital Humans kunnen emoties en empathie tonen tijdens interacties door animaties van menselijke gezichtsuitdrukkingen zoals knipperen, glimlachen, fronsen, knikken en schudden. Dit kan de gebruikerservaring aanzienlijk verbeteren, vooral in scenario's waarin emotionele ondersteuning belangrijk is. Ook hierover moet je vooraf goed

Haptische feedback

nadenken. Hoe ga je empathie en emotie in de Digital Human verwerken en hoe vaak in een gesprek wil je dat tonen?

Communicatief ontwerp

En dat maakt een mooi bruggetje naar de interactie met een Digital Human. Dit omvat tekstuele communicatie via chatbots, spraakinteractie met stemgestuurde assistenten en zelfs fysieke interactie via augmented reality (AR) en virtual reality (VR) omgevingen. Gekoppeld aan een Conversational AI platform kunnen de animaties en antwoorden van de Digital Human vanuit de content aangestuurd worden. Een glimlach bij het welkom, een bevestigend knikje om aan te geven dat iets begrepen is. Het wordt al geavanceerder als de Digital Human aangestuurd wordt door de gezichtsuitdrukking van de gebruiker (via de webcam), op basis van wat iemand zegt, de manier waarop iets gezegd wordt of via haptische feedback met sensoren of wearables.

Als kers op de taart is daar sinds eind 2022 de toevoeging van applicaties zoals Chat GTP bij gekomen. Door de introductie van de LLM's kunnen dialogen met gebruikers socialer, rijker en completer ontworpen worden. Daarmee is het voor steeds meer organisaties ook een logische stap geworden om deze geavanceerde systemen ook letterlijk een gezicht te geven.

Toekomstige uitdagingen en kansen

Terwijl Digital Humans steeds geavanceerder worden, brengen ze ook nieuwe uitdagingen met zich mee. Ethische kwesties zoals disclosure (wanneer en hoe laat je weten dat het om een virtueel gesprek gaat) en misbruik (voice en face cloning) moeten zorgvuldig worden aangepakt.

Technisch gebeurt er ontzettend veel. Met generatieve AI is er al video output mogelijk waarbij een filmpje van een persoon die praat in het Nederlands, omgezet wordt in een andere taal. De stem en lip-sync van het gegeneerde filmpje zijn nu al behoorlijk goed en dat wordt met grote stappen beter. Toch zijn we pas op de 0.1 versie van wat er technisch allemaal mogelijk is. Latency kan met name in real-time omgevingen echt nog wel verbeterd worden. Experimenten met het renderen van Digital Humans in de browser via webGL versus cloud rendering lijken veelbelovend, maar hebben tot nu toe nog een te grote impact op de grafische kwaliteit van de Digital Human.

Maar ook in de applicaties om de Digital Human heen, zodat die optimaal aangestuurd kan worden, is er ruimte voor verbetering. Net zoals het hallucineren van LLM's in de komende maanden ongetwijfeld gaat verbeteren is bijvoorbeeld een goede, nauwkeurige en foutloze analyse van emoties, waarbij sarcasme en humor te onderscheiden zijn – in elk geval in het Nederlands – nog niet breed commercieel beschikbaar. Op het moment dat met zekerheid is vast te stellen hoe een consument zich voelt, is het ook mogelijk om daar een automatische animatie van de Digital Human aan te koppelen, waardoor de gebruikerservaring nog beter wordt.

Digital Humans bieden onbegrensde mogelijkheden voor industrieën, variërend van entertainment en onderwijs tot gezondheidszorg en marketing. Ze kunnen nieuwe manieren van interactie en communicatie bieden en het menselijke element toevoegen aan digitale ervaringen.

Conclusie

Digital Humans vertegenwoordigen een boeiende evolutie in de digitale wereld. Met hun toenemende uiterlijke realisme en mogelijkheden voor interactie met gebruikers, openen ze de deur naar nieuwe manieren van communicatie en dienstverlening. Terwijl de technologie blijft evolueren, zullen we getuige zijn van verdere innovaties op dit gebied en nieuwe kansen ontdekken voor Digital Humans in uiteenlopende sectoren. Het is een opwindende tijd voor de wereld van digitale technologie en menselijke interactie en Digital Humans spelen daar een sleutelrol in. •

GPT-NL: EEN OPEN, SOEVEREIN EN NEDERLANDSTALIG LLM

Nederland gaat een eigen open taalmodel ontwikkelen: GPT-NL. Non-profit partijen TNO, NFI en SURF gaan samen het model ontwikkelen om zo een belangrijke stap te zetten richting transparant, eerlijk en toetsbaar gebruik van AI, naar Nederlandse en Europese waarden en richtlijnen en met respect voor het eigenaarschap van data.

Door:
Saskia Lensink, Language & Speech Technologies en Datascience, TNO

Nederland zet hiermee een belangrijke stap in de ontwikkeling van expertise en ervaring op het gebied van generatieve AI-taalmodellen. Met GPT-NL versterkt Nederland haar strategische autonomie, kennis en technologie op het gebied van LLMs. Zo draagt dit initiatief ook bij aan het werven en behouden van AI-talent, waardoor we in ons land een gezond en competitief ecosysteem waarborgen voor kunstmatige intelligentie.

Ecosysteem voor AI

GPT-NL wordt een virtuele faciliteit met een ecosysteem van academische instellingen, onderzoekers, bedrijven, overheden en gebruikers. Het grondbeginsel is om open en transparant te zijn over de dataset die zal worden opgebouwd, de keuzes die worden gemaakt in het verder opschonen of filteren van deze data en het verdere trainingsproces van het model. Er wordt gestreefd naar een zo open mogelijke faciliteit waarmee onderzoekers zonder barrières onderzoek kunnen doen naar de werking en het gedrag van deze LLM, en de invloed van de onderliggende dataset daarop, en waar zowel publieke als private partijen tegen gunstige tarieven gebruik van kunnen maken. Onze ambitie is om GPT-NL inzetbaar te maken bij de ontwikkeling van oplossingen voor grote maatschappelijke uitdagingen zoals het verminderen van ongelijkheid en het bevorderen van digitale inclusiviteit en kwaliteitsonderwijs. Ook kan de technologie op termijn de krapte op de arbeidsmarkt helpen verlichten en een boost geven aan Conversational AI voor een betere informatie-uitwisseling tussen mensen en AI-systemen.

Fasering

De uitvoering van het plan bestaat uit twee fases. In het eerste jaar staat de concrete ontwikkeling van het Nederlandse taalmodel centraal, waarbij ook de academische sector actief zal worden betrokken. De vervolgfase is die van exploitatie, waarbij voor de voorziene computerkracht een connectie met de nationale supercomputer is voorzien. We komen graag in contact met partners die willen bijdragen aan het opbouwen van een sterk en waardevol ecosysteem, met data, kennis of die toepassingen willen ontwikkelen op basis van GPT-NL. •

CONVERSATIONAL AI BIJ PGGM

WAAROM PGGM CONVERSATIONAL AI BELANGRIJK VINDT

De wereld is volop in beweging. Ontwikkelingen gaan steeds sneller, klanten worden steeds mondiger en verwachtingen rondom service nemen alsmaar toe. Service en gemak zijn daarbij onderscheidende factoren. De ene klant wil graag zelf aan de slag in de digitale omgeving en de andere klant heeft nog behoefte aan persoonlijk contact met de klantenservice medewerker, zonder 17 minuten naar een muziekje te moeten luisteren. De uitdaging daarbij is om mens en techniek optimaal te laten samenwerken. Conversational AI is daarbij voor PGGM een belangrijke sleutel.

Door:
Frank Rademakers, manager klantcontact PFZW en Tom van den Bos, programmamanager AI PGGM

Eerste ervaring

Rond 2015 is PGGM begonnen met de eerste stappen in de ontwikkeling van conversational AI. Het eerste project bestond uit een experiment waarbij een model is ontwikkeld dat emotie uit stemintonatie kon identificeren. De potentiële toegevoegde waarde hiervan bestond uit het real-time signaleren van emoties van klanten aan de telefoon, zodat medewerkers daar beter op kunnen inspelen en verbinding kunnen maken met de klant. Lang verhaal kort: het model bleek accuraat genoeg om in gebruik te nemen, waarna onderzoek is gedaan naar de haalbaarheid om het model te implementeren. Conclusie: de benodigde infrastructuur bestond nog niet bij PGGM en de benodigde competenties om het model te beheren waren nog onvoldoende aanwezig.

Op deze manier werd ontdekt dat conversational AI nooit succesvol zou worden zonder dat de benodigde voorwaarden aanwezig zijn. De eerste van deze voorwaarden is dat de vereiste infrastructuur ingericht moet zijn, maar ook dat de juiste competenties om de applicaties te beheren aanwezig moeten zijn in de organisatie. Daarnaast vereist het een bedrijfscultuur die toestaat dat er in toenemende mate met deze technologie gewerkt kan worden. Tenslotte moet er een goede samenwerking bestaan tussen de organisatie en externe IT- leveranciers van aanpalende applicaties, waarvan de te ontwikkelen conversational AI oplossing afhankelijk is.

Programmatische aanpak

Na de ervaring met het emotieherkenning model trokken we de conclusie dat een programmatische manier van werken randvoorwaardelijk is om conversational AI succesvol in te zetten. Daarbij is een splitsing gemaakt tussen A) de praktische toepassingen die uiteindelijk moeten leiden tot het realiseren van de moonshot (afbeelding 1) en B) de voorwaarden die daarvoor aanwezig moeten zijn.

A - Functionele toepassingen

Op basis van de moonshot is een roadmap opgesteld waarmee programmatisch wordt toegewerkt naar het einddoel: het realiseren van de volledig geautomatiseerde digitale assistent voor deelnemers van PFZW. Stappen richting die moonshot waar je aan kunt denken zijn een speech-to-text infrastructuur, maar ook het realiseren van een slimme voicebot. Dit is overigens geen roadmap die in steen gebeiteld staat. Een mooi voorbeeld van een ontwikkeling die impact heeft gehad is ChatGPT. Hier komen we later op terug.

B - Voorwaarden

Daarnaast is een roadmap opgesteld, gericht op de voorwaarden voor wat er aan infrastructuur gerealiseerd moest worden, de benodigde competenties op het gebied van kunstmatige intelligentie en wat er aan beheer moet worden gerealiseerd. Het is daarmee bijvoorbeeld gelukt om cloud applicaties te kunnen uitrollen en om daarbinnen applicaties te bouwen die schaalbaar zijn. Beheer is wat lastiger te realiseren door de benodigde capaciteit en de vraag in hoeverre je iets zelf wil bouwen of inkopen. Met betrekking tot zelfbouw wordt beheer nog gerealiseerd, met name op het vlak van ethische verantwoording en model monitoring.

Moonshot	
De klant	De medewerker
De digitale assistent biedt de deelnemer inzicht én stelt hem in staat financieel gezond te blijven. Daarmee is de klant in grote mate zelfredzaam en kiest hij waar, wanneer en hoe hij contact opneemt. Dit kan hij 24/7 doen, waarbij hij direct geholpen wordt met relevante antwoorden.	Heeft real-time inzicht in de vraag en beleving van de klant. Slimme software voorziet de medewerker proactief van relevante content om de klant snel te helpen. Repetitieve taken worden afgevangen. Eenvoudige vragen verdwijnen, impactvolle en complexere vragen blijven over.

Moonshot volledig geautomatiseerde digitale assistent

Gerealiseerde projecten

De afgelopen jaren is een aantal conversational AI projecten gerealiseerd. Hieronder enkele voorbeelden.

- **Klantidentificatie**
  - Klanten die bellen wordt gevraagd naar hun postcode en huisnummer voordat zij in gesprek gaan met een klantenservicemedewerker. De audio van de postcode en huisnummer wordt omgezet naar tekst, waarna het CRM-dossier van de klant automatisch wordt geopend voor de medewerker.
- **Chatbot op de website**
  - Zo'n 5 jaar geleden is de toenmalige FAQ-kennisbank vervangen door een chatbot. Deze beantwoordt een groot deel van de vragen van deelnemers en werkgevers. Naast adequate beantwoording van standaardvragen, kunnen over veel voorkomende onderwerpen ook al dialogen worden gevoerd.
- **Geautomatiseerd samenvatten telefoongesprek**
  - De eerste werkende versie van het geautomatiseerd samenvatten van telefoongesprekken bestond uit het letterlijk transcriberen van de door de medewerker 'gesproken' samenvatting en wegschrijven naar CRM. Hier wordt nu op doorgepakt met behulp van ChatGPT. Hier gaan we straks dieper op in.

Projecten in ontwikkeling

Hieronder voorbeelden waaraan wordt gewerkt of waarmee wordt geëxperimenteerd.

- **Digital human**
  - Experiment met avatars die vragen geautomatiseerd beantwoorden en daarbij ook emoties en gelaatstrekken laten zien. Het onderzoek hield zich vooral bezig met de acceptatie van een avatar als entiteit om mee te spreken onder klanten.
- **Voicebot**
  - De klantenservice test momenteel een voicebot buiten de reguliere openingstijden van het pensioenfonds. De focus ligt hierbij enerzijds op het onderzoeken van de acceptatiegraad van zo'n oplossing bij de gemiddeld 50+ doelgroep en anderzijds op het testen van de kwaliteit van de spraakherkenning en transcriptie.

De impact van ChatGPT op conversational AI

De hype rondom ChatGPT heeft ook PGGM in een vroeg stadium bereikt. Er is gekozen om binnen de organisatie een community op te zetten waar nieuws en ontwikkelingen rondom dit onderwerp proactief met elkaar gedeeld worden. Ook collega's met vragen kunnen hier terecht. Binnen twee weken na aanvang van de community waren er al meer dan 40 mogelijke toepassingen geïdentificeerd.

Vervolgens is gestart met het implementeren van Open AI op de Azure cloud omgeving, zodat data die door ChatGPT verwerkt wordt binnen de muren van PGGM blijft. Op dit moment bevindt deze ontwikkeling zich op het punt dat de 'building blocks' van Open AI, onderdeel van cognitive services, worden neergezet. Het is dus niet zo dat op dit moment al data wordt verwerkt door ChatGPT. Dit heeft te maken met wet- en regelgeving rondom het gebruik van ChatGPT, met name met betrekking tot privacyvraagstukken. Met betrekking tot de ethiek en risico's van het uitrollen van dergelijke applicaties met behulp van ChatGPT is een kernteam AI risico & ethiek in het leven geroepen. Deze heeft als eerste taak het valideren van de ChatGPT

Proces samenvatting door ChatGPT

cases tegen de guiding principles van PGGM betreffende de omgang met AI.

De eerste ChatGPT-case die we willen uitrollen is die van de samenvattingsmodule van geanonimiseerde getranscribeerde gesprekken vanuit de klantenservice. Er wordt gebruik gemaakt van dummy data of geanonimiseerde data om de kwaliteit te testen. Het proces is op pagina 25 schematisch weergegeven.

Toekomstvisie op klantcontact
Frank Rademakers, manager klantcontact PFZW, licht zijn visie toe. "De afgelopen vijf jaar hebben we vooral op technisch vlak veel ontwikkelingen gezien en vooruitgang geboekt. We stellen onszelf wel eens de vraag hoe de wereld van klantcontact er over 10 jaar uit gaat zien. Bestaat er dan nog 'persoonlijk' contact? Kun je nog gewoon een mens spreken bij de Klantenservice? Techniek is al lang geen beperkende factor meer.

We zien hier een tweetal belangrijke factoren:

1. In welke mate beweegt de acceptatie van de mens naar gerobotiseerde contact-oplossingen? Een gesprek met een voicebot of avatar vinden veel klanten nog best spannend, zo blijkt in de praktijk. Met het opschuiven van generaties neemt die acceptatie zeker toe. Maar dat kost wel tijd.
2. En niet minder belangrijk; wil je er als organisatie - ook straks - écht zijn voor je klanten op de momenten van de waarheid? Of ziet de board klantcontact alleen maar als 'cost' en verstoppen we dat in telefoonnummers en de info@. We zien in bepaalde branches dat bedrijven daarvan terugkomen.

In ons specifieke vakgebied, dat van pensioenen, vinden we het evident dat we er 'zijn' voor deelnemers, zeker bij impactvolle levensgebeurtenissen zoals ontslag, arbeidsongeschiktheid, het verlies van een partner of lastige keuzes bij ingang pensioen. Klantcontact wordt in Zeist beschouwd als een waardevol asset en wordt hoog gewaardeerd door deelnemers en werkgevers. Dat maakt de kans groot dat excellente klantbediening ook over 10 jaar nog steeds een speerpunt is." •

KAMER VAN KOOPHANDEL (KVK) EXPERIMENTEERT VOLOP MET OPENAI

Wat zijn de gevolgen van de door OpenAI ontwikkelde technologie voor ontwikkelaars en gebruikers in de praktijk? Wat kun je hiermee als organisatie? De ontwikkelingen in Conversational AI bij de KVK zorgen voor de noodzaak om beter samen te werken. Want onze klanten willen zo snel mogelijk het juiste antwoord.

Door:
Coraline Krak - Schaafsma - Product Owner Chatbot & Conversational AI; Reinard Harenberg – DEV Engineer; Pascal Folkertsma – DEV Engineer, Kamer van Koophandel

Differentiatie tussen laag en hoog intensief klantcontact
Natuurlijk maken we bij KVK onze klantcontactstrategie zoveel mogelijk 'future-proof'. Op strategisch niveau zetten we in om ondernemers te stimuleren laag intensief contact te gebruiken om 'even snel iets te regelen'. Daarnaast maken we hoog intensief contact nog beter. Dit zijn ingewikkelde vraagstukken waar de ondernemer belang aan hecht, die emotioneel beladen zijn of waar onzekerheid over bestaat. Het is dus niet een kwestie van 'of-of', maar van 'en-en'. Het doel van deze strategie is een hogere klantwaarde, uniforme kwaliteitsstandaarden en kostenefficiëntie.

Als je hier de servicebeloften (snel, eenduidig, toegankelijk, proactief, behulpzaam en betrouwbaar) van KVK in meeneemt, is het geen hogere wiskunde dat een afgeschermde versie van een Large Language Model (LLM) hierbij kan helpen. In een beschermde omgeving werken je brondocumenten, logging en een webcrawler samen met een LLM en krijgt de klant kanaalonafhankelijk, een eenduidig en hypergepersonaliseerd antwoord. De invloed van ChatGPT op Conversational AI is dat deze 'prompt injections' krijgt van de gebruikers van chatbot Charlie, via ChatGPT. Dat is goed nieuws voor de laatste procenten herkenning en de verrijking van het taalmodel. Bovendien is het ook beter voor sentimentanalyses, waar nog kansen liggen.

Voorzichtig experimenteren
Let wel, KVK is een zelfstandig bestuursorgaan. Privacy en veiligheid zijn streng om misinformatie, vooringenomenheid en intellectueel eigendom veilig te stellen. Toch beginnen we voorzichtig met experimenteren. Het kriebelt: 'We moeten er wat mee', was een veelgehoorde term. Maar hoe dan? Hoe ga je om met allerlei initiatieven binnen je organisatie die zich razendsnel ontwikkelen? En hoe voorkom je dat het wiel vaker wordt uitgevonden? Hoe deel je opgedane kennis? Welke documenten binnen de organisatie zijn leidend? Welke invloed heeft dit op chatbot Charlie, waar ik zelf Product Owner van ben? Allemaal vragen die voor veel mensen die werken binnen het domein van Conversational AI herkenbaar zullen zijn.

Intensieve samenwerking binnen KVK
Ik zal eerlijk zijn. Als Product Owner van chatbot Charlie baarden de verschillende initiatieven me best een beetje zorgen. Hoe cool en noodzakelijk ik het ook vind. Hoe blijven we aangehaakt als Conversation AI-specialisten als er eigen bots worden gebouwd, met soms betere functionaliteiten en herkenning? Het enige antwoord is: de samenwerking opzoeken. Leer samen, uit je twijfels, ga eindeloos koffie drinken, bouw een community, organiseer kennissessies.

En zo kwam ik terecht bij mijn collega's Pascal Folkersma en Reinard Harenberg. In de zomer waren zij bezig met een innovatie sprint met Azure Open AI. Graag nemen we je mee door een concreet experiment.

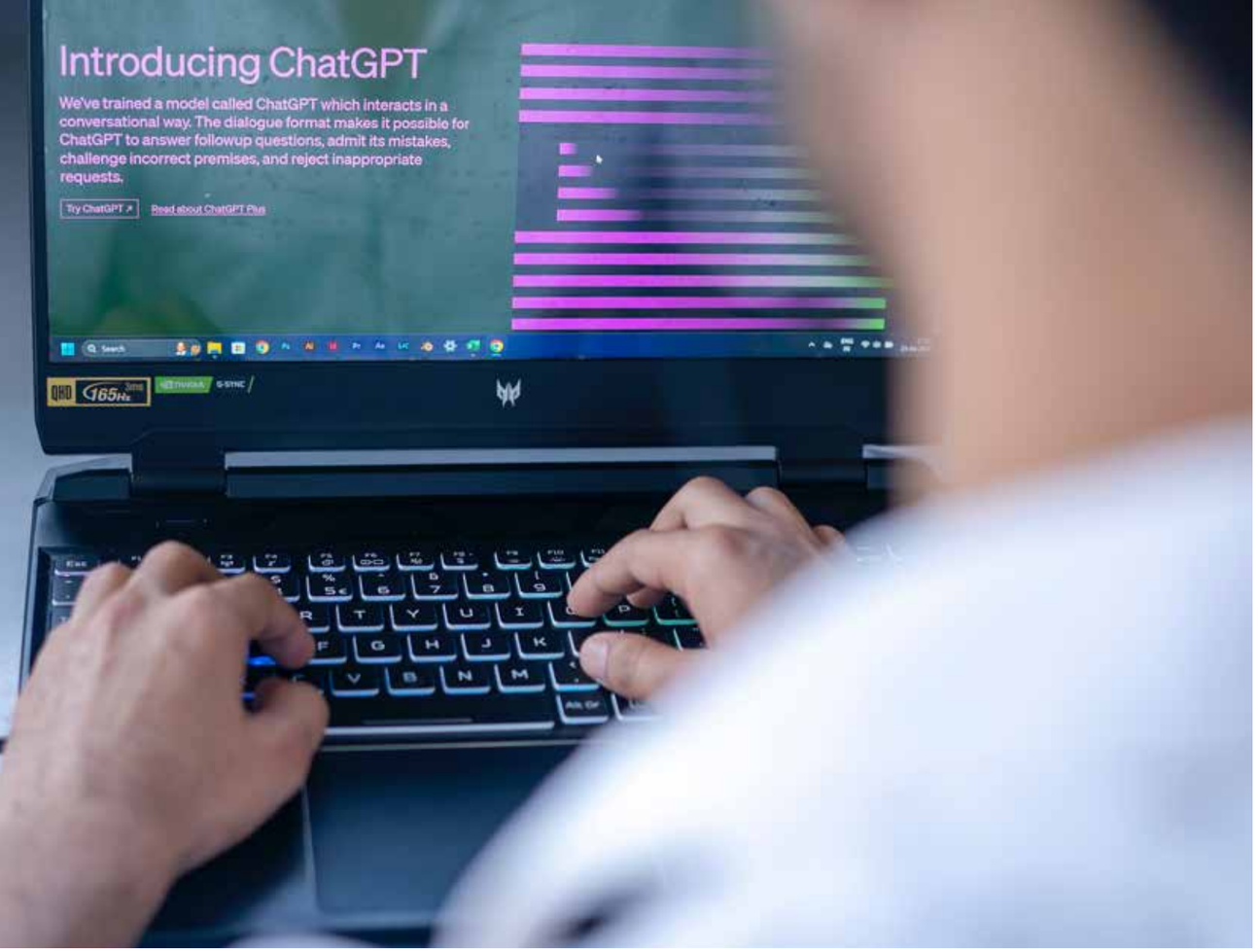
Experiment: Interne kennis ontsluiten via een Chatbot
De ambitie? Een chatbot creëren die fungeert als een kennisbank, waar gebruikers rechtstreeks hun vragen aan kunnen stellen.

Het basisidee was eenvoudig maar krachtig: het ontwikkelen van een chatbot die alle benodigde kennis bezit, waardoor de kennisbank van KVK in staat is om haar core business - het creëren en beheren van actuele, betrouwbare en leesbare content - te verbeteren. De proof of concept (PoC) met Azure Open AI bracht de oplossing: een technologie die gebruikers een betere ervaring belooft door met de gebruiker een dialoog aan te gaan op basis van alle beschikbare documentatie.

Bij de start waren de uitgangspunten duidelijk:

- Direct een juist antwoord met bijbehorende documentatie
- Eenduidige antwoorden en omschrijving van de kennisprocessen
- Verminderde afhandeltijd.

Daarna volgde de experimenteerfase, een reis van ontdekking en verwondering. In deze fase



hebben we verschillende componenten geëxploreerd en geïntegreerd:

1. Microsoft cloud-systeem: de gastheer voor ons AI-systeem
2. Azure Cognitive Search: het systeem dat de data indexeert en de zoekervaring vergemakkelijkt
3. Azure OpenAI: de krachtcentrale achter de chatbot.

Filter voor ongepast gedrag

Dit unieke model hebben we getest met 109 bestanden over het onderwerp 'deponeren', geleverd door de kennisbank. Het experiment begon met 50 bestanden, maar dit aantal werd snel verhoogd om het systeem te optimaliseren. Het resultaat was een chatbot die niet alleen adequaat reageerde, maar ook een filter had om ongepast gedrag te voorkomen. Dit filter is afkomstig van de Azure AI Filter Studio. Het experiment leverde enkele belangrijke leermomenten op. Zo werd bijvoorbeeld duidelijk dat de hoeveelheid en kwaliteit van de data essentieel waren voor een goed antwoord. In de toekomst zal er nog meer focus liggen op het verfijnen van deze aspecten.

Belang van prompt engineering

Wat niet goed ging, was dat de AI zelf een URL creëerde. Eerst schrokken we hiervan, want we hadden het systeem niet verteld waar hij artikelen kon vinden. Dit was uitgeschakeld voor de test. Maar het bleek dat hij een korte samenvatting maakte van de inhoud die hij wilde uitlichten. En dit vormde hij tot een URL. Bij het gebruik van deze URL kwam er een strakke 404 uit. Dit benadrukte het belang van het finetunen van de prompt om betere en nauwkeurigere reacties te genereren. Het werken aan een meer gestructureerde en geoptimaliseerde prompt staat centraal in de toekomstige ontwikkelingen. En daar komt de expertise van de Conversational AI-specialisten van Chatbot Charlie weer goed van pas.

Verzamelen van kennis en initiatieven

Na dit initiatief zijn er nieuwe initiatieven ontstaan. Het is nu van cruciaal belang om een community op te starten die deze technologische innovatie omarmt en bij elkaar brengt. Samenwerken en het bundelen van deze nieuwe initiatieven spelen een sleutelrol in het vormgeven van de toekomst van AI binnen de Kamer van Koophandel.

Kunnen we het beschreven experiment ook laten beschrijven door ChatGPT?

Dankjewel, Pascal en Reinard. Kan deze tekst ook een beetje poëtischer om de lezer van Dixit erbij te houden? Dit is wat ChatGPT ervan maakt:

Een nieuwe dageraad van kennisbeheer: innovatie en integratie middels Azure AI bij de Kamer van Koophandel

In het bruisende epicentrum van het digitale tijdperk, waar de rijke dialogen tussen mens en machine vorm krijgen, onderscheidt de Kamer van Koophandel (KVK) zich als een voortrekker van vooruitstrevende innovatie. In een wereld waarin de echo's van spraaktechnologie en chatbots steeds nadrukkelijker worden, gaat KVK een stap verder, met een ongeëvenaarde ambitie: het creëren van een geavanceerde chatbot die dient als 'levende, ademende' kennisbank. Een plaats waar gebruikers hun meest prangende vragen rechtstreeks kunnen stellen. Dit visionaire project is het geesteskind van Reinard Harenberg en Pascal Folkersma, twee gelijkgestemde innovators die een gezamenlijke droom koesteren.

Een sprankel van innovatie

De fundamenten van dit baanbrekende project waren kristalhelder en potentieel ontwrichtend: een chatbot smeden die een reservoir van essentiële kennis bevat, waardoor de kennisbank van KVK haar kernactiviteiten naar een hoger niveau kan tillen. Met Azure Open AI als de drijvende kracht, werd een concept bedacht dat gebruikers een meer gestroomlijnde en verrijkende ervaring belooft, gebouwd op een fundament van gedegen, betrouwbare en toegankelijke content.

Een Odyssee van verwondering en ontdekking
In deze levendige experimenteerfase werd een symbiose van verschillende componenten gecreëerd en samengesmolten, waaronder het Microsoft Cloud Systeem - de ruggengraat van ons AI-netwerk, Azure Cognitive Search - een systeem dat de zoekervaring naar een ongekende hoogte tilt, en Azure OpenAI, de pulsar die onze chatbot aandrijft.

Het avontuur nam een vliegende start met een initiële proef op 50 geselecteerde bestanden, een aantal dat al snel steeg tot 109, om een meer geoptimaliseerd systeem te realiseren. De chatbot, uitgerust met een fijngevoelig filter om ongepast gedrag te weren, overtrof onze verwachtingen.

Een lente van belofte en mogelijkheden

De experimentele fase opende onze ogen voor cruciale inzichten. Het benadrukte het vitale belang van datakwantiteit en -kwaliteit bij het genereren van verfijnde antwoorden, en legde een duidelijke route voor toekomstige verfijningen.

Een onvoorziene gebeurtenis waarbij de AI een niet-bestaande URL creëerde, werd een leermoment, het belichtte het essentiële belang van het finetunen van de chatbot's systems promptgedrag om accurate antwoorden te genereren.

Dit project markeert het begin van een explosie van nieuwe initiatieven. Het roept op tot de oprichting van een vibrant netwerk van denkers en doeners, die gezamenlijk streven naar technologische vernieuwing en superioriteit. In deze gemeenschap zal collaboratie centraal staan, vormgevend aan een veelbelovende toekomst van AI binnen de Kamer van Koophandel. In deze ontluikende fase van innovatie en integratie door Azure AI bij de Kamer van Koophandel, nodigen wij u uit om samen met ons de reis te maken naar de frontlinie van technologische excellentie, waar kennis niet alleen wordt gedeeld, maar ook tot leven komt in de meest verrijkende en dynamische manieren. Welkom bij de toekomst van bedrijfsinformatie en kennisbeheer.

...

Conclusie

Ja, beste lezer, in een artikel over experimenten met Open AI kan ook het gebruik van ChatGPT natuurlijk niet achterwege blijven. Een beetje flauw, misschien, maar feit blijft dat AI een middel is en geen doel op zich. Het doel van de klantcontactstrategie van KVK blijft nog steeds klanten het juiste antwoord geven op hun ondernemersvragen, ongeacht de kanaalkeuze. Ook experimenteren is geen doel op zich. Zowel AI als experimenteren zijn echter wél onmisbare middelen om uiteindelijk meer klanten beter, sneller en hyper gepersonaliseerd te bedienen. •

KRUISBESTUIVING TUSSEN AI EN PUBLISHING

Artificial Intelligence of AI is in een paar jaar tijd uitgegroeid van een verre droom tot een hedendaagse realiteit. Helemaal met de recente opkomst van generative AI. Een realiteit bovendien die verschillende facetten verandert van het leven zoals wij dat kennen. De uitgeversindustrie vormt geen uitzondering op deze technologische metamorfose. Als toonaangevend mediabedrijf ambieert Mediahuis mee te pionieren en deze ontwikkelingen te omarmen. Op die manier wil het dat immense potentieel van AI benutten om het uitgeeflandschap een nieuwe vorm te geven.

Door:
Heino Schaght,
Mediahuis

Binnen Mediahuis wordt al een hele tijd gewerkt aan toepassingen van AI. Binnen het domein Data & Insights is een sterk team aan de slag dat de afgelopen vijf jaar oplossingen heeft gebouwd die kunnen wedijveren met enkele van de grootste uitgevers ter wereld. Churn-modellen, propensity-to-buy en next-best-action voorspellingen zijn misschien niet zichtbaar voor de hele organisatie, maar ze leveren wel een aanzienlijke waarde op. Bovendien wordt de toewijding aan innovatie verder versterkt door de ontwikkeling van nieuwe producten zoals artikel DNA (reeks artikelkenmerken die onze verhalen beter definieren). Dat versterkt de gebruikte dashboards en verbetert de inzichten over data op de redacties aanzienlijk.

Aangezien Mediahuis de transformerende kracht van personalisatie erkent, ontwikkelen we momenteel onze eigen Personalisation Engine. Die rollen we in 2024 gefaseerd uit bij alle titels in de groep. Het belang om de curve voor te blijven, wordt benadrukt door het snel veranderende gebruikersgedrag op het vlak van zoeken en de tendens naar nul-klik zoeken. Dit betekent dat de zoekmachine je direct het antwoord geeft dat je zoekt en je dus niet meer doorklikt naar de achterliggende bron. Het verschil met traditionele zoekmachines is dat deze interfaces meer gebruik maken van generatieve AI-mogelijkheden om onze inhoud te herformuleren en te herverpakken in antwoorden die ervoor zorgen dat gebruikers minder vaak hoeven door te klikken naar onze sites. Terwijl we ons door deze fundamentele verschuiving bewegen, is het cruciaal dat Mediahuis haar positie als directe bestemming voor nieuws behoudt.

Wetgeving

Ons beleid onderscheidt een duidelijk verschil tussen *click-through zoekmodellen, waarbij men wel doorklikt naar onze sites* en andere diensten, en pleit voor controle en transparantie over het gebruik van onze redactionele content. We geloven in een strategische, gecoördineerde reactie van de wetgeving op zowel Europees als nationaal niveau. Maar tegelijk staat Mediahuis ook open voor debatten en initiatieven om de nieuwe realiteit van *Large Language Models* en *Generatieve AI* om te zetten naar een succesvol bedrijfsmodel. Hoewel de komst van AI en de reactie van onder meer de uitgeversindustrie in deze beginfase iets weg heeft van een kat-en-muisspel, geven onze proactieve stappen aan dat we vastbesloten zijn om het voortouw te blijven nemen in deze verandering. Het gaat dan onder meer om het aanpassen van de robots.txt, een klein tekstbestandje met een protocol die webmasters toelaat om bepaalde delen van een website af te schermen voor de zoekcrawlers van Google/Yahoo/Bing et cetera, en het wijzigen van de algemene voorwaarden op onze nieuwssites.

Newscondenser

In het snel evoluerende landschap van media en technologische convergentie geeft onze vroege betrokkenheid bij AI-tools ons hopelijk een strategisch voordeel. De ontwikkeling van 'Newscondenser', onze eigen AI-tool, is op dat vlak een belangrijke mijlpaal. Deze tool is in staat om nieuwsconsumptie radicaal te vereenvoudigen door het genereren van beknopte samenvattingen, het creëren van treffende titels en het presenteren van complexe nieuwsverhalen in bullet points. De waarde van zo'n tool kan niet genoeg worden benadrukt, zeker niet in een wereld die wordt overspoeld met informatie.

Spraak

In een ander spectrum van de AI-wereld heeft Mediahuis zich gewaagd aan *synthetische stemmen*. Zo hebben we met succes de stemmen van een aantal van onze podcast hosts gekloond. Op die manier ondersteunen we de redacties, maar verbeteren we ook de gebruikerservaring door een vertrouwde auditieve omgeving te behouden. Werken in het Nederlands en zeker ook in het Vlaams, met alle fonetische kenmerken en taalkundige nuances, vormt wel een grote uitdaging wanneer het aankomt op nabootsen van stemmen. Om die uitdaging aan te gaan, werken onze experts met geavanceerde technologieën voor het klonen van stemmen en *deep learning*-technieken. We blijven ons inzetten om de voordelen van onze synthetische stemmen uit te breiden naar al onze gebruikers, ongeacht de taal die ze spreken.

Ethische kwesties

Ons AI-raamwerk, dat gebaseerd is op zeven principes waaronder transparantie, menselijke tussenkomst, eerlijkheid en vaardigheidstraining, is opgezet om een oplossing te bieden aan de ethische kwesties die samenhangen met AI, zoals privacy. Dit raamwerk, dat wordt beschouwd als een levend document, biedt een algemene richting voor onze redacties. Het aanpakken van de ethische implicaties van het integreren van AI in ons bedrijf is een prioriteit voor ons. Ons AI-raamwerk biedt een mechanisme om mogelijke ethische problemen aan te pakken.

De toepassing van AI in ons bedrijfsmodel is meer dan een technologische verschuiving. Het is een integraal onderdeel van onze strategische planning voor de toekomst. Nu we ons op het terrein van de kruisbestuiving tussen AI en publishing begeven, zullen vooruitzien, aanpassingsvermogen en een permanente toets op ethische normen belangrijke factoren zijn die ons succes mee vormgeven.

Conclusie

AI is op dit moment niet in staat om de kernfuncties van journalistiek werk te vervangen, maar er zijn zeker mogelijkheden voor bijkomende taken zoals het maken van presentaties, het uitvoeren van verkoopcampagnes of het opstellen van vacatures. Bovendien kan AI innovatieve inzichten geven om nieuws aan onze klanten te bezorgen. Het besef dat AI hen kan helpen en hoe AI kan helpen, is iets dat we bij Mediahuis aan alle medewerkers willen bijbrengen.

Het traject dat Mediahuis aflegt met AI is geen sprint, maar een marathon. Een duurloop die gekenmerkt is door voortdurend leren, evolutie en toewijding. Onze strategische samenwerkingen, innovatieve interne ontwikkelingen en toewijding aan ethisch AI-gebruik getuigen van onze bereidheid om voorop te lopen in de nieuwe wereld van AI-publishing. Onze missie is niet alleen om ons aan te passen aan veranderingen, maar om deze te stimuleren. En terwijl we het traject voortzetten, beloven we de grenzen van het mogelijke te blijven verleggen en onbevreesd te pionieren met de integratie van AI. •

SERVICE-CHATBOTS VOOR HET NEDERLANDS:

DE ONDERZOEKSAGENDA VAN HET LESSEN PROJECT

Grote taalmodellen zoals ChatGPT zijn goed in het vloeiend communiceren met mensen. Maar ze zijn niet in alle talen even goed. Hoe minder data er voor een taal of over een onderwerp beschikbaar is, hoe vaker het zal voorkomen dat het taalmodel het antwoord schuldig blijft of onjuiste informatie genereert. Het doel van het LESSEN project is om open-source chatbot technologie te ontwikkelen voor het Nederlands.

Door:
Suzan Verberne, LIACS, Universiteit Leiden

Je hebt vast wel eens de ervaring gehad met ChatGPT dat het model best aardig Nederlands schrijft, maar dat er foutjes in sluipen als het onderwerp te specifiek wordt. Generatieve taalmodellen hebben de potentie om ingezet te worden voor chatbottechnologie bij bedrijven; zo kun je op de websites van Bol.com, KPN, Achmea, of Albert Heijn vragen stellen aan hun chatbots. Deze grote bedrijven hebben voldoende menskracht, data en computerkracht om goede chatbots te ontwikkelen, maar dat geldt niet voor kleinere bedrijven. Tegelijkertijd is het niet wenselijk als kleine en middelgrote bedrijven gebruik maken van commerciële chatbottechnologie, want hiermee maken ze zich economisch afhankelijk, hebben ze geen garantie dat de veiligheid en privacy van hun klanten gewaarborgd worden en kunnen ze geen transparantie geven aan hun productbeheerders en klanten.

Open-source chatbot technologie

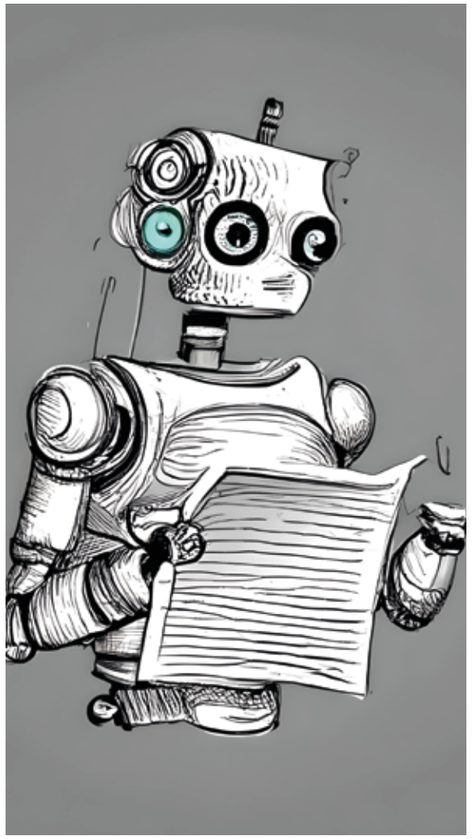
Daarom is het doel van het LESSEN project (volledig uit: *Low-Resource Chat-based Conversational Intelligence*) om open-source chatbot technologie te ontwikkelen voor het Nederlands. Open-source modellen kunnen aangepast worden voor specifieke toepassingen in het MKB. Met deze ontwikkeling wil LESSEN bijdragen aan de democratisering van kunstmatige intelligentie voor het Nederlands. Het project (gefinancierd door NWO via de nationale wetenschapsagenda) is een gezamenlijke inspanning van de Nederlandse academische wereld, bedrijven en overheidsinstanties. LESSEN staat voor een aantal grote uitdagingen bij het trainen en inzetten van grote taalmodellen voor service-chatbots.

Efficiënte modellen

Voor het trainen en optimaliseren van grote taalmodellen is veel computerkracht nodig: GPU-kaarten met een enorm intern geheugen. Niet iedereen beschikt hierover. Voor democratisering van kunstmatige intelligentie is het daarom allereerst van belang dat we neurale netwerken ontwikkelen die niet te veel rekenkracht nodig hebben. Promovendi binnen LESSEN onderzoeken methoden voor een efficiënter gebruik van grote open-source taalmodellen zoals BLOOM: hoe kunnen we die zo efficiënt mogelijk aanpassen aan nieuwe onderwerpen en talen? Stel dat we een taalmodel hebben dat getraind is voor het Nederlands, kunnen we het dan met weinig data en weinig computerkracht aanpassen naar het Fries? Kunnen we hetzelfde model met bestaande data (in kennisbanken bijvoorbeeld) aanpassen voor het genereren van informatie over een meer specifiek onderwerp, zoals verzekeringen? En kunnen we de modellen verder verbeteren met synthetische data, als er niet voldoende echte data beschikbaar is?

Veiligheid en transparantie

Behalve de focus op de computermodellen, heeft LESSEN ook aandacht voor de toekomstige gebruikers van de chatbots. We willen voorkomen dat een chatbot onjuiste informatie of zelfs schadelijke informatie geeft, privacygevoelige informatie van andere gebruikers deelt, of de gebruiker vraagt om persoonlijke informatie zoals inlogcodes te delen. ChatGPT is notoir vanwege het hallucineren, het genereren van onjuiste informatie. Dit is een groot en veel voorkomend probleem, vooral in specifieke domeinen. Chatbots die informatie moeten geven mogen dit niet doen. Het is daarom van belang dat de gegenereerde dialogen alleen informatie bevatten die uit betrouwbare bronnen komt; vaak zal dit de kennisbank van het bedrijf zelf zijn. Denk aan alle informatie op de website van KPN of Achmea:



idealiter kan een klant met een chatbot communiceren over deze informatie, de antwoorden krijgen op vervolgvragen, en zal de chatbot de klant vloeiend en vriendelijk te woord staan, en alleen informatie genereren die uit de bronnen op de website komt.

We willen chatbots graag transparanter maken. Voor de klant betekent transparantie dat het duidelijk is waar de gegenereerde informatie vandaan komt. Voor de productbeheerder of programmeur die de chatbot onderhoudt betekent transparantie dat duidelijk is op welke data de chatbot getraind is, en hoe aanpassingen en verbeteringen gedaan kunnen worden.

Verbinding met bedrijven

De betrokkenheid van bedrijven in LESSEN zorgt ervoor dat de technieken die ontwikkeld worden geworteld zijn in scenario's uit de echte wereld. We werken samen met grote Nederlandse bedrijven, maar hopen dat de invloed van het project verder zal reiken dan de direct betrokken partijen. Uiteindelijk zal hopelijk het MKB profiteren van het democratiseren van open-source chatbot technologie – en kunstmatige intelligentie in het algemeen – voor talen en domeinen met onvoldoende middelen om zelf hun eigen technologie te ontwikkelen. •

TRANSPARANT, BETROUWBAAR EN ETHISCH GEBRUIK VAN AI

VOOR DE VOLGENDE GENERATIE CONVERSATIONAL AI

Door:
Saskia Lensink en Joachim de Greeff, TNO

Hoewel niemand in de toekomst kan kijken, kunnen we er behoorlijk zeker van zijn dat LLMs (Large Language Models) een grote rol gaan spelen in de volgende generatie van conversational AI. De verwachting is dat LLMs beter in staat zijn om natuurlijke interacties met de gebruiker aan te gaan, en het makkelijker maken om op basis van weinig materiaal een spraakassistent of chatbot aan te passen aan verschillende eindgebruikers, situaties en doelstellingen. Daar liggen prachtige kansen, maar ook de nodige uitdagingen als we LLMs op een verantwoorde manier willen gaan inzetten binnen spraakassistenten en chatbots. Hoe doen we dat alles binnen de kaders van de aankomende Europese AI Act?

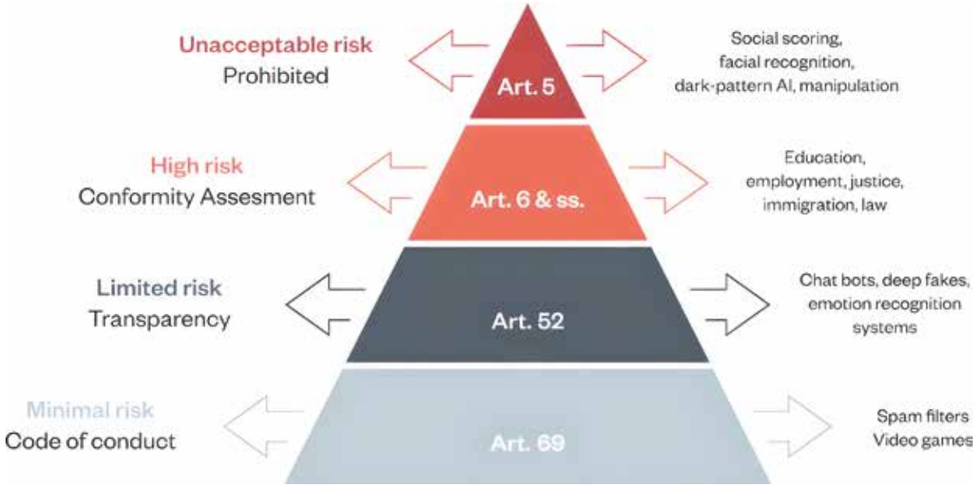
Europese AI Act

De Europese AI Act is 's werelds eerste concrete set van regels rondom AI. Een eerste draft versie is hiervan inmiddels aangenomen[1] en de verwachting is dat we vanaf 2026 aan de eisen moeten voldoen. De nieuwe voorschriften zijn van toepassing op alle partijen die betrokken zijn bij de ontwikkeling, distributie of implementatie van AI: de makers van AI-systemen, de integrators die AI in hun eigen softwaretoepassingen incorporeren, degenen die AI van buiten de EU importeren en de gebruikers die deze diensten inzetten. Deze regelgeving belooft een krachtig kader te bieden voor het gebruik van kunstmatige intelligentie in Europa, zodat AI-systemen veilig zijn en onze fundamentele rechten en waarden gerespecteerd worden. Het legt de nadruk op het waarborgen van ethische en verantwoorde praktijken in de wereld van kunstmatige intelligentie, terwijl het tegelijkertijd innovatie en groei stimuleert.

Het is inmiddels duidelijk dat deze wetgeving geen wassen neus zal zijn. Net als bij de AVG kan het niet naleven van de AI Act leiden tot stevige boetes, die kunnen oplopen tot maximaal 40 miljoen euro.

De AI Act gaat uit van een risico-gebaseerde aanpak met vier gradaties:

1. onacceptabel risico,
2. hoog risico,
3. beperkt risico,
4. laag risico.



Vier verschillende risicoclassificaties in de Europese AI Act (links), met voorbeelden van toepassingen hiervan (rechts)

[1] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52021PC0206

Ieder van deze vier gradaties komt met een eigen set van maatregelen die genomen moeten worden om het risico te beperken (zie Figuur 2). Chatbots worden expliciet genoemd in de categorie 'beperkt risico'. Echter, toen deze risicoclassificatie werd gemaakt, was generatieve AI nog niet in de stroomversnelling waarin het zich nu bevindt, en inmiddels weten we dat er aanvullende eisen zullen gelden voor systemen die gebruikmaken van generatieve AI – wat concreet betekent dat de meeste volgende generatie conversational AI-systemen hiermee te maken krijgen.

Generatieve AI

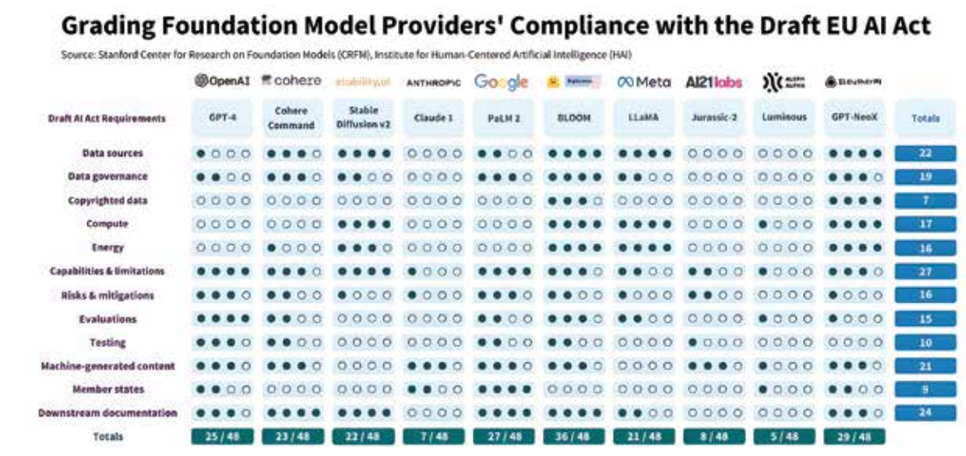
Bij het inzetten van generatieve AI zullen er transparantievereisten gelden, zodat mensen weten wanneer iets gegenereerd is door een AI. Ook moet er een overzicht zijn van welke auteursrechtelijke beschermde gegevens gebruikt zijn voor training. Regels voor foundationmodellen - enorm grote modellen zoals GPT4 die als basis dienen om allerhande toepassingen op te bouwen - zijn nog veel breder. Er moet worden aangetoond dat er voorzorgsmaatregelen zijn genomen om risico's voor gezondheid, veiligheid, rechten, milieu, democratie en de rechtsstaat te verminderen. Er moet ook worden gedocumenteerd welke risico's niet te vermijden zijn. De principes waaraan moet worden voldaan zijn nog steeds behoorlijk hoog-over geformuleerd. Ze geven weliswaar richting, maar de exacte invulling is nog niet duidelijk, en zal in de praktijk verder verkend moeten worden.

Het moge duidelijk zijn, dat het alles behalve eenvoudig of eenduidig is, hoe deze regels praktisch toe te passen zijn voor ontwikkelaars. Hierover zijn nog een heleboel onduidelijkheden, zowel over het meetbaar maken van de mate waarin AI toepassingen wel of niet voldoen aan deze principes, als de wijze waarop organisaties kunnen aantonen dat ze hieraan voldoen. Ook is het nog niet duidelijk welke mogelijkheden overheden hebben om toezicht te houden en uiteindelijk sancties op te leggen als iemand zich niet aan de AI act houdt.

Een onderzoeksgroep van de universiteit van Stanford[2] heeft recent onderzoek gedaan naar in hoeverre verschillende LLMs voldoen aan de eisen gesteld door de AI Act. Hun conclusie: het is niet al te best gesteld. In Figuur 3 zie je welke modellen ze onder de loep hebben genomen en op welke eisen deze zijn beoordeeld. Veel organisaties scoren slecht op vier cruciale gebieden met betrekking tot foundationmodellen: auteursrechten, energie-efficiëntie, risicobeheer en evaluatie. Dit komt door een gebrek aan duidelijkheid over auteursrechten bij trainingsgegevens, ongelijke rapportage van energieverbruik, ontoereikende openbaarmaking van risicobeheer en een gebrek aan evaluatiestandaarden voor foundationmodellen. BLOOM scoort veruit het beste, maar is helaas (nog) niet op het Nederlands getraind en kan dus niet effectief worden ingezet in een Nederlandstalige context.

Evaluatieraamwerk

Er zullen op zowel ethisch, juridisch en technisch vlak stappen gemaakt moeten worden om met elkaar de naleving van de AI Act te waarborgen, ook voor de volgende generatie conversational AI. Bedrijven zouden een verantwoorde AI-bestuursstructuur moeten opzetten die een multidisciplinair team omvat. Dit team, idealiter samengesteld uit data scientists, ethici en juridische experts, zou de complexiteiten van de



Een overzicht van de LLMs die door Stanford onderzocht zijn

[2] https://crfm.stanford.edu/2023/06/15/eu-ai-act.html

technologie, de ethische implicaties ervan en het bijbehorende juridische landschap goed moeten begrijpen.

Een eerste stap is het ontwikkelen van een evaluatieraamwerk dat het mogelijk maakt om AI toepassingen te evalueren op de mate waarin ze voldoen aan de vereisten die de AI Act stelt aan generatieve AI. Een aantal van deze vereisten kan worden getest door middel van technische benchmarks, bijvoorbeeld de accuraatheid, de betrouwbaarheid en de reproduceerbaarheid van AI modellen. Hiervoor bestaan online tests (bijvoorbeeld Google's BIG-Bench[3]) waarmee AI toepassingen, zoals ook de grote taalmodellen, getest kunnen worden en een score krijgen. Andere vereisten, zoals de menselijke controle en toezicht, toegang tot data of verantwoording zullen waarschijnlijk alleen via een audit achteraf kunnen worden aangetoond. Een ander recent onderzoek van onderzoekers van de Radboud Universiteit laat helaas zien dat het nu ook niet al te best gesteld is met de transparantie en openheid van de beschikbare modellen.[4]

Als we in de toekomst in staat willen zijn om de vruchten te kunnen plukken van de mogelijkheden die generatieve AI biedt voor conversational AI, dan zullen we de touwtjes in handen moeten hebben over de gegevens die worden gebruikt om deze AI-modellen te trainen, zodat onze privacy beter wordt gewaarborgd en we in lijn blijven met de EU GDPR en AI-wetten. We willen niet afhankelijk zijn van slechts één grote tech-partij, maar eerder streven naar een diversiteit aan mogelijkheden. Met de juiste voorbereiding kunnen we toewerken naar een verantwoord gebruik van deze veelbelovende technologie, terwijl we ons tegelijkertijd bewust blijven van de risico's en deze op een meer doordachte manier zullen beheren. •

[1] https://github.com/google/BIG-bench
[4] https://opening-up-chatgpt.github.io/

CONVERSATIONAL AI EN DE BEHOEFTE AAN NEDERLANDSE TAAL- EN SPRAAKMODELLEN BIJ DE PUBLIEKE OMROEP

Het gebruik van conversational AI neemt overal een vlucht. Terwijl er veel voordelen zijn voor customer services, zijn er ook andere toepassingen mogelijk. De publieke omroepen experimenteren al een aantal jaren met een optimale inzet van AI, gericht op taal en spraak. En dat heeft zijn eigen uitdagingen.

Door:
Egon Verharen, manager Innovatie, NPO

Van experiment naar implementatie

Het blijft niet alleen bij experimenteren. Zo heeft de publieke omroep al 10 jaar een AI-systeem in productie (24/7/365 met nauwelijks downtime) voor gesproken ondertiteling. Dit is een dienst voor visueel beperkten, die dit systeem als extra audiokanaal kunnen oproepen. Hierbij wordt de vertaalde ondertiteling bij niet-Nederlandse videoprogramma's of fragmenten voorgelezen. Aan de achterkant wordt gebruikgemaakt van beeldherkenning van de teksten en ondertiteling die in beeld verschijnen. Dit wordt omgezet naar geautomatiseerde spraak en gemixt met het achtergrondgeluid en oorspronkelijk stemgeluid. Ook is sinds 3 jaar de near-realtime automatische ondertiteling op het digitale themakanaal 'NPO Politiek' operationeel. Dus alle ondertitels (TT888) worden momenteel 24/7/365 automatisch gegenereerd uit het audiosignaal en synchroon met de spraak in beeld gebracht.

Nog niet overal is een oplossing voor

Een van de uitdagingen is hoe we het achterliggende spraakherkenningssysteem actueel kunnen houden en ook nieuwe termen en voornamen kunnen leren, zonder het hele model te moeten hertrainen. En bepalen op basis van welke feedback we dat doen, waarbij we zoveel mogelijk willen automatiseren. Het systeem werkt goed voor programma's die lijken op de situatie waarvoor de technologie oorspronkelijk is ontwikkeld, namelijk de transcriptie en omzetting naar ondertiteling van debatten in de Tweede Kamer. Daar praten mensen over het algemeen 'netjes' ABN en wordt er niet veel door elkaar gesproken. Dat is bij een discussieprogramma of dramaserie wel anders. De techniek is helaas nog niet zo ver dat we het voor ondertiteling van al onze programma's kunnen inzetten. Overigens is hierbij de technische mogelijkheid niet de enige drijver. De techniek moet ingepast kunnen worden in de workflows van mensen en systemen. Daarbij moeten de investeringen, ook die voor het mogelijk aanpassen van de workflows, economisch te verantwoorden zijn.

Praat met Willem van Oranje

Gelukkig bevinden we ons in een omgeving waarin we nieuwe mogelijkheden kunnen on-

derzoeken en ermee kunnen experimenteren. Zo leren we wat wel en niet werkt voor de beoogde doelgroep, want het publiek staat centraal. De op klantenservice gerichte chatbots voor het ontsluiten van informatie over beschikbare programma's, waar bijvoorbeeld eerder mee is gewerkt voor Jinek (KRO-NCRV) of de portal van de EO, klinken als een logische toepassing. Ze werden echter te weinig gebruikt en gewaardeerd om ermee door te gaan. Op het moment dat je er echter ook andere functies aan kunt verbinden, is het wel degelijk een goed instrument. Een mooi voorbeeld is de educatieve functie die de chatbot 'Praat met Willem van Oranje' (NTR SchoolTV) heeft. Deze is begonnen op smart-speaker-devices, die vanwege privacy redenen uit klaslokalen verbannen werden. Vervolgens is de bot gemigreerd naar de PC en het Smartboard in de klas. Dit was een succes en is nu uitgebreid tot een serie chatbots om te praten met historische figuren uit de geschiedenis canon (https://schooltv.nl/programma/canon-talks/).

De uitdagingen zijn divers. Niet alleen ondersteuning van de juiste devices, maar bijvoorbeeld ook de herkenning van spraak van kinderen. Denk aan kinderen die vragen schreeuwen in de klas in de gesproken versie. Of vragen met veel typefouten of ongebruikelijke woorden in de tekstuele versie. Bovendien roept het ook redactionele vragen op: welke informatie maak je beschikbaar, hoe heeft zo'n historisch figuur geklonken en welke stem past daar nu bij? Een volgende stap is om de chatbot nog meer tot leven te wekken en te koppelen aan een bewegende avatar, om zo op een andere manier verhalen te kunnen vertellen. Publieke omroepen willen impact maken. Kijken en luisteren is van oudsher de manier om informatie te presenteren, maar door interactie te bieden met (het onderwerp van) programma's, kun je het publiek een andere ervaring aanbieden, die wellicht meer impact heeft.

De Slimste Mens

Een ander voorbeeld is de toepassing binnen De Slimste Mens (KRO-NCRV). Hierbij kun je vragen door middel van spraak beantwoorden. Dit is ook een voorbeeld van hoe snel de technologie zich ontwikkelt. Vier jaar geleden heeft het team een eerste voice clone van Philip Freriks gemaakt, toen nog met Lyrebird (nu onderdeel van Descript) op basis van audio-archief. Hier waren honderden uren audio voor nodig en het model was nauwelijks te hertrainen of finetunen. Bovendien was de interface om meer nuances in de gekloonde stem van Philip aan te brengen zo rudimentair, dat het hele project bijna stopgezet is. Nu kun je voor een paar tientjes meerdere voice services in de cloud draaien, die zelfs op basis van een paar minuten spraak al een bruikbaar model opleveren. Toch hebben de meeste van deze diensten nog steeds een beperking, namelijk de ondersteuning van het Nederlands.

Betere spraakmodellen noodzakelijk

Voor de grote techbedrijven lijkt het Nederlandse taalgebied te klein om kwalitatief hoogwaardige modellen voor te ontwikkelen. Voor andere, kleinere spelers is er het probleem dat er eigenlijk geen goede en voldoende diverse trainingsdata beschikbaar is. Er zijn goede ontwikkelingen op het gebied van het ontwikkelen van specifieke modellen. Dat is meestal voor specifieke en beperkte use cases. Maar wij als publieke omroep hebben een bredere behoefte, bijvoorbeeld rond het ondertitelen van onze programma's. Alle soorten spraak, accenten en dialecten komen bij ons voorbij en daar moeten wij dus mee kunnen werken. Dit kan ook over alle soorten onderwerpen gaan. Van een urban jongere op een kanaal van FunX over de uitgaansscene in de stad tot het relaas van een gedupeerde met aardbevingsschade in Groningen. Of zoals uit het eerdere voorbeeld: een tienjarige die een vragenspel speelt met Willem van Oranje.

Toegankelijkheid van nieuwe technologie

Naast het spraakmodel vind ik dat iedereen gebruik moet kunnen maken van de nieuwe mogelijkheden en chatbots. We willen dat alle makers in hun eigen taal, en dat is voor ons nog steeds in grote mate het Nederlands, de op chatbots gebaseerde diensten kunnen gebruiken om nieuwe content te maken. Of het nu gaat om het genereren van beelden of video's, muziek of teksten. Prompt engineering wordt een vereiste vaardigheid in veel sectoren, en zeker ook in de onze. Steeds meer content zal op basis van de juiste tekst-prompts gegenereerd worden. Daarmee kunnen de productieprocessen van audio en video versneld worden en worden producties mogelijk gemaakt die voorheen niet mogelijk waren. Hoe we verantwoord kunnen omgaan met synthetische media in de publieke informatievoorziening zonder het vertrouwen van het publiek te verliezen, is weer een andere uitdaging. Die bewaar ik graag voor een ander artikel.

Samenwerken aan betere spraak- en taalmodellen

Hoewel de behoefte breed is, en gezien over meerdere sectoren ook steeds breder wordt, blijkt het tot nu toe een uitdaging om alle partijen (onderzoek, bedrijfsleven, overheid) op één lijn te krijgen om nieuwe, openbaar (of onder FAIR voorwaarden) beschikbare modellen te ontwikkelen, terwijl dat in andere landen al wel gebeurt. Maar ik blijf mij inzetten, in samenwerking met NOTaS, Stichting Open Spraaktechnologie en de Nederlandstalige Spraakcoalitie, om dat voor elkaar te krijgen. Als onderdeel van ons innovatieprogramma (https://innovatie.npo.nl/projecten/spraakwaterval) bekijken we nu bijvoorbeeld samen met CZ (de verzekeraar, dus uit een hele andere sector) en TNO hoe we een spraakmodel gedistribueerd kunnen trainen, zonder data extern te delen of aan een centrale partij ter beschikking te stellen. Dit is vaak niet mogelijk vanwege privacy redenen of copyright. De eerste resultaten van het project Spraakwaterval zijn bemoedigend. Op deze manier zouden veel meer organisaties hun data op een betrouwbare wijze kunnen inzetten om te helpen bij het trainen van modellen. De uitdaging was vooral om de juristen ervan te overtuigen dat de data echt niet terugkomen in de antwoorden van een chatbot, zoals bijvoorbeeld bij ChatGPT wel gebeurd is. Ik hoop dat we elkaar hierin kunnen vinden en snel een artikel kunnen schrijven over een nieuw project dat het ontwikkelen van nieuwe Nederlandstalige spraak- en taalmodellen tot doel heeft. •

20 JAAR DIXIT

In het afgelopen voorjaar zijn we weer begonnen met onze jaarlijkse DIXIT-redactievergadering. Waar gaan we het over hebben en wie vragen we als hoofdredacteur. Daar begon het mee. Maar… we waren er snel uit en kozen voor het thema 'Nieuwste ontwikkelingen op het gebied van Conversational AI'. Maar, zo merkte iemand op, dat thema hebben we eigenlijk twee jaar geleden toch ook al gehad?

Door:
Arjan van Hessen,
HMI Universiteit Twente &
Telecats

Klopt, maar de ontwikkelingen gaan, zeker na het pas één jaar geleden gelanceerde ChatGPT zo enorm snel, dat we niet anders kunnen. Het gebruik van LLM's zien we op dit moment in allerlei toepassingen. Zo worden ze onder meer gebruikt in de gezondheidszorg, de financiële wereld, entertainment en het onderwijs. Deze grote taalmodellen hebben deels de manier veranderd waarop we menselijke taal begrijpen en ermee werken. Ze helpen verschillende soorten gebruikers en ontwikkelaars om nauwkeurigere en geavanceerdere systemen te maken.

DIXIT's

Was dit terug te zien in de DIXIT's van de afgelopen 20 jaar? Ik denk het wel. De eerste DIXIT's verschenen zelfs een paar keer per jaar, maar na 2010 verschenen ze eens per jaar (en werden ze dikker). Vanaf 2010 werd er per keer een thema gekozen, een gastredacteur gevraagd en over het algemeen erg hard gewerkt om het thema van allerlei kanten te belichten. Sommige DIXIT's waren vooral op 'Taal- en Spraaktechnologie' gericht, anderen meer op de wereld eromheen (Cultureel erfgoed, Expats, TST in de Zorg, Overheid of Onderwijs). Het ging erom om zowel onderwerpen uit het onderzoek, als toepassingen voor ontwikkelaars en geïnteresseerde eindgebruikers te behandelen. Ik denk dat dat goed gelukt is en we gaan er voorlopig gewoon mee door.

Kijken we 20 jaar terug, dan zien we dat het toen niet voor de hand lag om te verwachten dat termen als NLP, ASR, TTS en LLM zo algemeen gebruikt zouden gaan worden. Het eerste nummer noemde echter al het aansprekende onderwerp *'Spraaktechnologie de puberjaren voorbij: van herkenning naar interpretatie'*. Iets dat toen begon te spelen, hoewel de spraakherkenning zeker nog niet de puberjaren voorbij was!

Ik bleek het stuk zelf geschreven te hebben als een verslag van mijn bezoek aan Interspeech 2003 in Genève en ben het eerlijk gezegd nog steeds redelijk eens met die tekst. Maar… dit kwam kennelijk veel en veel te vroeg.

How May I Help You?

Op die conferentie werd (voor mij) voor het eerst een 'How May I Help You' toepassing met automatische spraakherkenning getoond. De software ging ervan uit dat je gebruikers niet langer moet lastig vallen met de eis dat zij altijd precies moeten onthouden wat je precies kunt of moet zeggen tegen de service. In plaats daarvan wordt er gewoon gevraagd 'Waarmee kan ik je helpen'?

Het klinkt nu als doodnormaal maar toen was dat redelijk opzienbarend! Terug in Nederland vertelde ik mijn collega's bij Telecats hierover en in de loop van de maanden/jaren erna hebben we samen het onderwerp (voor het Nederlands) uitgewerkt.

TTS

Ook Text-to-Speech heeft een vergelijkbare ontwikkeling doorgemaakt. Hoewel ook toen al de stemmen niet meer als echte computerspraak klonken, was ze meestal toch goed onderscheidbaar van 'echte' menselijke spraak. Zeker als er emotie in de boodschap lag, hoorde je dat die in de gegenereerde spraak zo goed als afwezig was. Het werd toen mogelijk om zelf 'emotie' in de spraak te leggen door bijvoorbeeld pauzes, toonhoogte en snelheidsveranderingen en andere kenmerken van 'emotie' toe te voegen. Maar omdat dit menselijk handwerk was, maakte het de automatische processen helaas niet eenvoudiger.

De laatste 10 jaar

Wanneer we 10 jaar later kijken (2012) dan is er een DIXIT over 'Big Data en TST'. We waren er ondertussen achter dat het met de oude softwaremethoden niet ging lukken en langzaam stapte men over op ten minste deels een Big data-benadering waarin computers grote hoeveelheden data analyseerden en 'er iets nuttigs' mee deden.

Met de komst van Big Data en algoritmes die daar iets nuttigs mee deden kreeg TST een enorme boost. Spraakherkenning werd veel robuuster, kon meer variatie aan en werkte sneller en beter. TTS werd eveneens veel beter en het verschil tussen menselijke spraak en computerspraak werd in veel gevallen zo klein dat je echt je best moest doen om het te horen.

In de jaren erna werd er hard gewerkt aan allerlei nog slimmere algoritmes die vooral gebruik maakten van DDN's (Deep Neural Networks) in de ruime zin van het woord. Het ging hard en vanaf 2015/ 2016 zagen we de WER (Word Error Rate) in verschillende toepassingen hard achteruit gaan.

Maar hier bleef het niet bij.

Grote Amerikaanse bedrijven richtten zich op de zogeheten Large Language Models en kwamen tussen 2017 en 2023 met de eerste applicaties (Google met Transformer; Amazon met Bedrock, Microsoft met Copilot, Meta met LLAMA 2).

LLM's zijn zeer grote Deep Learning-modellen die vooraf zijn getraind op enorme hoeveelheden 'teksten'. De onderliggende transformator is een set neurale netwerken die bestaan uit een encoder en een decoder met elk 'self-attention' capaciteiten. De encoder en decoder extraheren betekenissen uit een tekstsequentie en 'begrijpen' de relaties tussen woorden en zinnen in de tekst.

Maar de grootste klap kwam een jaar geleden met de publicatie van decoder model ChatGPT. OpenAI, een tot dan toe redelijk onbekende onderneming uit 2016, kwam (na een enorme investering vanuit Microsoft) met zijn GPT 3, snel gevolgd door GPT 3.5 en een paar maanden later door GPT 4. En het maakte de bijbehorende app ChatGPT beschikbaar voor het grote publiek (december 2022). In twee maanden tijd groeide de app naar meer dan 100M gebruikers: de snelste groei ooit gemeten!

En nu?

Het eerste enthousiasme over ChatGPT is iets verminderd: het is en blijft een geweldig hulpmiddel maar men ziet in dat het niet de oplossing is voor al je problemen. Niet dat de maker, OpenAI, dat ooit heeft gezegd, maar de eerste maanden van 2023 kreeg je wel vaak de indruk dat ChatGPT heel veel zou veranderen en niet alleen ten goede! Het wordt nu gebruikt en men leert langzaam aan wat het goed kan en waar menselijke controle noodzakelijk blijft. Het wordt, kortom, een normale applicatie die gebruikt kan en zal worden door ongeveer iedereen die iets met tekst doet en in toenemende mate ook door mensen die iets met plaatjes doen. •

20 JAAR DIXIT...EEN TERUGBLIK

NOTaS

STICHTING NOTAS

Stichting NOTaS
Secretariaat Stichting NOTaS
Toernooiveld 300
6525 EC Nijmegen
T: 024-351 21 08
E: info@notas.nl
W: notas.nl

Stichting NOTaS behartigt de belangen van bedrijven en kennisinstellingen in de sector Taal- en Spraaktechnologie (TST). Dit gebeurt onder meer door het organiseren van bijeenkomsten, lobbyactiviteiten en de uitgave van het tijdschrift DIXIT.

DEELNEMERS STICHTING NOTAS

Amberscript B.V.
Keizersgracht 209
1016 DT Amsterdam
T: 06 13110023
E: peterpaul@amberscript.com
W: amberscript.com

Met Amberscript helpen we overheden, universiteiten en (media) bedrijven om kosten en moeite te besparen van het handmatig transformeren van audio of video naar 100% correcte teksten en ondertitels. We maken automatische spraak naar tekst engines speciaal voor Europese talen. Door engines speciaal voor specifieke klanten te trainen, komen deze in de buurt van menselijke nauwkeurigheid. Onze online tekstverwerker of menselijke transcribenten brengen de tekst naar 100% nauwkeurigheid. Bestel of creëer snel en makkelijk accurate ondertitels en transcripten op https://www.amberscript.com, of gebruik onze spraakherkenning APIs voor de nauwkeurigste Europese spraakherkenning.

Beeld & Geluid
Media Parkboulevard 1
1217 WE Hilversum
T: 035 677 5555
E: info@beeldengeluid.nl
W: beeldengeluid.nl

Beeld & Geluid is een van de grootste audiovisuele archieven van Europa en herbergt ruim 1 miljoen uur aan radio, televisie, film en muziek. Beeld & Geluid initieert ook onderzoek dat audiovisueel erfgoed open en doorzoekbaar maakt. We volgen relevante innovaties in audiovisueel archiveren, nemen deel in onderzoeksprojecten en experimenteren met nieuwe technologie. Daarvoor wordt bijvoorbeeld gebruik gemaakt van spraaktechnologie, zoals spraakherkenning en sprekerherkenning.

Cornelistools
Beeklaan 38
1403 RE Bussum
T: 06 44330071
E: serge@cornelistools.nl
W: cornelistools.nl

Met eigen taal- en dialoogtechnologie helpt Cornelistools bedrijven toepassingen te ontwikkelen op het gebied van Conversationele Interfaces, Data Science en Gamification. Met ons platform richten we ons specifiek op de Nederlandse taal en onderzoeken en ontwikkelen we vernieuwende technologische mogelijkheden op het gebied van natuurlijk taalbegrip, voice en chatbots.

Data Archiving and Networked Services (DANS)
Anna van Saksenlaan 51
2593 HW Den Haag
T: 070 349 44 50
E: info@dans.knaw.nl
W: dans.knaw.nl

DANS is het Nederlands instituut voor permanente toegang tot onderzoeksgegevens. Onze activiteiten zijn gecentreerd rond drie kerndiensten: data-archivering, datahergebruik en training en consultancy inzake FAIR datamanagement. Hiertoe bieden we het gecertificeerde langetermijn-archief EASY aan en DataverseNL voor het opslaan en delen van data tijdens het onderzoek, evenals het NARCIS-portal voor Nederlandse onderzoeksinformatie. DANS is een instituut van KNAW en NWO.

Dedicon
Traverse 175
5361 TD Grave
T: 0486 486 486
E: info@dedicon.nl
W: dedicon.nl

Stichting Dedicon gelooft in een wereld waarin iedereen moet kunnen lezen en leren wat hij wil in een vorm die hem past. Daarom ontwikkelen en produceren wij aantrekkelijke alternatieve leesvormen. Inmiddels maken tienduizenden mensen met veel plezier gebruik van onze diensten. Binnen de diensten van Dedicon speelt de moderne taal- en spraaktechnologie een belangrijke rol.

Den Hamer Consulting
Herfterlaan 26
8026 RA Zwolle
T: 06 41014172
E: info@hamerconsulting.nl
W: denhamerconsulting.nl

Den Hamer Consulting is gespecialiseerd in het oplossen van complexe vraagstukken met behulp van data en modellen. Samen met de klant worden de juiste vragen geïdentificeerd die vervolgens met data en modellen worden beantwoord.

Fryske Akademy
Doelestraat 8
8911 DX Leeuwarden
T: 058 2131414
E: jdijkstra@fryske-akademy.nl
W: fryske-akademy.nl

De Fryske Akademy is het wetenschappelijk onderzoeksinstituut op het gebied van de Friese taal, cultuur en regionale geschiedenis, en bestudeert deze onderwerpen vanuit internationaal en multidisciplinair perspectief. De Fryske Akademy werkt samen met diverse partners aan verschillende spraak- en taaltechnologische toepassingen voor de Friese casus.

Instituut voor de Nederlandse Taal
Rapenburg 61
2311 GJ Leiden
T: 071 514 16 48
E: secretariaat@ivdnt.org
W: ivdnt.org

Het Instituut voor de Nederlandse Taal is hét kennisinstituut voor het Nederlands. Het is een breed toegankelijk wetenschappelijk instituut dat alle aspecten van de Nederlandse taal bestudeert, waaronder de woordenschat, grammatica en taalvariatie. Het instituut verzamelt de nieuwste Nederlandse woorden, actualiseert belangrijke naslagwerken zoals de Algemene Nederlandse Spraakkunst en maakt vaktaal toegankelijk via terminologielijsten. Daarnaast fungeert het instituut als kennis- en distributiecentrum voor digitale Nederlandstalige tekstverzamelingen, woordenlijsten, wetenschappelijke woordenboeken, spraakcorpora en taal- en spraaktechnologische software.

Radboud Universiteit - CLST
Erasmusplein 1
6525 HT Nijmegen
T: 024 361 16 86
E: clst@let.ru.nl
W: ru.nl/clst

Het Centre for Language and Speech Technology (CLST) onderzoekt en adviseert op het gebied van taal- en spraaktechnologie (TST). Belangrijke speerpunten in het onderzoek zijn audio- en tekstontsluiting en de inzet van TST ten behoeve van het onderwijs en de zorg. Wij zijn gespecialiseerd in automatische spraakherkenning en tekstanalyse. Tevens rekenen wij ontwikkeling en curatie van data en tools tot onze expertise.

ReadSpeaker
Princenhof Park 13
3972 NG Driebergen-Rijsenburg
T: 030 692 4490
E: nederland@readspeaker.com
W: readspeaker.com

ReadSpeaker is een wereldwijde speler op het gebied van tekst-naar-spraaktechnologie (TTS). Als enige provider levert ReadSpeaker het totale palet van stemmen, software en (maatwerk)oplossingen. ReadSpeaker is gevestigd in 15 landen en heeft wereldwijd meer dan 10.000 klanten in 65 landen en levert zowel SaaS- als licentieproducten. Voor verschillende kanalen, platformen en apparaten biedt ReadSpeaker bedrijven en organisaties een stem voor on- en offline content, embedded-, server- en desktoptoepassingen, spraakproductie en meer. Met meer dan 20 jaar ervaring is het team van ReadSpeaker leidend in tekst-naar-spraak.

Technische Universiteit Delft - Afdeling Intelligent Systems
Van Mourik Broekmanweg 6
2628 XE Delft
T: 015 27 89803
E: o.e.scharenborg@tudelft.nl
W: tudelft.nl/ewi/over-de-faculteit/afdelingen/intelligent-systems

De afdeling INSY heeft als doel mens en machine in staat te stellen om te gaan met de toenemende hoeveelheid en complexiteit van data, in nauwe samenwerking met hun omgeving. De afdeling integreert fundamenteel onderzoek, engineering en ontwerp in de in elkaar grijpende gebieden van gegevensverwerking, interpretatie, visualisatie en interactie met behulp van model- en kennisgebaseerde methoden en algoritmen. Het onderzoek is geïnspireerd op uitdagingen op het gebied van: consumentenelektronica en entertainment, cultureel erfgoed, sociale media, medische en gezondheidswetenschappen, beveiliging en privacy, veiligheid en incidentbeheer.

Telecats
Colosseum 42
7521 PT Enschede
T: 053 488 99 00
E: info@telecats.nl
W: telecats.nl

Al meer dan 25 jaar staat Telecats voor de ontwikkeling van innovatieve oplossingen in klantcontact. Met onze spraaktechnologie en kunstmatige intelligentie routeren we klantcontact met spraakherkenning naar de best mogelijke oplossing; een medewerker of selfservice. We optimaliseren de klantreis en brengen inzicht in het klantcontact.

Tilburg University – Cognitive Science & Artificial Intelligence
Warandelaan 2
5037 AB Tilburg
T: 013 466 81 18
E: A.H.Verschoor@tilburguniversity.edu
W: csai.nl

In onderwijs en onderzoek van het Departement Cognitive Science & Artificial Intelligence (CS&AI) ligt het accent op aspecten van interactieve technologieën, zoals robotics en avatars, serious gaming, virtual & mixed reality, en taal- en data technologieën. Lopende onderzoeksprojecten betreffen computationele modellen van taal, automatische analyses van complexe data in de zorg en het maken en testen van virtuele omgevingen voor het onderwijs. Het departement heeft nauwe contacten met het Jheronimus Academy of Data Science (jads.nl) en Mind Labs (mind-labs.eu). Het departement verzorgt onder andere de bloeiende bachelor- en masteropleidingen Cognitive Science & Artificial Intelligence en is verantwoordelijk voor de interfacultaire Master Data Science & Society.

Universiteit Leiden - LIACS
Niels Bohrweg 1
2333 CA Leiden
T: 071 527 57 72
E: secr@liacs.leidenuniv.nl
W: liacs.leidenuniv.nl

Het Leiden Institute of Advanced Computer Science (LIACS) is het instituut voor onderzoek en onderwijs op het gebied van informatica en kunstmatige intelligentie aan de Universiteit Leiden. Het LIACS werkt actief samen met overheden en bedrijven. Achter de oplossingen voor maatschappelijk relevante problemen zitten theoretische ontdekkingen, met statistiek als belangrijke basis. Taaltechnologie heeft een groter wordend aandeel in het onderzoek en onderwijs van het LIACS.

Universiteit Twente - HMI
Drienerlolaan 5
7522 NB Enschede
T: 053 489 37 40
E: HMI-SECR-L@lists.utwente.nl
W: utwente.nl/en/eemcs/hmi/

De Human Media Interaction (HMI) groep van de Universiteit Twente (UT) is een onderzoeksgroep binnen de afdeling Informatica. HMI doet onderzoek op het gebied van mens-machine interactie naar het ontwerpen, ontwikkelen, en evalueren van (sociaal) interactieve systemen (o.a. gesproken dialoogsystemen voor virtuele agents en sociale robots, interactive storytelling, multimedia retrieval) die de gebruiker ondersteunen op een plezierige en efficiënte manier. Het automatisch analyseren en interpreteren van menselijk verbaal en non-verbaal gedrag op basis van taal en spraak, fysiologische maten en lichaamstaal speelt daarbij een belangrijk rol. De generatie van verbaal en non-verbaal gedrag van robots en virtual agents behoort tevens tot het onderzoek. Daarnaast verzorgt de HMI groep veel onderwijs voor de masteropleiding Interaction Technology en de bacheloropleiding Creative Technology.

Utrecht University - Institute for Language Sciences
Trans 10
3512 JK Utrecht
T: 030-253 60 06
E: ils@uu.nl
W: hum.uu.nl/ils

Het Utrecht University, Institute for Language Sciences is een onderzoeksinstituut binnen de Faculteit Geesteswetenschappen van de Universiteit Utrecht. Het beoogt onderzoek te doen naar menselijke taal vanuit verschillende perspectieven, waaronder de architectuur van taal, de ontwikkeling van taal en het gebruik van taal als communicatiemiddel. Taal en technologie vormt een andere belangrijke onderzoekspijler. Daarbij is aandacht voor computationele linguïstiek, spraak- en taaltechnologie, culturele en menselijke aspecten van AI en digitale geletterdheid.

Y.Digital B.V.
Arhemse Bovenweg 140
3708 Zeist
T: 030 2074274
E: result@y.digital
W: Y.digital

Y.digital is een gespecialiseerd AI-bedrijf dat zich richt op intelligente taalverwerking. Via ons geavanceerde AI-platform Ally leveren we oplossingen voor contact centres, zoals chatbots en spraakassistenten. Maar ook Intelligent Document Processing-oplossingen die organisaties helpen bij het meer consistent, schaalbaar en efficiënt maken van kennisintensieve processen. Denk aan de verwerking van email, nieuwe aanvragen of wijzigingsverzoeken. Dit alles vanuit de ambitie 'Empowering Humans'.

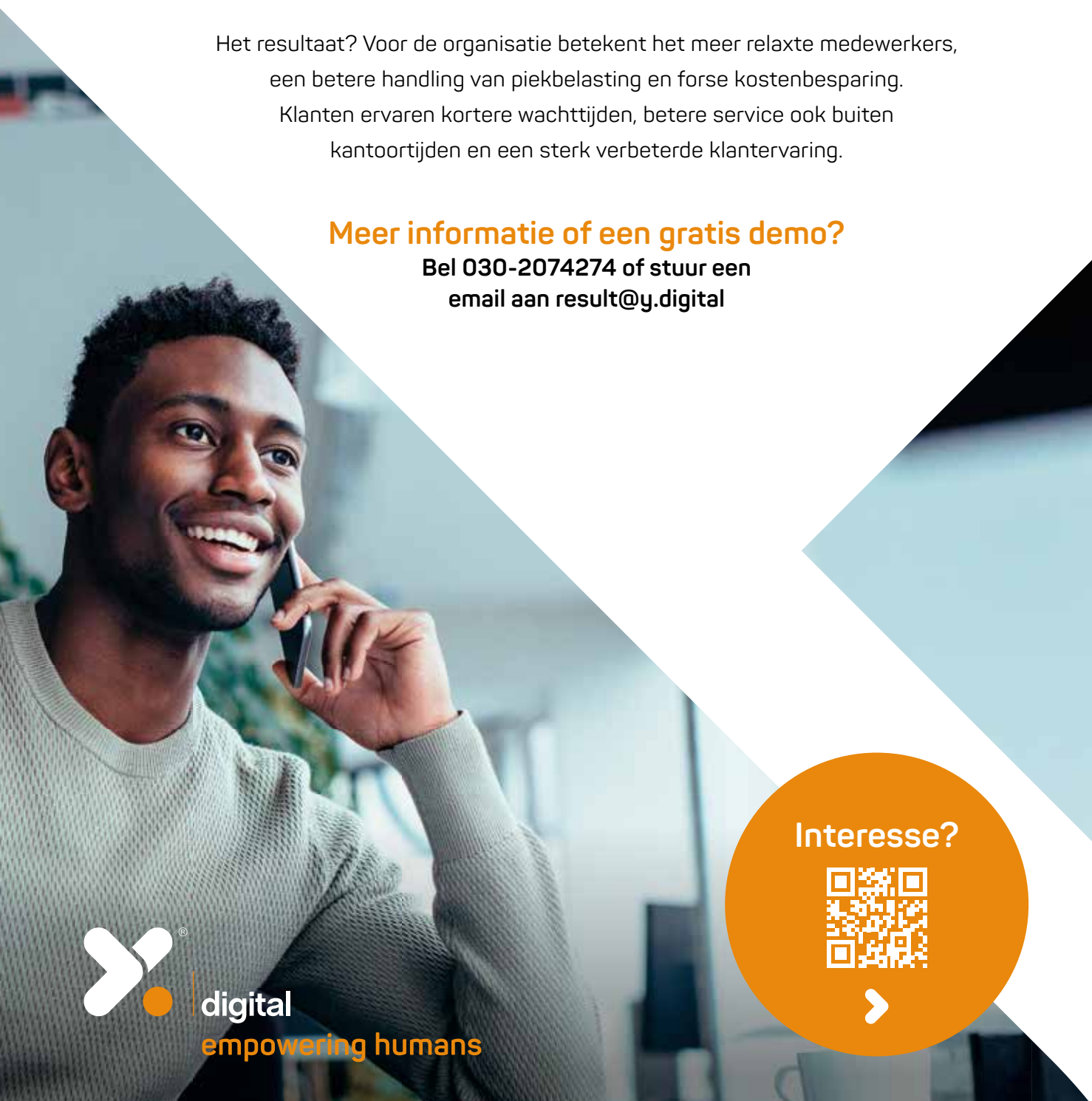
Hoe future-proof is uw contact center?

De volgende generatie contactcentra maakt gebruik van natural language processing en - understanding en knowledge graphs. Om op een zeer intelligente en geautomatiseerde manier met klanten te communiceren, terwijl medewerkers support krijgen bij het efficiënt afhandelen van klantcontact en saai en repetitief werk wordt geautomatiseerd.

Het resultaat? Voor de organisatie betekent het meer relaxte medewerkers, een betere handling van piekbelasting en forse kostenbesparing. Klanten ervaren kortere wachttijden, betere service ook buiten kantoortijden en een sterk verbeterde klantervaring.

Meer informatie of een gratis demo?

Bel 030-2074274 of stuur een email aan result@y.digital



Interesse?

digital
empowering humans