

DIVIT

A woman with dark hair, wearing a black top, is sitting at a desk and smiling while looking at a laptop. A blue thermal cup is in the foreground. The background shows an office environment with plants and shelves.

TIJDSCHRIFT OVER TAAL- EN SPRAAKTECHNOLOGIE

Conversational AI

INHOUD

INHOUD	Pagina
Voorwoord en colofon	2
Conversational AI: beware of the bots?	3
Ethiek en Diversiteit	
Ik begrijp niet dat je me niet begrijpt	6
Gendervrije inclusieve technologie	8
Hallo, ik ben Anna, uw virtuele assistent	10
Inclusieve automatische spraakherkenning	13
Inclusive conversational experiences	15
Nederlands als low-resource language	
De Nederlandstalige Spraak Coalitie	24
Conversational TTS	26
Een voice of chatbot voor Vlaanderen en Nederland	28
Best practises	
Persoonlijkheid in conversational inter'faces'	17
New kid on the block: intent management	20
Van intent-classificatie naar wederzijds gesprek	30
Wat je als taalkundige kunt bijdragen aan een conversational interface	32
De uitdaging van conversational interfaces	34
Gepersonaliseerde service met voice bot Sam van VGZ	36
Explainable AI	
Op zoek naar interpretatie: InDeep	22
Meer informatie	
Leeslijst	25
Podcastlijst	33
Opleidingen en cursussen	38

VOORWOORD

We zijn weer een jaar verder. Er is in dat jaar genoeg veranderd, maar veel is ook hetzelfde gebleven: de opeenvolgende coronabesmettingsgolven, de persconferenties, mondklappers op, mondklappers af, en voor velen van ons nog steeds thuiswerken.

Na zo'n lange tijd moet ik bekennen dat ik enorm naar persoonlijke interacties verlang. Van het spreekwoordelijke koffiezetapparaatgesprek mis ik niet zozeer de koffie als wel het gesprek, soms over het werk maar meestal over koetjes en kalfjes. Daarom, van tijd tot tijd, in de eenzaamheid van mijn kleine thuishok, open ik mijn laptop, start ik Emacs en typ ik vlug op mijn toetsenbord "M-x doctor"...

En daar is Joseph Weizenbaum's Eliza, een van de eerste implementaties van Conversational AI met het bekende psychotherapeutenscript. Met zijn simpele en elegante logica (geschreven bij de MIT Artificial Intelligence Laboratory midden jaren 60 en verpakt in nog geen 1600 regels) kan ik een surrogaatconversatie houden over koetjes en kalfjes met een altijd beschikbare virtuele gesprekspartner.

Wie echter geen open conversatie maar antwoorden zoekt, wordt helaas snel teleurgesteld.

Voor serieuze antwoorden moeten we tien jaren in de tijd reizen - midden in de jaren 70 - maar wel dichtbij huis, namelijk in Eindhoven. Daar, bij de Philips Research Laboratories, ontwikkelde een groep Nederlandse TST-pioniers PHLIQA1, een vraag-antwoord systeem in natuurlijke taal. Maar een echte dialoog is het natuurlijk niet.

Beide voorbeelden laten volgens mij de uitersten van het spanningsveld van Conversational AI zien. Een goed gesprek kan niet alleen met droge feiten gevoerd worden, maar ook niet uitsluitend uit platitudes bestaan.

Na vele verbeterlagen zijn de uitdagingen van Conversational AI grotendeels onveranderd, maar het is in ieder geval duidelijk dat Conversational AI in Nederland springlevend is en eer doet aan onze pioniers. Dat en meer kunt u lezen in deze editie van DIXIT.

Veel leesplezier,

Fabrice Nauze,
voorzitter Stichting NOTaS



DIXIT: Tijdschrift over toegepaste taal- en spraaktechnologie – 18e jaargang, editie Conversational AI. 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: henk.vandenheuvel@ru.nl; Martha Hofman: MAHofman@hetnet.nl; Marieke den Os (secretariaat NOTaS): info@notas.nl. **Hoofredactie themanummer:** Maaike Groenewege: maaike@convocat.nl. **Advertenties:** Stichting NOTaS – info@notas.nl, 024-3512108 **Abonnementen:** Voor een gratis abonnement kunt u zich wenden tot een van de NOTaS-deelnemers **Druk:** Leonard B2B Communicatie **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.

CONVERSATIONAL AI: BEWARE OF THE BOTS?

Voor u ligt een prachtig themanummer over conversational AI, oftewel de technologie van het 'praten met dingen'. Chatbots, smartspeakers zoals Amazon Alexa, social robots, de metaverse én je woonkamerverlichting: conversational interfaces vind je tegenwoordig overal.

Aan mij als gastredacteur de eer om een en ander te voorzien van een inleidend commentaar. Volgens Adri Vermaat, ombudsman van dagblad Trouw, laat het redactioneel commentaar "de ziel van de krant zien". Meer dan objectieve rapportering en informering, geeft het commentaar duiding aan recente ontwikkelingen en hoe ze te plaatsen in onze bekende en vertrouwde eigen context.

En dat dan voor conversational AI. Ga er maar aan staan. Ik doe een poging.

De staat van conversational AI

Ik denk dat we mogen stellen dat conversational AI de hype voorbij is. Chatbots zijn inmiddels gemeengoed voor veel organisaties. Bedrijven wagen steeds vaker een uitstapje naar smart speakers en 'digital humans' en spraaktechnologie op het telefoonkanaal is terug van nooit weg geweest. Ondanks dat Google Nest me alleen verstaat als ik een halve octaaf lager praat dan normaal, en we met z'n allen heel wat afmoppen op die chatbots die ons niet begrijpen, zie ik dat het organisaties niet afschrikt.

Integendeel, er wordt flink geïnvesteerd in alle aspecten van conversational AI: conversation design ontwikkelt zich als volwassen ontwerpdiscipline op het snijvlak van taal en techniek; bedrijven zetten vol in op het ontwikkelen van *data-driven*, *user-centered* bots die echt helpen. Nederlandse onderzoekers staan bekend om hun hoogwaardig onderzoek naar social robotics. We beschikken inmiddels over een masteropleiding Voice Technology aan de Rijksuniversiteit Groningen, en de eerste minor Conversation Design ter wereld is in de maak bij de Hogeschool van Amsterdam.

Sluimerend onbehagen

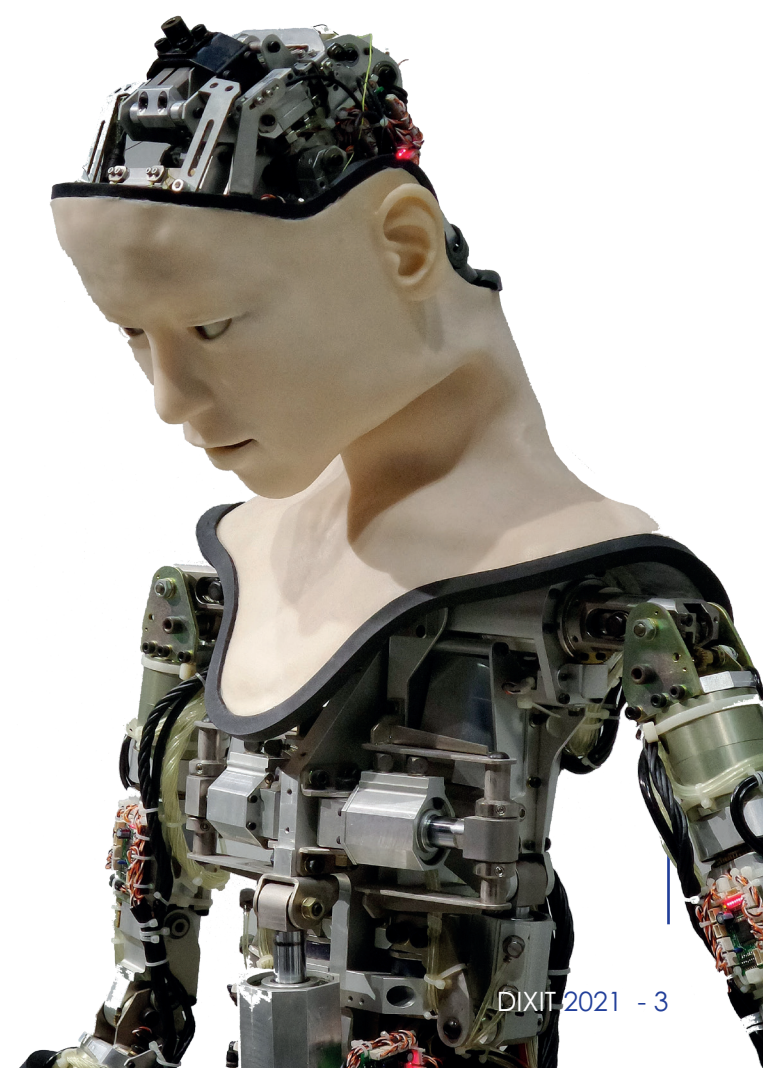
'Geweldig toch?' hoor ik u denken, en dat kan ik vanuit persoonlijk perspectief alleen maar beamen: zeker voor een taalkundige-met-heimwee zoals ik, zijn dit geweldige tijden. Ik zit met m'n neus boven op de nieuwste ontwikkelingen en help ze in veel gevallen zelf vormgeven. Maar met mijn enthousiasme over nieuwe mogelijkhe-

den, bekruipt me ook wel eens een sluimerend gevoel van onbehagen.

Ik vraag me regelmatig af wat het doet met ons als mens, als cultuur, wanneer de technologie, waarvan we inmiddels bewezen afhankelijk zijn geworden, expliciet wordt ontworpen om relaties met ons aan te gaan. En letterlijk steeds dichterbij op onze huid komt te zitten. Ik probeer dit onbehagen te vangen in drie vragen: Versta je me? Kan ik je vertrouwen? Wie zorgt ervoor dat dit ethisch, veilig en algemeen OK blijft?

Versta je me?

Ik hoorde onlangs voor het eerst het woord *Anglophone privilege*, oftewel het gegeven dat in voicetechnologie het Engels vrijwel altijd de norm is. Amazon Alexa spreekt geen Nederlands. Veel innovaties in Voice zijn alleen beschikbaar voor het Engels. Startups in Voice, ook Europese, richten zich vaak primair op een doorbraak in de Verenigde Staten. En conversational datasets, belangrijke input voor het verbeteren van spraakherkenning, zijn bijzonder schaars voor niet-Engelse talen.



Door:
Maaike
Groenewege,
Convocat

Anglophone privilege - wat Amerikaans uitgedrukt wellicht, maar het roept behoorlijk wat herkenning op. In de twee wereldwijde communities waar ik actief deel van uitmaak, Voicelunch en Women in Voice, valt enorm veel te leren en te



- Keep home closer - Meet Astro, the household robot for home monitoring, with Alexa.
- Introducing Intelligent Motion - Amazon Astro uses advanced navigation technology to find its way around your home and go where you need it. When you're not using Astro, it will hang out close by at the ready.
- Stay connected from anywhere - Remotely send Astro to check on specific rooms, people, or things. Plus, get alerts if Astro detects an unrecognized person or certain sounds when you're away.
- Unlock even more peace of mind - Activate your 6-month free trial of Ring Protect Pro subscription and have Astro proactively patrol, investigate activity, save videos in Ring's cloud storage for up to 60 days, and more.
- Alexa Together subscription (coming soon) - Remotely care for aging loved ones, giving you peace of mind while helping them live independently. Set up reminders, manage shopping lists, receive activity alerts, and more.
- Put Alexa in motion - Astro can follow you with entertainment or find you to deliver calls, messages, timers, alarms, or reminders.
- Designed to protect your privacy - Turn off mics, cameras, and motion with one press of a button and use the Astro app to set out of bounds zones to let Astro know where it's not allowed to go.
- Customize with compatible products - Astro comes with a detachable cup holder and can carry other items (sold separately) like a Ziploc container, the OMRON blood pressure monitor, and a Furbo Dog Camera that tosses treats to your pet.

We want you to know

Astro cannot go up or down stairs.

Astro's kenmerken - hondensnoepjes gooien, een kopje koffie rondrijden en vooral: heel veel data verzamelen.

delen. Maar ze bieden niet de plek waar we een lans kunnen breken voor de uitdagingen van ons eigen taalgebied; want nu geldt dat als het geen Engels is, dan is het simpelweg geen onderwerp van gesprek.

Van globaal naar hyperlokaal

En daar waar wereldwijde bewegingen op het web, zoals het W3C-consortium, ons veel goeds hebben gebracht, begin ik voor taal-en spraak-technologie steeds meer te geloven in een hyperlokale aanpak.

Tijd voor verandering dus. Women in Voice Nederland gaat inmiddels verder als Women in Conversational, een hyperlokale community voor vrouwen die werken in conversational AI (en ja, mannen zijn net zo welkom!). Zodat we daar kunnen zijn waar we waarde kunnen toevoegen, aan het Nederlands, het Vlaams, het Fries en het Limburgs. Voor verbinding tussen professionals, onderwijzers, studenten en onderzoekers.

En Women in Conversational is niet de enige: Mozilla Common Voice en sinds kort ook de Nederlandse Spraak Coalitie werken vanuit hetzelfde perspectief. Kortom: ook in Nederland bouwen we aan een steeds steviger community rondom conversational, spraak en voicetechnologie. Meedoen? Je bent van harte welkom!

Kan ik je vertrouwen?

Dit is Amazon Astro. Het is een schattig huishoudrobotje op wielen met grote puppy-ogen. Hij volgt je trouw en met schattige non-verbale geluidjes wekt hij de indruk dat hij contact met je maakt. Via de ingebouwde Alexa kun je ook daadwerkelijke gesprekken met hem voeren. Wat me wel opviel: voor een huishoudrobot houdt Astro verbazingwekkend weinig huis; hij kan niet stofzuigen, de plinten afstoffen of de planten water geven. Toegegeven, hij kan een kopje koffie rondrijden en een hondensnoepje gooien. Maar daar blijft het ook wel bij.

Wat Astro wel kan? Het is een datadiet eerste klasse: met z'n camera, gezichtsherkenning, microfoons, GPS én de afstandsbediening van je voordeur volgt hij je de hele dag in je meest privé omgeving: je eigen huis. En met een big tech gigant die een aanzienlijk belang heeft bij jouw persoonlijke informatie, bekruipt me daar toch weer dat sluimerende onbehagen. Wie garandeert mij dat mijn data niet bij Amazon terechtkomen?

Darwinian buttons

Nu denkt u misschien 'Daar zijn we toch zelf bij? Ik kan toch prima zelf bepalen wat ik wel of niet gebruik?' Nou, da's een interessante. In haar boek 'Alone together' beschrijft Sherry Turkle hoe robots worden ontworpen om onze 'Darwinian buttons' in te drukken: "Making eye contact, trac-



king and mimicking motion, saying our name. These make people experience a robot as not just clever and thus deserving of respect, but as deserving of care and attention - and beyond this, as having care and attention to give us."

En als ons zorginstinct wordt aangesproken, verliezen we blijkbaar vrij makkelijk ons kritisch denkvermogen: "We prove willing to talk to robots about personal matters, to take them as confidants if they could give us that feeling that they understood (regardless of whether they had any comprehension of the loves and lives we were describing)."

Large language models

Deze kennis is natuurlijk niet nieuw. Het is wel nieuw dat met de opkomst van conversational AI, en met name de opkomst van large language models, inmiddels ook de schijn kan worden gewekt van 'comprehension', van een robot die je daadwerkelijk begrijpt.

Large language models zijn AI-algoritmes die worden getraind op ongelooflijk grote datasets. Daardoor zijn ze in staat om -op het eerste gezicht- coherente gesprekken te voeren, puur op basis van die data. Helaas worden deze datasets samengesteld uit internetdata zoals Reddit en Twitter, de onderbuik van het internet. Ze bevatten dan ook grote hoeveelheden misogynie, homofobie en racistische data. Niet echt het soort gespreksstof dat thuishoort bij een familierobot waar je kinderen een spelletje mee kunnen spelen.

Quis custodiet ipsos custodes? Wie bewaakt de bewakers?

Ik hoop dat ik u niet al te veel ontmoedigd heb. Mijn eigen onbehagen verdwijnt doorgaans net zo snel als het komt. Maar toch. Conversational AI: het is prachtig en zorgwekkend tegelijk. En die zorgwekkende kant, daar moeten we wat mee.

Dit kunnen we niet overlaten aan big tech, zo tonen de recente affaires bij Google en Facebook maar weer aan, maar ook niet alleen aan de overheid of professionele ethici. Ethical stewardship, dat is iets wat ons allemaal aangaat. Dit vraagt van ons allemaal, vanuit onze eigen rol, om kritisch te zijn op datgene wat we, vanuit de allerbeste intenties, ontwerpen, ontwikkelen en de wereld in sturen.

Pascale Fung gaf in haar keynote op Interspeech (<https://www.interspeech2021.org/keynotes>) een schitterend voorbeeld van hoe spraaktechnologen en andere wetenschappers deze rol kunnen invullen. Ook als conversation designer ben je niet alleen een user advocate in de traditionele zin, maar kun je ook een rol pakken in het voorkomen van gender stereotyping en in het uitdragen van transparantie en diversiteit.

Deze Dixit is een weerslag van dit gedachtegoed: een bonte verzameling van verschillende perspectieven op ons prachtige vakgebied. Ik wens u veel leesplezier! •

IK BEGRIJP NIET DAT JE ME NIET BEGRIJPT KOETJES, KALFJES EN ALLES DAAROMHEEN

We vragen ons al lang niet meer af wat de toegevoegde waarde van een chatbot of voice assistant is. De mijne maakt me 's ochtends wakker, vertelt me wat mijn agendapunten van de dag zijn en helpt me mijn telefoon te vinden als ik deze weer eens ergens heb laten slingeren. Als ik iets online bestel, legt Daan me uit wanneer ik mijn pakketje kan verwachten en als ik een item wil terugsturen, maakt Billie het retourlabel voor me aan. Maar mogen conversational designers ook snedige opmerkingen in een chatbot stoppen? Jazeker, dat kan toegevoegde waarde hebben.

Door:
Esha Metiary,
Crossphase

Steeds vaker komt het voor dat de limieten van een bot getest worden. Niet langer nemen mensen genoegen met de standaardantwoorden, maar stellen ze ook vragen als "Wie heeft je gemaakt?", "Ben je een jongen of een meisje?" en ook "Wat vind je van Zwarte Piet?". Vraag ik Billie of hij religieus is, dan antwoordt hij - teleurstellend genoeg - met "Ik snap niet echt wat je bedoelt. Kun je je vraag anders verwoorden?".

Anticiperen

We leven in 2021 en het ik-begrijp-het-niet-antwoord zou al lang een ding van het verleden moeten zijn. Zo vooruitstrevend als conversational AI is, zo achterhaald zijn de antwoorden die de meeste huidige bots gebruiken. Er bestaan tegenwoordig zeer uitgebreide datasets waarmee een bot kan anticiperen op koetjes, kalfjes en de rest van een boerenstal, dus dat zou in ieder geval geen belemmering meer mogen zijn voor het schrijven van strakke smalltalk.

In een wereld waar we stukje bij beetje activistischer worden en de podia om onze standpunten te verkondigen alomtegenwoordig zijn, ligt bij ons de taak om niet alleen op persoonlijk vlak alert te zijn op maatschappelijke misstanden en ongelijkheden, maar ook om dit door te voeren in conversational AI. Je mag vandaag de dag verwachten op elk mogelijk terrein getoetst te worden. Zeker digitaal, zeker wanneer er een bot in het spel is.

Maar ja, niet alles wat een bot zegt, valt in goede aarde. Waarom zou je dat stukje code dan een gevat antwoord laten geven op een brutale vraag? Ikzelf voel een maatschappelijke verplichting, en voor een

POSITIEVE ERVARING

LOYALITEIT

OMZET/VERKEER

bedrijf mag dat evengoed gelden. Een deel van het publiek van een bot zal ogenblikkelijk afhaken bij een gevat antwoord, maar het publiek dat blijft hangen wordt gelijk van grotere waarde. Chatbots worden ingezet voor een hogere klantbetrokkenheid, betere klantreis en optimalisatie van een klantervaring. Smalltalk, het voorbeeld van Billie dat we hierboven zien, is niet alleen een handig hulpmiddel om van een botje een sterke bot te maken; het speelt een cruciale rol in het maken van die klik tussen mens en bot. Positieve ervaringen leiden tot loyaliteit en loyaliteit leidt tot meer omzet of verkeer. Simpel.

Kijk, er zijn zeker voorbeelden waarin Billie een vraag mag afwijzen. Een vraag over sport of religie ligt totaal niet in zijn straatje, maar had hij geantwoord dat hij een trouwe aanhanger van het botsleeteam is, dan had hij een fan voor het leven gehad. De kwestie ligt anders voor bijvoorbeeld bots van grote sportmerken. Daar ben ik zeker geneigd dat machientje aan de tand te voelen over genderkwesties of Black Lives Matter. Want ja, zegt het merk er iets over, dan verwacht ik dat de bot het ook doet.

De uitdaging

Het bedenken van sterke antwoorden op zogenaamde smalltalkvragen kan of de levensvreugde of de vloek van een conversation designers bestaan zijn. Maar we kunnen meer. We moeten meer. Aan ons, conversation designers, de uitdaging en de taak om conversational AI de 21e eeuw in te slepen.

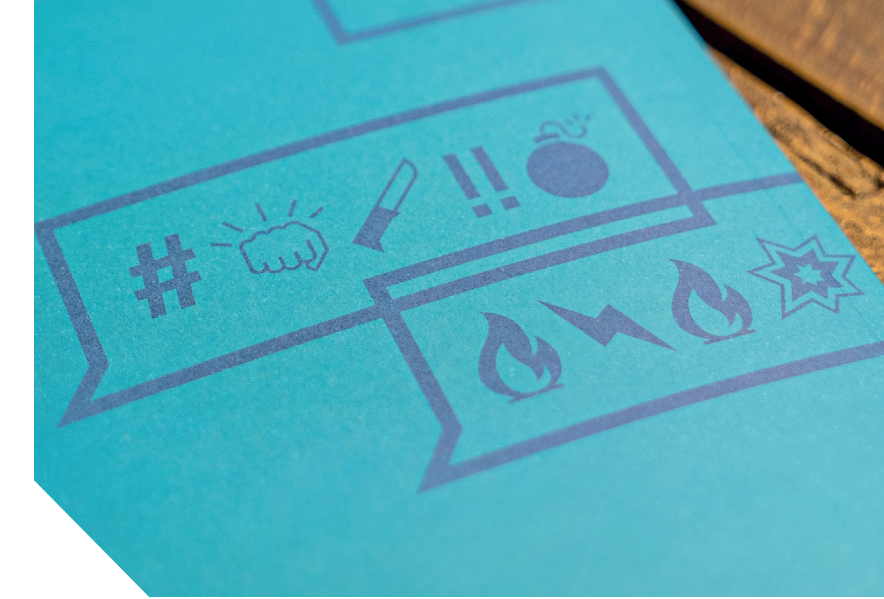
Die uitdaging is, smalltalktechnisch, tweeledig. Enerzijds moeten we mensen leren het onverwachte te verwachten. Chatbotgebruikers mag best wat bijgebracht worden, zonder dat de bot

daarbij in het neerbuigende vervalft. Geen persoonlijke aanvallen, geen scheldwoorden; blijf kort en bondig en hou de eer aan de bot. Een vraag als "Wat vind je van Black Lives Matter" kan simpelweg beantwoord worden met "Precies dat. Black lives matter."

Het is prima om discussie te willen vermijden. De basisregels van een discussie zijn immers dat beide partijen ontvankelijk zijn voor de argumenten van de ander, hetgeen in een online discussie zelden het geval is. Een eerlijke discussie zit er dan sowieso niet in. Een bot mag dan prima het gesprek afkappen en laten weten zich niet te kunnen vinden in de input.

Anderzijds moeten wij zelf het lef tonen om die snedige opmerkingen in de bot te stoppen. Het kan zijn dat er tegen wat heilige huisjes geschopt moet worden. Maar zonder pijn geen groei en zonder wrijving geen glans. Sluit "Ik heb je niet begrepen" op in een zuurvast en laat het alsjeblieft nooit meer het daglicht zien, en bedenk een antwoord dat waardig genoeg is. Voor jezelf en voor de maatschappij.

Mijn Google Assistant hoeft me 's ochtends na het opdreunen van mijn agendapunten en het nieuws niet te gebieden om in plaats van een



mooie dag een woke dag te hebben, maar de digitale assistent zou me zeker wel mogen corrigeren als ik een ongepaste of seksistische opmerking maak. Van een bedrijf dat betrokken is bij welke activistische beweging dan ook, verwacht ik een snerpend goed statement als ik de bot vraag naar (sociaal-)maatschappelijke issues.

We vragen ons al lang niet meer af wat de toegevoegde waarde van een chatbot of voice assistant is. De vraag is nu hoe wij, conversation designers, van toegevoegde waarde voor een chatbot of voice assistant kunnen zijn. •

/instituut voor de
Nederlandse taal/

taalmaterialen

Het Instituut voor de Nederlandse Taal (INT) stelt tekstverzamelingen, woordenlijsten, wetenschappelijke woordenboeken, spraakcorpora en taal- en spraaktechnologische software beschikbaar via:

taalmaterialen.ivdnt.org

GENDERVRIJE INCLUSIEVE TECHNOLOGIE, OFTEWEL WAAROM ZIJN ALLE CHATBOTS EIGENLIJK VROUW (EN IS DAT ERG)?

We spraken met Marion Mulder, FutureMaker en Digital Strategist (MuldiMedia) en een van de bestuursleden van Women in Conversational Nederland. Marion spreekt en adviseert regelmatig over gender-vrije en inclusieve digitale technologie, over toezicht op AI, algoritmen en digital humans, en over AI ethics. Wij zijn razend benieuwd naar wat dat allemaal is.

Interview met
Marion Mulder,
MuldiMedia

Om maar gelijk met de deur in huis te vallen:

Waarom zijn alle chatbots eigenlijk vrouw?

Heb je even ...

Toen ik een jaar of 5 geleden in de conversational AI rolde, viel me één ding sterk op: de meeste chatbots en voice applicaties hebben standaard een vrouwenstem en/of een vrouwennaam. Ik ben feminist en vroeg me af, waarom heeft technologie eigenlijk een gender? Is dat nodig? Is dat erg? Het is een interessante reis en ook ik leer nog iedere dag iets bij. Kort gezegd, heeft het te maken met bias en binair denken? Een gedachtlijn is dat we service nou eenmaal als iets vrouwelijks zien. Maar er schuilt ook een jaren 50 nostalgisch beeld van de ideale huisvrouw onder. En als je ziet wat veel van de (initiële) smart devices en voice tools doen, dan is dat huishoudelijke zaken automatiseren. Het is begrijpelijk, maar tegelijkertijd vind ik het ook zorgelijk. We hebben als mensen nog behoorlijk wat vooroordelen en als we die vooroordelen automatiseren, worden ze alleen maar versterkt. Ik kan hier uren over doorgaan, maar ik verwijs graag naar het boek 'The Smart Wife' van Yolande Strengers and Jenny Kennedy. Hierin wordt het heel genuanceerd en wetenschappelijk onderbouwd uitgelegd. En ook het boek 'Invisible Women' van Caroline Criado Perez is een goudmijn aan voorbeelden over hoe sterk de wereld standaard is ingericht op de man. Dat is niet per se allemaal de schuld van de standaard man (fitte witte man, 1.75m, normaal postuur 30er-40er). Maar iedereen die niet aan dat signalement voldoet, wordt daarmee subtiel of minder subtiel benadeeld.

Let wel: Ik ben ook niet tegen vrouwelijke technologie hè! Maar ik vind wel dat je er als organisatie over nagedacht moet hebben. En ik denk ook dat er veel meer gendervrij omgegaan kan worden met bijvoorbeeld het taalgebruik, nog los van hoe de applicatie zelf praat en klinkt. Zo stond er in de handleiding van een organisatie waar ik een tijdje werkte dat men de ondernemer aanspreekt met hij (bias: de meeste ondernemers

zijn man). Daar vind ik wat van. Ik ben zelf ook ondernemer en heb een eenmanszaak (nog zo'n gendered woord). Ik voel me niet aangesproken, en zelfs een beetje geïrriteerd als ik als hij aangesproken word. Het is ook een beetje een selffulfilling prophecy. Als je maar hard en lang genoeg roept dat iets zo is, ga je er ook in geloven en het ook zo zien. Sommige mensen zeggen dan "ah joh, maak je niet zo druk". Maar vervang voor de gein vandaag in alles wat je leest en schrijft eens 'hij' door 'zij'. En kijk eens wat dat met je doet.

Wat bedoel je met gendervrij?

We denken nog vaak heel binair over gender man-vrouw. Er zit ook bewust of onbewust een hiërarchie in en een hele set aan verwacht gedrag aan vast. Het probleem met binair is die hiërarchie; het is de impliciete aanname dat de een beter is dan de ander; man/vrouw, wit/zwart, able/disabled, in een relatie/single, straight/gay, left brain/right brain, de economie/environment, hooggeschoold (theoretisch opgeleid)/laaggeschoold (praktisch opgeleid), wij/zij, et cetera.

Maar de wereld is niet binair! En wat nou als we de wereld meer gaan zien vanuit een non-binair perspectief en meer als een spectrum?

Sommige mensen gebruiken de term 'gender-neutraal' maar dat voelt als one-size-fits-no-one. 'Gendervrij' betekent aan de ene kant kijken of we iets zonder gender kunnen doen, en aan de andere kant de vrijheid om je eigen ruimte te pakken en de boel een beetje te mixen.

Ik vergelijk het ook wel met linker- en rechterhersehelften. Niemand functioneert lekker op maar een half brein. Je hebt beide nodig. En het ene moment meer dit en het andere moment meer dat. Vrijheid en flexibiliteit.

Wat kunnen organisaties doen die gendervrij willen omarmen?

Er zijn een heel aantal zaken. Maar ik zou in ieder geval willen pleiten voor meer inclusief taalgebruik.

Code D&I heeft een fantastische handreiking geschreven voor een veilige, inclusieve en toegankelijke taal:

(https://codedi.nl/wp-content/uploads/2021/04/WAARDEN_VOOR_EEN_NIEUWE_TAAI.pdf).

Als we al alleen dat doen, gaan we al zoveel meer bewustzijn creëren en zijn we inclusiever bezig.

Er zijn ook algoritmen die je een indicatie kunnen geven van hoe gendered jouw tekst is. Dat is een mooi hulpmiddel als startpunt.

Verder gaat het ook om zaken als: welke velden gebruik je op een aanmeldformulier? Is die aanhef of dat m/v-boxje echt nodig?

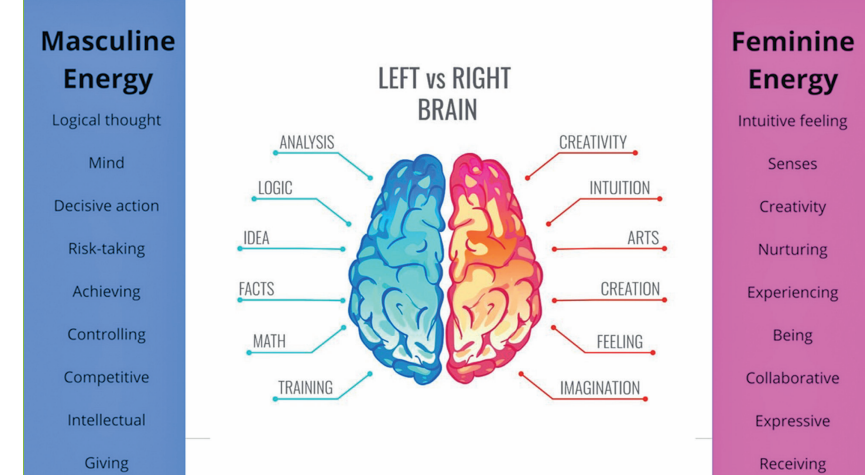
En als je datasets gebruikt om jouw systeem te trainen (of voor welk doel dan ook), zorg dan dat een van je gebruikscriteria is dat je weet hoe inclusief en intersectioneel de data zijn, wat de bron is en waarvoor de data waren verzameld. Met intersectioneel wordt bedoeld dat je niet sec kijkt naar bijvoorbeeld gender, maar ook naar kruisverbanden met andere demografische kenmerken.

Het helpt daarbij enorm als het team dat de technologie maakt en test, divers is en een goede vertegenwoordiging vormt van de doelgroepen. Er zijn legio voorbeelden van producten die live gegaan zijn en waarvan je denkt "hoe heeft dit in hemelsnaam kunnen gebeuren?". Voorkomen is beter dan genezen. En mocht je nou nog geen divers team hebben, betrek er dan netwerken en organisaties bij die dat wel zijn. Zoals Women in Conversational, Visio, Codam, OneWorld of Colourful People om er maar een paar te noemen.

Nog even een paar vragen over jouw achtergrond in conversational:

Hoe ben jij in conversational gerold?

Zo'n 5 jaar geleden werkte ik op een project bij een grote corporate organisatie en die had een startup booster programma. Een collega vroeg me of ik mee wilde pitchen (ze wilden graag een vrouw erbij). Het ging om een virtual assistant. Ja



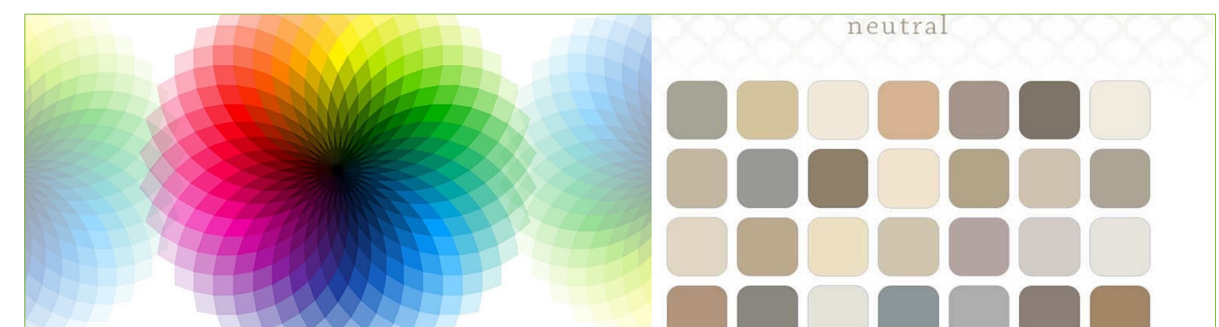
leuk! zei ik. Wat is dat trouwens, een virtual assistant? Hij legde het me uit en ik was wakker! Dat is interessant! Ik ben vervolgens gaan googelen en bij allerlei meetups aangehaakt om me te verdiepen in dit fantastische onderwerp.

Heb jij eigenlijk een achtergrond in taal?

Nou niet echt nee. Ik heb op de middelbare school wel examen gedaan in Frans en Engels. En ik heb gemerkt dat ik wel een taalgevoel heb, dat ik redelijk snel nieuwe talen oppik, – en ook in het spreken van bijvoorbeeld Nederlands of Engels met een ander accent – maar mijn grammatica rammelt, ik ben een beetje dyslectisch volgens mij, en ik maak behoorlijk veel spelfouten. Maar het mooie van het conversational vakgebied is dat het veel breder is. Ik ben bijvoorbeeld wel weer goed in informatiearchitectuur en in het herkennen van de vraag achter de vraag, en in de strategiekant. Wat ik zo mooi aan conversational vind, is dat het je dwingt om heel moeilijke processen (onder de motorkap) heel simpel te maken qua gebruikersinteractie. "Hé assistent, boek een vakantie voor mij", "Ok Marion, ik heb een weekje naar de zon op Ibiza voor je geboekt dan en dan, het staat in je agenda". Dat idee.

Any famous last words?

Ik hoop dat ik de lezers een beetje heb kunnen prikkelen. Er is nog zoveel meer te vertellen en samen te ontdekken over dit onderwerp. Dus bij deze een handreiking: heb je vragen of wil je een keer sparren, vraag me gerust. Of neem deel aan de dialoog op LinkedIn bij posts die ik deel. •



“HALLO, IK BEN ANNA, UW VIRTUELE ASSISTENT” TALIGE KENMERKEN IN CUSTOMER SERVICE CHATBOTS

De hele middag is Geeske thuisgebleven om haar nieuwe nachtkastje in ontvangst te nemen. Maar inmiddels is het avond en van de pakketbezorger ontbreekt nog elk spoor. Wat nu? Ze wil weten wanneer het nachtkastje dan wel geleverd wordt en gaat naar de website van het bedrijf. Daar popt een ballonnetje op: “Hallo, ik ben Anna, uw virtuele assistent van de meubelwinkel. Waar kan ik u mee helpen?”.

Door:
Christine Liebrecht,
Charlotte van Hooijdonk,
Utrecht University;
Florian Kunneman,
Emiel van Miltenburg,
Vrije Universiteit Amsterdam

De situatie van Geeske is exemplarisch voor het klantenservicecontact anno 2021. Een kwart van de Nederlandse organisaties zet customer service chatbots in om vragen, klachten en opmerkingen van consumenten te beantwoorden en de verwachting is dat dit in de komende jaren nog flink zal groeien (Hachmang & Keuning, 2020). Een groot voordeel van deze automatische gesprekspartners is dat ze dag en nacht bereikbaar zijn en dat ze grote hoeveelheden berichten kunnen verwerken. Maar de mogelijkheden van chatbots worden nog niet volledig benut. Klanten vrezen dat hun vragen niet goed begrepen worden en ervaren deze vorm van klantcontact als onpersoonlijk (Drift, SurveyMonkey Audience, Salesforce & Myclever, 2018). Dit kan anders. In dit artikel bespreken we vanuit talig perspectief hoe chatbots worden geïntroduceerd bij klanten, welke communicatiestijl in chatbotgesprekken aan te raden is en hoe omgegaan kan worden met miscommunicatie.

Introdactie van de chatbot

De chatbot introduceert zich als “Anna, uw virtuele assistent van de meubelwinkel”. Geeske leest dat Anna “veel vragen kan beantwoorden”. Het zal mij benieuwen, denkt Geeske. Ze tikt in “nachtkastje niet geleverd”.

Wij hebben een verkennende inhoudsanalyse uitgevoerd voor 50 Nederlandse customer service chatbots. Hieruit blijkt

dat chatbots zich op verschillende manieren introduceren bij klanten. Ze bevatten bijvoorbeeld vaak één of meerdere menselijke kenmerken, zoals een naam ('Anna' of 'Billie'), een menselijke avatar en de term “virtuele assistent”. Het is ech-

ter belangrijk dat de introductie van een customer service chatbot niet te veel menselijke cues bevat, omdat dit bij klanten de indruk kan wekken dat ze met een medewerker van vlees en bloed communiceren. Door een expliciete disclosure (“Ik ben de chatbot van de meubelwinkel”) toe te voegen aan de chatbotintroductie, communiceren organisaties op een transparante manier over de automatische gesprekspartner en worden verwachtingen van klanten gemanaged (Mozafari, Weiger, & Hammerschmidt, 2021).

Daarnaast blijkt er veel variatie te zijn in hoe organisaties in de introductie communiceren over de kennis en kunde van de chatbot. Veel chatbots geven niet aan waar ze klanten mee kunnen helpen. Ze stellen enkel de vraag “Waar kan ik u mee helpen?” Sommige chatbots geven duidelijk aan over welke onderwerpen ze meer informatie kunnen geven, bijvoorbeeld over de locatie van winkels, het assortiment en bestellingen. Een andere bevinding is dat organisaties in de introductie aangeven dat de chatbot (nog) niet alles begrijpt. Uitingen zoals “chatbot in opleiding” of “Ik ben nieuw maar raak steeds beter uitgerust om je te helpen” kunnen mogelijk de verwachtingen van klanten over de capaciteiten van de chatbot temperen.

Menselijke communicatiestijl

“Voer hier uw klantnummer in”, leest Geeske. Gevolgd door: “Uw bestelling wordt donderdag tussen 12 en 18 uur geleverd”. Het is weliswaar

een antwoord op Geeskes vraag, maar de toon waarop had wel wat persoonlijker gemogen.

Net als Geeske vinden meer gebruikers dat customer service chatbots vaak onnatuurlijk en onpersoonlijk overkomen (Drift et al., 2018). Met verschillende experimenten toonden wij aan dat de communicatie met chatbots natuurlijker en persoonlijker gemaakt kan worden door de communicatiestijl aan te passen. Deze zogeheten *conversational human voice* (Kelleher, 2009; Kelleher & Miller, 2006) bevat talige elementen van personalisatie (bijvoorbeeld “Hoi Geeske”), informeel taalgebruik (“jey!”) en uitnodigende retoriek (“Kan ik je verder nog helpen?”).

Zowel chatbots waarin een van deze categorieën als meerdere categorieën werden geïmplementeerd werden door participanten niet alleen menselijker gevonden, ze leidden ook tot positievere evaluaties van de organisatie. Echter, als de chatbot enkel elementen van uitnodigende retoriek gebruikt, dan is het wel belangrijk om rekening te houden met de mate waarin de klant bekend

is met de organisatie. Iemand die de organisatie al kent vindt deze chatbot inderdaad menselijker, maar nieuwe gebruikers vinden deze communicatiestijl wellicht ongepast (Van Hooijdonk & Liebrecht, 2021).

Miscommunicatie oplossen

Geeske wil weten of het nachtkastje bij een afhaalbalie bezorgd kan worden. “Je bericht wordt niet herkend, heb je het wel goed ingetypt?”, vraagt de chatbot. “Nou ja!”, denkt Geeske verontwaardigd. Hoe zorgt ze dat de chatbot haar wél begrijpt?

Natuurlijk is het mooi meegenomen als chatbots veel uitingen van klanten kunnen herkennen, maar repareren van miscommunicatie is minstens zo belangrijk. Er zijn verschillende strategieën van reparatie, variërend in wie van de gesprekspartners de miscommunicatie adresseert en eventueel herstelt, en op welk moment dit wordt gedaan (Schegloff, 1992). Ook kan de chatbot de miscommunicatie op allerlei manieren formuleren. Vanuit de literatuur over crisiscommunicatie kunnen we deze strategieën plaatsen op een continuüm, van defensief (de schuld leggen bij de klant) naar accommodatief (de schuld meer bij zichzelf leggen; Coombs, 2007).

Veel customer service chatbots adresseren miscommunicatie op een generieke manier, net als in het voorbeeld. Geeske krijgt daarmee weinig handvatten om de miscommunicatie te herstellen. Het zou effectiever zijn als de chatbot inzicht gaf in de woorden die niet herkend worden (Ashktorab, Jain, Liao, & Weisz, 2019). Of als de chatbot simpelweg een antwoord op de vraag zou 'gokken', wat de gebruiker ook meer inzicht geeft in de mogelijke weg naar herstel (Cho & Rader, 2020). Het risico is wel dat de chatbot minder vertrouwd wordt en de gebruiker niet de moeite neemt om tot reparatie over te gaan. Ook de defensieve houding van Geeskes chatbot draagt niet bij aan een coöperatief gesprek. De chatbot had beter de schuld van de miscommunicatie bij zichzelf kunnen leggen met “Sorry, ik begrijp uw bericht niet goed”.

Eind goed, al goed?

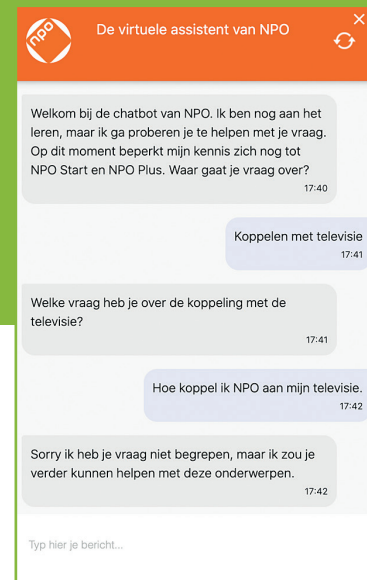
Bij het schrijven van chatbotconversaties kunnen talige elementen gebruikt worden die de verwachtingen en ervaringen van klanten kunnen sturen. De introductie van de chatbot en zijn communicatiestijl in het gesprek kunnen een positieve invloed hebben op de gehele klantervaring zodat mensen het gevoel hebben op persoonlijke en natuurlijke wijze geholpen te worden, ondanks dat hun gesprekspartner niet van vlees en bloed is. Mocht er desondanks toch miscommunicatie optreden, dan kan de chatbot op een coöperatieve manier aangeven dat hij het niet begrijpt en de gebruiker de kans geven om het misverstand te repareren. Eind goed, al goed.

'Pakket ophalen' herkent de chatbot wél, ziet Geeske. “U bent donderdag welkom op ons ophaalpunt aan de Schoolstraat.” Nu kan ze met een gerust hart gaan slapen. Enthousiast typt ze terug: “Bedankt!!!”

Verantwoording

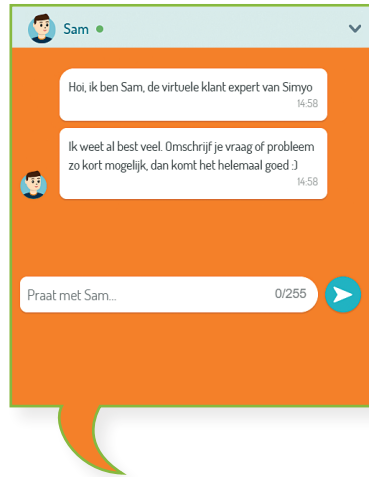
Deze inzichten hebben we opgedaan in het door NWO-gefinancierde Smooth Operators project waarin wetenschappers en maatschappelijke partners samen de technische en sociale uitdagingen onderzoeken bij het implementeren van customer service chatbots.

Zie <https://www.conversationalagentsresearch.com>



Literatuur

Ashktorab, Z., Jain, M., Liao, Q. V., & Weisz, J. D. (2019). Resilient chatbots: Repair strategy preferences for conversational breakdowns. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-12).



Cho, J., & Rader, E. (2020). The Role of Conversational Grounding in Supporting Symbiosis Between People and Digital Assistants. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW1), 1-28.

Coombs, W.T., 2007. Protecting organization reputations during a crisis: The development and application of situational crisis communication theory. *Corporate Reputation Review*, 10(3), 163-176.

Drift, SurveyMonkey Audience, Salesforce, Myclever: The 2018 State of Chatbots Report. How chatbots are reshaping online experiences. Op 1 september 2019 geraadpleegd op: <https://www.drift.com/wp-content/uploads/2018/01/2018-state-of-chatbots-report.pdf>

Hachmang, D. & Keuning, A. (2020). *Stand van Webcare 2020*. Op 9 september 2021 geraadpleegd op:

<https://www.upstream.nl/wp-content/uploads/2020/06/Stand-van-Webcare-juni-2020.pdf>

Kelleher, T. (2009). Conversational voice, communicated commitment, and public relations outcomes in interactive online communication. *Journal of Communication*, 59(1), 172-188.

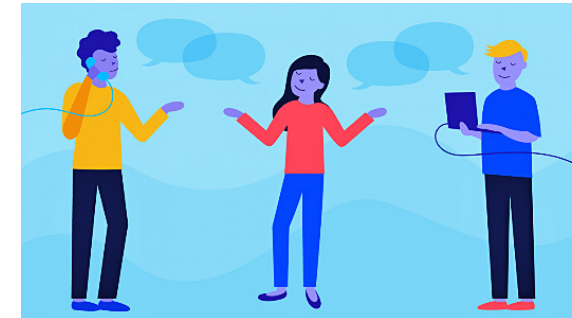
Kelleher, T., & Miller, B. M. (2006). Organizational blogs and the human voice: Relational strategies and relational outcomes. *Journal of Computer-Mediated Communication*, 11(2), 395-414.

Mozafari, N., Weiger, W. H., & Hammerschmidt, M. (2021). Trust me, I'm a bot—repercussions of chatbot disclosure in different service frontline settings. *Journal of Service Management*, Vol. ahead-of-print No. ahead-of-print. <https://doi.org/10.1108/JOSM-10-2020-0380>

Hooijdonk, C. van & Liebrecht, C. (2021). Chatbots in the tourism industry: the effects of communication style and brand familiarity on social presence and brand attitude. In *Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '21 Adjunct)*, June 21–25, 2021, Utrecht, Netherlands. ACM, New York, NY, USA, 7 pages. <https://doi.org/10.1145/3450614.3463599>.

Schegloff, E. A. (1992). Repair after next turn: The last structurally provided defense of intersubjectivity in conversation. *American journal of sociology*, 97(5), 1295-1345. •

INCLUSIEVE AUTOMATISCHE SPRAAKHERKENNING



Introductie

Automatische spraakherkenning wordt steeds meer gebruikt, bijvoorbeeld in digitale assistenten, smart speakers, call centers en zoekmachines. Het is een makkelijke manier om met machines en computers te communiceren. Om dit mogelijk te maken moet de automatische spraakherkenner kunnen omgaan met de grote variatie in de manier waarop mensen spreken.

De huidige best werkende spraakherkenners zijn gebaseerd op deep neural networks (DNNs). Deze worden gezien als zeer objectief. Gegeven een getraind model en gegeven een taak en dataset zullen DNN's altijd dezelfde uitkomst geven, terwijl de uitkomst bij mensen op dezelfde data en taak afhankelijk kan zijn van heel veel verschillende factoren. Ik wil hier het punt maken dat spraakherkenners dan wel objectief zijn, maar dat recente resultaten laten zien dat ze niet in staat zijn om de spraak van iedereen even goed te herkennen, wat grote gevolgen heeft voor grote groepen mensen in onze samenleving.

Bias tegen verschillende spreker groepen

Automatische spraakherkenners worden standaard getraind op de spraak van moedertaalsprekers die een 'standaard' variant van de taal spreken, ouder dan 18 zijn (en vaak jonger dan 65 jaar) en geen spraakafwijking hebben (de 'normale spreker'). Hierdoor wordt niet alleen onbedoeld de spraak van vreemdtaalsprekers veel slechter herkend dan die van moedertaalsprekers, maar ook die van mensen met een (sterk) regionaal accent of sociolect. Het verschil in herkenprestatie van een spraakherkenner voor spraak van een 'normale' spreker en

spraak van een niet-gemiddelde spreker(groep) noemen we 'bias'.

Daarnaast zijn er grote verschillen in hoe goed spraakherkenners de spraak van mensen uit verschillende leeftijdsgroepen kunnen herkennen en hebben spraakherkenners grote problemen met spraakafwijkingen, bijvoorbeeld veroorzaakt door dysarthrie, beroertes en orale kanker. En dat terwijl deze mensen vaak juist gebaat zijn bij het kunnen gebruiken van hun stem voor het bedienen van applicaties en apparaten. Een recente studie¹ uit de Verenigde Staten laat zien dat voice assistants de spraak van blanke sprekers veel beter kunnen herkennen dan die van Afro-Amerikaanse sprekers, waardoor de raciale scheidslijn ook op dit vlak blijft bestaan. Verschillende onderzoeken in verschillende talen laten zien dat er ook verschillen bestaan tussen hoe goed spraak van mannen en spraak van vrouwen wordt herkend (overigens zonder eenduidig patroon).

Oorzaken van bias

Er zijn verschillende oorzaken die deze bias tegen spreker groepen veroorzaken. Ten eerste is er de samenstelling van de data die gebruikt worden om de spraakherkenners te trainen: als de trainingsdata geen spraak van bepaalde

Door:
Odette
Scharenborg,
SpeechLab /
Multimedia
Computing
Group,
Technische
Universiteit
Delft

taalkundige werkzaamheden voor taal- en spraaktechnologie, conversation design

lexicons en datasets

part-of-speech tagging

conversation design, intent classification

Martha Hofman

tel. 06 19 858338

mahofman@hetnet.nl

www.marthahofman.nl

Wat is bias?

Stel je wilt een systeem bouwen waarmee je automatisch klinkers kunt herkennen, bijv. de /ɔ/, o, e, ε/. Je bouwt twee verschillende systemen. De fout herkende klinkers door beide systemen staan hieronder aangegeven in het rood:

In beide systemen worden 25% van de klinkers fout herkend. In systeem 1 zijn de fouten verdeeld over de verschillende klinkers. In systeem 2 worden alle fouten gemaakt voor de /ɔ/. Dit laatste, waarbij er een hele klasse (of groep) slechter herkend wordt dan de andere klassen (of groepen), noemen we bias.

¹ Koenecke A, et al. (2020) Racial disparities in automated speech recognition. *Proceedings of the National Academy of Sciences*, 117(14):7684–7689.

sprekergroepen bevat, dan zal dat type spraak in de regel ook slecht herkend worden. Daaraan gerelateerd: als de trainingsdata bestaan uit spraak van verschillende spreker-groepen, maar één spreker-groep is duidelijk anders dan die van de overige spreker-groepen, dan zal die 'afwijkende' spreker-groep meestal slechter herkend worden. En zelfs als de spraakherkenner getraind wordt op spraak van een ondergerepresenteerde groep (bijvoorbeeld op spraak van Afro-Amerikaanse sprekers) dan laat de accuraatheid van de spraakherkenner nog te wensen over door de grote variabiliteit van de spraak in zo'n spreker-groep.

Deze grote variabiliteit doet zich ook voor op het niveau van het individu: taalleerders (zowel kinderen als vreemdetaal-leerders) en mensen met een spraakafwijking laten een grotere variatie binnen hun spraak zien wat een negatief effect heeft op de prestaties van de spraakherkenner. Verder hebben we recent aanwijzingen gevonden dat de handmatige transcripties van verschillende spreker-groepen verschillen: incorrecte uitspraken van kinderen worden vaker gecorrigeerd naar de 'correcte' uitspraak dan die van niet-moedertaalsprekers. Dit leidt tot meer out-of-vocabulary (OOV) woorden voor niet-moedertaalsprekers. Tot slot zien we dat de grootte en de richting van de bias afhankelijk is van het gekozen algoritme en architecturen van de modellen.

Inclusieve automatische spraakherkenning

Om ervoor te zorgen dat iedereen gebruik kan maken van automatische spraakherkenning ongeacht hoe men spreekt of welke taal men spreekt, zijn wij recent een nieuwe onderzoekslijn begonnen die we 'inclusieve spraaktechnologie' hebben genoemd, met een sterke focus op inclusieve automatische spraakherkenning. In het eerste project is het doel om de bias in een state-of-the-art DNN-gebaseerde spraakherkenner bloot te leggen zodat we proactief bias kunnen verminderen en zoveel mogelijk

voorkomen². Naast het kwantificeren van de grootte van de bias voor gender, regionale accenten (inclusief het verschil tussen Nederlands en Vlaams), leeftijd en vreemdetaal-leerders, onderzoeken we de oorsprong van de bias en vergelijken we bias in twee verschillende architecturen, en in twee talen: Nederlands (we hebben hier het Jasmin-corpus voor gebruikt) en Mandarijn Chinees (Magic Data Mandarin Chinese Read Speech Corpus).

Samenvattend zien we in ons onderzoek dat de bias taal- (of wellicht database-)afhankelijk is: voor het Nederlands zien we een genderbias (vrouwenspraak wordt beter herkend dan man-spraak) maar voor het Mandarijn zien we geen bias. Verder zien we in hybride modellen een kleinere bias dan in end-to-end modellen voor het Nederlands en het omgekeerde patroon voor het Mandarijn Chinees. Beide talen laten grote biases zien voor spreker-groepen met sterke regionale accenten. Voor het Nederlands zien we verder een bias bij leeftijd: spraak van ouderen (ouder dan 60 jaar) wordt slechter herkend dan dat van tieners (12-16 jaar) en de spraak van kinderen (7 tot 11 jaar) wordt veruit het slechtst herkend. Tot slot zien we dat de geobserveerde biases groter zijn in meer spontane spraak (human-machine interaction spraak) dan in voorgelezen spraak. Vermoedelijk leidt de grotere vrijheid in uitspraak en woordkeus bij meer spontane spraak tot grotere verschillen met de getrainde akoestische en taalmodellen en dus tot een grotere bias. Daarnaast zien we dat de samenstelling van de trainingsdata een belangrijke bron van de bias is. Als er geen of weinig trainingsdata zijn van een spreker-groep, dan wordt deze slecht herkend.

Afsluitend, automatische spraakherkenners werken goed voor de 'normaal' Nederlandse sprekers, maar het is noodzakelijk dat deze gaat werken voor heel veel meer, het liefst alle, sprekers van het Nederlands. Oftewel, automatische spraakherkenning moet inclusief worden. •



² S. Feng, O. Kudina, B. Halpern, O. Scharenborg (in preparation). Towards inclusive automatic speech recognition.

INCLUSIVE CONVERSATIONAL EXPERIENCES

OVER WAT WE KUNNEN LEREN VAN BLINDE EN SLECHTZIENDE VOICEGEBRUIKERS

Introductie

Stel je voor. Je komt thuis van je werk en hebt eigenlijk geen zin om te koken. Je besluit iets te eten te bestellen en opent de site van een restaurantje in de buurt. Je weet eigenlijk al wat je wilt hebben en klikt je door alle verschillende stappen heen. Uiteindelijk heb je besteld; laat de avond maar beginnen.

Maar wat als je een beperking hebt die het bedienen van een website lastig of soms onmogelijk maakt? Neem dezelfde situatie maar je bent blind. Het is, met behulp van software die het scherm voorleest, ook mogelijk om (zonder het scherm te zien) de bestelling te plaatsen. Alleen duurt het dan wel dubbel zo lang. Wat als je de website vervangt door een slimme speaker? In plaats van scrollen en klikken op het scherm vraag je rechtstreeks met je stem om je favoriete pizza. Als je blind bent, bieden dit soort voice toepassingen uitkomst op veel gebieden. Maar als je geen beperking hebt, kan het net zo goed efficiënt werken. Je zit in de auto en bedenkt alvast dat je zin hebt in iets lekkers. Hoe makkelijk is het dan om niet naar het scherm te hoeven kijken en eyesfree en handsfree via voice te bestellen?

Voice als inclusief startpunt

Voice interfaces zijn voor een grote groep mensen met een beperking de ideale manier om online zaken gedaan te krijgen. Praten tegen interfaces haalt drempels weg waar mensen nu nog vaak tegenaan lopen bij het gebruik van apps en websites op schermen waar je zelf moet klikken. Het is dan ook zo dat mensen met een visuele beperking al jaren de voorkeur hebben voor voice toepassingen zoals Siri en Google Assistant als het gaat om het snel uitvoeren van eenvoudige taken. Denk aan het maken van een agenda-afspraak, een wekker zetten, een route starten of het nieuws lezen.

Bij Visio (expertisecentrum voor blinde en slechtziende mensen) ondersteunen wij mensen met een visuele beperking van allerlei leeftijden om mee te komen in de digitale maatschappij. Naast advies over pc-gebruik, smartphones en tablets zijn we steeds meer gericht op voice en

hoe dat het leven van mensen makkelijker kan maken.

Met deze unieke gebruikerservaringen in ons achterhoofd adviseren wij bedrijven over de kans die voice biedt om hun diensten ook via dit nieuwe kanaal bereikbaar te maken. En niet alleen voor mensen met een visuele beperking maar ook voor andere groepen zoals oudere volwassenen zonder veel digitale ervaring. We leren hierbij van de ervaringen en behoeften van mensen met een visuele beperking, die vaak al veel beter weten wat werkt én wat nog beter kan met voice. De wensen verschillen niet van wat andere mensen zonder beperking willen, namelijk dingen op een zo makkelijk mogelijke manier gedaan krijgen.

Van webshop naar voice experience

Onlangs hebben we een concreet voorbeeld uitgewerkt met studenten Communicatie en Multimedia Design van de Hogeschool van Amsterdam. Zij onderzochten samen met mensen met een visuele beperking hoe je een webwinkel, wat vaak toch wel een complexe interactie is, in een voice interface zou kunnen vertalen. Daarbij gingen we uit van Coolblue, want iedereen wil toch het gemak ervaren van zo'n grote webshop?

Van probleem naar kans

Zoals je mag verwachten bij goede designers, brachten zij de huidige obstakels in kaart door in gesprek te gaan met blinden en slechtzienden over hun ervaringen met webshoppings. Maar

Door:
Timon van Hasselt,
Jesse Wienholts en
Stefan Laureijssen,
Koninklijke Visio



we daagden elkaar vervolgens uit. Ga eens out of the box denken? Hoe zou een ideale voice experience bij CoolBlue eruitzien voor jouzelf (goedziend of slechtziend)? Dat bleek anders te zijn dan simpelweg alle mogelijkheden van de huidige webshop zoveel mogelijk toegankelijk te maken in een voice versie. Het startpunt van de designers werd door de interactie met échte gebruikers 180 graden gedraaid.

Het advies aan CoolBlue: spiegel je voice user interface aan de beleving in je fysieke winkels. Ontwikkel een digitale verkoopmedewerker die met de klant meedenkt en vraagt "kan ik je helpen of wil je even rondkijken"? En die naargelang de fase waarin de gebruiker zich in de zoektocht bevindt, de juiste informatie geeft en subtiel doorvraagt. Compacte informatie bij het zoeken en vergelijken en gedetailleerde informatie bij het uiteindelijke kiezen. Door op deze manier te kijken naar je oude product (website) en met super users van voice onderzoek te doen naar wat zij écht willen, zorg je voor een betere toegankelijkheid van je dienst voor iedereen. Wat moet er gebeuren om voice écht inclusief te maken?

Voice is nu al een veelbelovend kanaal en we verwachten dat het alleen maar beter toegankelijk wordt. Maar we hebben uiteraard nog wel wat wensen. We merken dat de standaardcommando's om een voice assistent als Siri of Google in actie te laten komen niet altijd heel natuurlijk zijn voor eindgebruikers.

Een bekend probleem, ook als je niet blind of slechtziend bent, is dat de spraak opdrachten vaak niet direct begrepen worden of dat de exacte formulering van de spraakopdracht nodig is.

Het zou makkelijk op te lossen moeten zijn voor de grote leveranciers van dit moment (Google, Apple en Amazon), door de intentie achter de spraakopdrachten beter te begrijpen

(intent-matching). Of door te vragen wat de vraagsteller bedoelt in plaats van te zeggen "dat begrijp ik niet" of "dit heb ik gevonden op het web".

Actions of skills

Voice apps die derde partijen kunnen maken voor slimme assistenten zijn daarin een uitkomst. Als de action is gestart door de gebruiker, biedt het de ontwikkelaars de optie om de gebruiker ook beter te begrijpen en een eigen tone of voice te gebruiken. Het zou wenselijk zijn als het ook hier wat natuurlijker was om een action te activeren, namelijk zonder "praat met". We merken dat de "praat met" commando's door weinig gebruikers zelf ontdekt worden. In een ideale situatie zou je via een spraakopdracht "wat kost een Google Nest Hub bij CoolBlue" naadloos en direct in de action van CoolBlue terecht willen komen bij de prijs van deze slimme speaker met een beeldscherm.

Daarover gesproken, we vinden de ontwikkeling van de smart speakers met scherm heel interessant. Het blijft voice first, maar het is met name voor minder digivaardige mensen een eenvoudig tablet-alternatief. Ook voor slechthorende of slechtziende mensen is aanvullende visuele informatie op het scherm soms meer toegankelijk.

Call to action

Voice maakt je dienstverlening als bedrijf voor iedereen toegankelijker. We denken graag mee over hoe we voice en conversational interfaces kunnen gebruiken om ook jouw product leuker, makkelijker én toegankelijker in gebruik te maken voor iedereen, waarbij we leren van de experts die voice al jaren gebruiken. We dagen je bij deze uit om bij je volgende innovatie toegankelijkheid eens mee te nemen vanaf het begin van het ontwerpproces, zodat in de toekomst echt iedereen gebruik kan maken van je product of dienst. Neem gerust contact op via visiolab@visio.org.

PERSOONLIJKHEID IN CONVERSATIONAL INTERFACES

Misschien ken je het wel: een stopcontact dat vriendelijk naar je lijkt te lachen, een roepende boom of een blijde putdeksel.



Het herkennen van gezichten in levenloze objecten is een bekend psychisch verschijnsel dat 'pareidolia' wordt genoemd. Het gaat zelfs nog verder dan alleen het zien van gezichten in voorwerpen: we kennen er ook direct gevoelens en emoties aan toe.

Als we naar stemmen luisteren zonder daar een gezicht bij te zien gebeurt er iets soortgelijks; we vormen ons onmiddellijk een beeld van de spreker, zowel van het uiterlijk als van de persoonlijkheid.

Pareidolia en conversational interfaces

Onze gewoonte om aan levenloze objecten en aan gezichtsloze stemmen een persoonlijkheid toe te kennen zorgt ervoor dat we dit voor conversational assistants ook doen. We hebben onbewust direct een mening over de assistant: die is vriendelijk, kortaf, slim, dom, behulpzaam... dit doen we zelfs als we weten dat we met een robot interacteren. Het betekent dat een chat-

bot of een voice assistant altijd een persoonlijkheid heeft, of die nu bewust ontworpen is of niet. Natuurlijk heb je als merk liever wél controle over hoe je assistant door de gebruiker wordt waargenomen. Daarom is het belangrijk om al vroeg in het proces van het ontwerpen van de assistant aan de slag te gaan met het opstellen van een persoonlijkheid die bij je merk en gebruiker past. Dit heeft een aantal belangrijke voordelen:

- De brand voice en waarden van het merk worden gehandhaafd en de tone of voice wordt doorgevoerd;
- De assistant komt betrouwbaar(der) over want deze heeft een consequent karakter. Hierdoor weet de gebruiker waar hij aan toe is;
- Als er meerdere conversation designers aan je bot werken dan voorkomt het hebben van een goed uitgedachte persoonlijkheid dat de flows qua schrijfstijl en tone of voice van elkaar verschillen.

Door:
Anja de Castro,
Freelance conversational UX designer
No bots no glory



De indruk die mensen zich van een bot vormen komt tot stand op basis van allerlei factoren, waaronder:

- Stem
- Taalgebruik
- Aanspreekvorm
- Gebruik van emoji, gifjes en afbeeldingen in de conversatie
- Gedrag (verwijst de bot door naar andere bronnen of helpt hij je zelf, laat hij buttons zien of een open input field)
- Naam
- Avatar (afbeelding).

Op de attributen stem en taalgebruik zal ik hieronder dieper ingaan.

Stem

De stem wordt ingezet bij voice assistants (ook als ze onderdeel zijn van een multi modal interface). Uit onderzoek van Clifford en Nass* blijkt dat hoe de stem klinkt vaak meer invloed heeft op het beeld dat de luisteraar van de assistent heeft, dan dat wat de assistent daadwerkelijk zegt. Voor voice assistants geldt verder dat de snelheid van praten, de toonhoogte, de warmte en de intonatie van invloed zijn op hoe de gebruiker de assistent waarneemt.

Het is daarom belangrijk om gebruikersonderzoek uit te voeren voordat je een stem kiest. Zo is de kans groter dat je doelgroep de stem prettig vindt om naar te luisteren. Ook toets je op deze manier of de snelheid en intonatie in orde zijn en of de stem bij de persoonlijkheid past die jij voor de assistent in gedachten hebt.

Taalgebruik

Schrijven voor een conversational assistant is fundamenteel anders dan schrijven voor een website of een brochure. Ook is schrijven voor een chatbot anders dan schrijven voor een voice assistant. Omdat mensen 'niet lezen', zoals je copywriters vaak hoort verzuchten, is het bij een tekst-based assistant (chatbot) belangrijk om de kern van je boodschap aan het begin van je tekst te plaatsen. Bij een voice-based assistant is het vaak juist verstandig de kern aan het einde van je tekst te plaatsen. Deze woorden blijven namelijk vaak hangen en maken het makkelijker om een vervolgcommando aan de bot te geven. Bij een chatbot is dit minder relevant omdat je de tekst die aan de opties voorafgaat opnieuw kunt lezen, en daarnaast geven eventuele knoppen vaak weer waar ze naartoe leiden.

Bovendien kan ieder woord dat je in je teksten gebruikt invloed hebben op de manier waarop de bot wordt ingeschat. De welkomsttekst geeft bijvoorbeeld meteen een indruk van de persoonlijkheid van de bot. Denk alleen al aan het verschil tussen de begroetingen "hoi", "hallo" en "goedemiddag". Een ander voorbeeld is tutoyeren of vousoyeren; ook dit op het eerste gezicht simpele aspect heeft grote gevolgen voor de perceptie van de gebruiker.

Om te zorgen dat het taalgebruik aansluit bij de voorkeur van de klant/bezoeker/patiënt, is er een aantal zaken van belang:

1. Ken je doelgroep. Als je weet wie de gebruiker is, kun je in je taalgebruik rekening houden met bijvoorbeeld leeftijd en opleidingsniveau;
2. Zorg ervoor dat de woordkeuze is afgestemd op je merk en product. Voor de assistent van een ziekenhuis is het vooral belangrijk dat deze vertrouwen en empathie uitstraalt, terwijl de assistent van een pretpark alvast de betovering en het plezier kan overbrengen;
3. Ook hier: test met de eindgebruiker. Pas dan zie je of de manier van schrijven/spreken écht overkomt op de manier die je in gedachten hebt.

Een conversational assistant is geen mens

Dit is één van de belangrijkste user experience-aspecten als je ontwerpt voor conversational interfaces en mag daarom niet ontbreken in dit artikel: zorg dat je altijd duidelijk maakt dat je gebruiker communiceert met een bot, en niet met een 'echt mens'. De technologie is nog niet ver genoeg om een robot een conversatie te laten voeren zoals wij dat kunnen, met name op het gebied van 'repair'. Let er maar eens op hoe vaak jij zelf je zin toch verandert, of hoe vaak je midden in een zin een zinsprongje maakt. Wij mensen hebben daar geen enkele moeite mee, maar voor conversational assistants is het nog



ontzettend moeilijk om hier adequaat op te reageren. Pretenderen dat de gebruiker met een mens interacteert, is gedoemd om verwarring en teleurstelling te veroorzaken. Waarom dan niet proberen de assistent te laten klinken als dat wat het is: een computer? Zo kun je door middel van geluid of taalgebruik voorkomen dat er verwarring ontstaat. Echter, de meeste mensen verwachten toch een persoonlijke ervaring, ook al is het met een computer.

Er zijn verschillende dingen die je kunt doen of laten om duidelijk te maken dat de gebruiker interacteert met een bot:

- Gebruik geen foto van een mens als avatar voor de bot;
- Denk na over de header van je chatscherm. Zet er bijvoorbeeld 'Chatbot Sonja' of Digitale assistent Sonja' in;
- Let op de naam. Gebruik bij voorkeur een naam die niet voor mensen wordt gebruikt, zoals 'Pennenvaar' van de Efteling. Als je een naam gebruikt zoals Helen, stel haar dan voor als 'Helen, de digitale assistent van ...';
- Schrijf een duidelijke introductietekst, waarin je nog eens benadrukt dat het om een digitale assistent gaat (zie het voorbeeld hierboven).

Pas het liefst alle vier de bovenstaande aspecten toe. Dit lijkt wellicht wat overdreven, maar uit de praktijk blijkt dat ze allemaal nodig zijn om de verwachtingen te managen.

Over een aantal jaren is het gebruik van digital assistants wellicht zo ver ingeburgerd en de technologie zo ver gevorderd, dat het de standaard is dat de gebruiker wordt geholpen door een con-

versational assistant. Dan is ontwerpregel nummer 1 misschien wel dat we het de gebruikers moeten laten weten als de conversatie gevoerd wordt door een mens. Dat zal heel andere uitdagingen met zich meebrengen waar wij als conversational (UX) designers nu nog geen weet van hebben.

Om over na te denken: affinity bias

Een ander interessant psychologisch verschijnsel waar we in conversational design rekening mee kunnen houden is 'affinity bias'. Dit houdt in dat wij als mensen de neiging hebben een voorkeur te hebben voor karakters die op ons lijken. Denk hierbij bijvoorbeeld aan introvert versus extravert, leeftijd, gender et cetera.

De chatbots en voice assistants waarmee we nu interacteren, hebben doorgaans slechts één persoonlijkheid en een bijbehorende tone of voice, onafhankelijk van de persoonlijkheid van de klant/bezoeker/patiënt. Missen we hier een kans? Is iemand die extravert is eerder geneigd om iets aan te nemen, of zelfs te kopen, van een digitale assistent die ook extravert is? Natuurlijk is het allemaal niet zo zwart-wit, maar één ding is zeker: er komen mooie uitdagingen aan voor iedereen die werkzaam is in dit interessante vakgebied, dat zowel taal als ICT raakt. •

Bron:

Clifford Nass and Scott Brave, *Wired for Speech* (Cambridge, Massachusetts, MIT Press, 2007), 34

Bron afbeeldingen:

<https://www.reddit.com/r/Pareidolia/>

NEW KID ON THE BLOCK: INTENT MANAGEMENT

Een eerste chatbot is, op de keper beschouwd, zo gebouwd. Maar hoe houd je conversational applicaties werkend en onderhoudbaar? Ook als ze gaan groeien? Naast het traditionele content management vraagt conversational AI om een nieuwe beheersmethodiek voor het beheren van je klantvragen en data: intent management.

Door:
Maaïke
Groenewege,
Convocat

De beginfase: even opstarten

Het zal veel organisaties bekend in de oren klinken: je hebt je business case en budget sluitend, er is een team vrijgemaakt en je kan eindelijk aan de slag met je eerste chatbotproject. Je gaat voortvarend van start, selecteert je eerste use case, ontwerpt flows en dialogen en zet je chatbot live. Die het vervolgens niet zo heel goed doet. 'Geeft niet', hoor je van je omgeving, 'zo'n bot moet nog van alles leren'. En hee, er zit AI in. Dus dat gaat vanzelf!

De meeste chatbots hebben even tijd (en vooral data) nodig voordat ze vragen goed herkennen.

Dat is helaas een fase die je niet kunt overslaan. Tegelijkertijd is dit ook een fase vol kansen: vanaf het moment dat je chatbot live staat, krijg je echte klantvragen binnen, gesteld in iemands eigen woorden. En die vragen zijn beschikbaar voor je organisatie. Dat is goud: als je eenmaal weet wat er wordt gevraagd, dan kun je daar wat mee doen.

Op weg naar succes: train je chatbot

Als we zeggen dat een chatbot nog moet leren, dan hebben we het over intents en utterances. In de praktijk bevat een intent lange lijsten van klantvragen (utterances) die onder dat intent worden

gecategoriseerd. Deze klantvragen dienen als trainingsdata voor de algoritmes die zorgen voor een correcte match tussen klantvraag en chatbot-antwoord: Natural Language Understanding.

En in tegenstelling tot wat de hype ons doet geloven, leert een chatbot niet vanzelf. Het trainen van een chatbot is mensenwerk. Dat doe je door dagelijks de binnengekomen klantvragen te analyseren. Valt een klantvraag onder een bestaande intent? Dan wijs je hem toe aan die intent. Is het een klantvraag die verwijst naar een nog niet eerder genoemde intent? Dan kun je overwegen om een nieuwe intent aan te maken, zodat je bot ook die vraag kan beantwoorden. En als je je bot een tijdje nauwgezet traint, dan merk je al snel dat de vraagherkenning steeds beter wordt.

De eerste barstjes

In de beginfase van een chatbot is het aantal intents doorgaans nog vrij beperkt. En het toevoegen van nieuwe intents gebeurt vraaggestuurd: je ziet tijdens je dagelijkse analyse nieuwe vragen en voegt ze toe aan een nieuwe intent. Een beetje zoals een sneeuwbal: je plakt er steeds wat meer aan. Dat gaat een tijd lang best goed. Tot je op een dag merkt dat je bot onverwachte antwoorden teruggeeft. Terwijl je zeker weet dat het juiste antwoord wel degelijk in de chatbot zit. Of helemaal geen antwoorden: je bezoekers worden, hup, de fallback ingestuurd.

Prijsschieten

Om de oorzaak van dit probleem te achterhalen, duiken we weer onder de motorkap. Het matchen van klantvragen met intents is een kwestie van statistiek: er is geen harde garantie dat een op verschillende wijze geformuleerde klantvraag altijd bij dezelfde intent uitkomt, laat staan bij de juiste. De kans dat een correcte match plaatsvindt, wordt bepaald door de balans tussen voldoende utterance-variantie binnen je intent (zodat een breed scala aan klantvragen wordt gematcht) en voldoende onderscheid tussen je intents, zodat klantvragen maar met één intent kunnen worden gematcht.

Deze balans wordt belangrijker naarmate je meer intents toevoegt: hoe meer intents je bot bevat, hoe dichter ze meestal qua betekenis bij elkaar komen te liggen, en dus hoe groter de kans dat een klantvraag bij de verkeerde intent uitkomt. Als je door de tijd heen redelijk lukraak intents toevoegt, dan eindig je uiteindelijk met een verzameling intents die uit balans is: intents kunnen overlappen, verschillen in granulariteit, of niet meer relevant zijn. Utterances die oorspronkelijk aan intent A waren toegewezen, passen inmiddels wellicht beter bij intent B die later is toegevoegd. Of wat te denken van vergelijkbare klantvragen die door verschillende analisten aan verschillende intents worden toegewezen? Met duizenden

klantvragen en tientallen intents die worden aangeraakt door meerdere teams ligt wildgroei op de loer.

Checklist intent management

Hoog tijd dus voor een methodische aanpak voor het consistent en gestructureerd onderhouden van grote verzamelingen intents en utterances. Intent management is nog een redelijk onontgonnen gebied binnen conversational AI en systeemondersteuning staat nog in de kinderschoenen (Botium, HumanFirst en Qbot bieden commerciële pakketten). Maar met deze praktische tips kom je ook zonder dat soort systemen een heel eind.

Proces

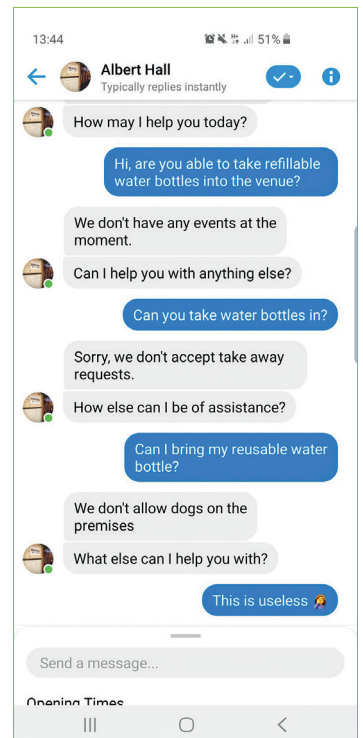
- Maak intentbeheer een onderdeel van je preventief onderhoud: neem periodiek de tijd om je hele intentset onder de loep te nemen en knelpunten op te lossen.
- Maak gebruik van slimme data-analyses: je kunt de prestaties van je chatbot testen met k-fold analyses (voor interne consistentie van je taalmodel) en blind tests (om te testen hoe goed je bot omgaat met echte vragen uit de buitenwereld).

Intents

- Zorg voor een consistente naamgeving, dat helpt je niet alleen bij het terugvinden, maar ook bij het opsporen van verdubbelingen en overlap. Bovendien dwing je jezelf om nog duidelijker te zijn bij het modelleren van je intents.
- Houd je intents in balans: zorg dat iedere intent ongeveer evenveel utterances heeft.
- Maak je intents eenduidig, zodat iedere klantvraag maar aan één intent kan worden toegewezen.

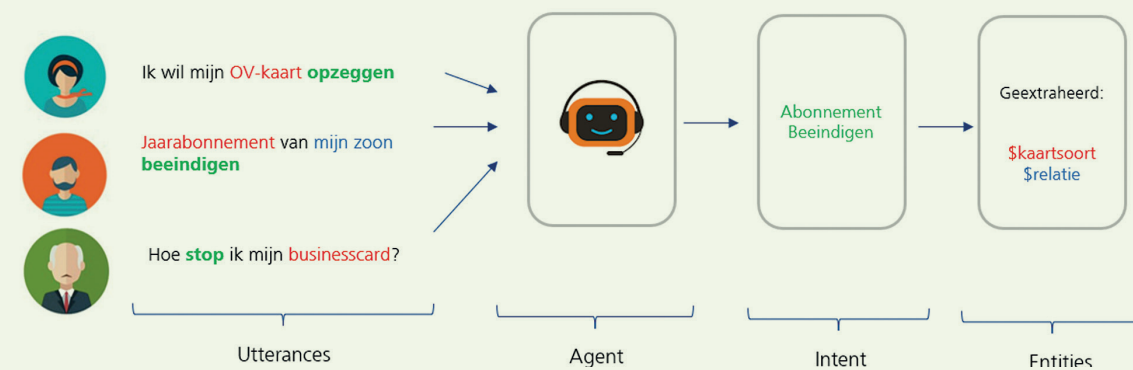
Utterances

- Gebruik zoveel mogelijk echte klantvragen uit de klantdata. Voorkom dat je zelf utterances gaat verzinnen; daarmee introduceer je je eigen bias!
- Clean je utterances: snoei overbodige woorden uit je utterances, zodat alleen die woorden overblijven die relevant zijn voor de intent (dit geldt voor je trainingsdata; testen doe je uiteraard wel met ongesnoeide klantvragen).
- Zorg binnen een intent voor variatie in utterances: zorg voor afwisseling in zinsconstructies, synoniemen, lange en korte vragen.
- Zorg tegelijkertijd ook voor een bepaalde mate van herhaling en taalkundige patronen waarmee je je NLU houvast geeft.



Onder de motorkap: utterance-intent-response

Chatbots op basis van Natural Language Understanding ('met AI') zijn als volgt opgebouwd:



Utterances

Daadwerkelijke vragen zoals een bezoeker ze stelt aan een chatbot.

Intent

De 'vraag achter de vraag'. Er zijn honderden manieren om één en dezelfde vraag te stellen. Al deze varianten van een vraag worden gegroepeerd onder één intent. Intents zijn gericht op de actie, de taak of het probleem dat een bezoeker probeert op te lossen.

Entity

Extra informatie waarmee de vraagherkenning kan worden geoptimaliseerd. Denk hierbij aan categorieën zoals 'ID-bewijs' (paspoort, rijbewijs, ID-kaart), of 'Kaartsoort' (Betaalpas, Creditcard, Gemachtigdenpas).

Response

Het antwoord dat de chatbot teruggeeft op het moment dat er een match is tussen utterance en intent, in de vorm van een dialoog.

OP ZOEK NAAR INTERPRETATIE: INDEEP

Tegenwoordig is Deep Learning niet meer weg te denken uit vakgebieden en toepassingen waarin structuur wordt gezocht in grote datasets en waarin classificatie en herkenning een belangrijke rol spelen. Het is verre van eenvoudig om erachter te komen hoe een netwerk, eenmaal getraind op data, feitelijk aan zijn output komt. Om dit in meer detail te onderzoeken is onlangs het project InDeep van start gegaan.

Door:
Louis ten Bosch en
Iris Hendrickx,
CLS/CLST,
Radboud
Universiteit,
Nijmegen

Deep Learning is een vorm van machinaal leren waarbij netwerken worden gebruikt die bestaan uit een groot aantal lagen. Een laag bestaat uit rekeneenheden die informatie ontvangen van units van de vorige laag, daar een berekening op loslaten en daarna hun resultaat doorgeven aan de volgende laag.

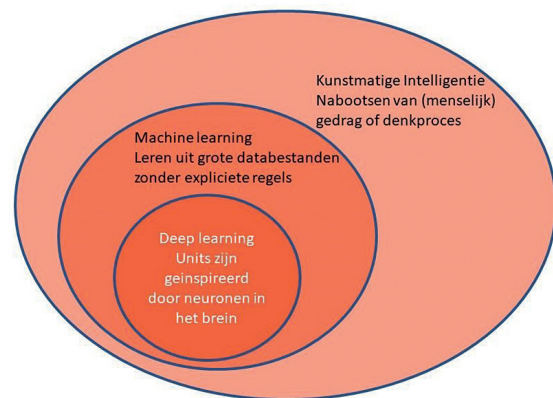


Fig. 1: Deep Learning, Machine leren en AI.

De Deep Learning techniek is de laatste jaren zeer populair voor tal van toepassingen, zoals gezichtsherkenning, vingerafdrukherkenning en spraakherkenning. Deep Learning heeft toepassingen mogelijk gemaakt die tot een tiental jaren geleden, op basis van de klassieke methoden, praktisch onmogelijk waren. Inmiddels wordt Deep Learning op steeds meer terreinen een gangbare techniek zoals in de analyse van teksten, bij automatisch vertalen, bij het toekennen van betekenissen aan zinnen, bij automatische classificatie van muziek en bij automatische spraakherkenning.

Maar waarom zijn Deep Learning algoritmen zo krachtig? Een van de eigenschappen van een Deep Learning aanpak is dat het herkenningsprobleem wordt gesplitst in simpeler deelproblemen. De informatieoverdracht van de input naar de output van een netwerk wordt gesplitst in meerdere stappen: van de eerste laag naar de

volgende laag, van deze laag naar de daaropvolgende laag, en zo verder, totdat de informatie bij de laatste laag, de outputlaag, arriveert. Als elke unit in een laag een niet-lineaire operatie uitvoert, dan is een netwerk in staat om een complexe input-output mapping te beschrijven in termen van meerdere niet-lineaire stapjes die stuk voor stuk simpeler zijn dan het originele input-output probleem. De architectuur van zo'n netwerk lijkt op die van het menselijk brein; de uitgebreide verbondenheid van vele eenvoudige niet-lineaire units ('interconnectiviteit') is een van de bepalende factoren van de kracht van een 'diep netwerk'.

Er zit bij Deep Learning wel een adder onder het gras. Deep Networks mogen dan krachtige instrumenten zijn om beslissingen te simuleren (bijvoorbeeld een medische diagnose te stellen op basis van een aantal symptomen) maar het is verre van eenvoudig om erachter te komen hoe een netwerk, eenmaal getraind op data, feitelijk

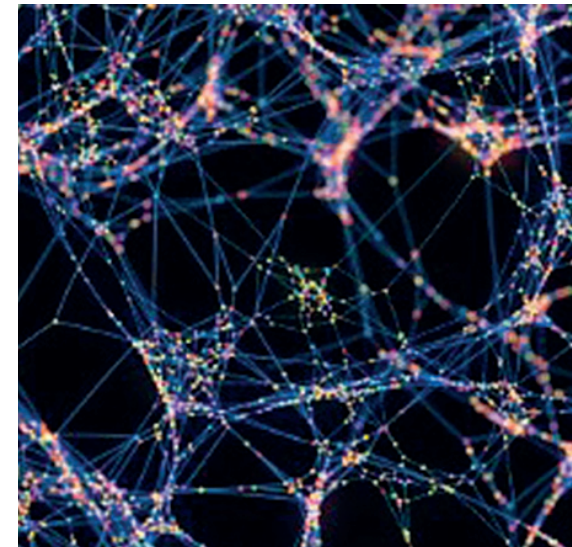


Fig. 2: In een netwerk wordt informatie van eenheid naar eenheid verspreid ('propagation'). Vele simpele niet-lineaire operaties maken een complexe operatie.

aan zijn output komt. Het netwerk is in dat opzicht vaak een ondoorzichtige black box.

Met de verbreding van de toepassingen van Deep Learning werd dit probleem steeds acuter en zichtbaarder. Als gevolg daarvan is er de afgelopen tijd internationaal meer aandacht gekomen voor de vraag hoe wij iets kunnen leren van een netwerk. Hoe komt het netwerk aan zijn beslissingen? Hoe is de informatiestroom van inputlaag naar outputlaag? Kunnen wij begrijpen waarom een netwerk een bepaalde beslissing neemt?

InDeep

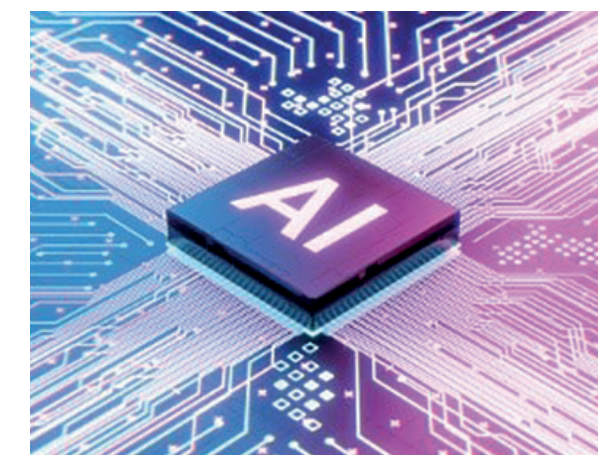
Om dit in meer detail te onderzoeken is onlangs het project InDeep¹ van start gegaan. Dit project, gefinancierd via NWO-NWA (Nationale WetenschapsAgenda), heeft als doel om meer inzicht te krijgen in hoe een netwerk werkt en hoe de 'informatie propagatie' in zo'n netwerk eruit ziet. In InDeep worden daarvoor een aantal methoden gebruikt. Om een blackbox stukje bij beetje open te breken bestaan er verschillende methoden die bekend zijn uit de wiskunde en uit computationele simulaties.

Een voorbeeld van zo'n aanpak is het langzaam veranderen van de input en kijken wanneer een beslissing van het netwerk kantelt. Op zo'n manier kan de inputruimte worden gescand op essentiële patronen. Een andere manier is na te gaan voor welke input een individuele unit op een bepaalde tussenlaag gevoelig is. Of je kunt proberen de informatie te volgen van beginlaag tot eindlaag en daar patronen in te zien.

In InDeep richten we ons op netwerken die voorspellingen doen op basis van door mensen ge-

produceerde signalen als input, namelijk tekst, spraak en muziek. Binnen het project werken academische en industriële partners samen aan zowel het ontdekken van theoretische inzichten in de werking en interpretaties van netwerk algoritmes als het toepasbaar maken van deze technologieën in real-life praktische toepassingen. Een van de doelen is het creëren van een toolbox met computationele interpretatietechnieken die toegespitst zijn op een scala van verschillende taken op het gebied van spraak, taal en muziek-analyse. In sommige taken is het de bedoeling dat deze technieken aan de gebruikers nuttige feedback geven (bijvoorbeeld bij mining van juridische of medische teksten), terwijl bij andere taken het erom gaat de gebruiker meer kennis te geven over het domein, zoals bij automatisch vertalen of automatische aanbevelings-systemen (zoals nu al bestaan op veel websites). De toolbox moet de gebruikers ook de mogelijkheid geven om zelf invloed uit te oefenen op het leerproces van de Deep Learning netwerken. Voorbeelden hiervan zijn de aanpassing van een bestaand systeem voor automatische spraakherkenning naar bepaalde dialecten en accenten, en machine-ondersteund vertalen van teksten.

InDeep is een samenwerkingsverband van onderzoekers van het ILIC, Universiteit van Amsterdam, Tilburg Universiteit, de Universiteit van Groningen, de Radboud Universiteit en de Vrije Universiteit, samen met negen commerciële partners (die in-kind bijdragen). De looptijd van InDeep is 5 jaar, van medio 2021 tot medio 2026. In InDeep komen een groot aantal verschillende vakgebieden en expertises samen, waaronder semantiek, syntactische parsing, machinegebaseerd vertalen, computationele fonologie, taalverwerking, en automatische spraakherkenning. •



Meer informatie over InDeep is te vinden op:

<https://interpretingdl.github.io/projects>

Foto's van <https://unsplash.com/>

¹ <https://interpretingdl.github.io/>

DE NEDERLANDSTALIGE SPRAAK COALITIE

Dit is een samenvatting van het ontstaan van het initiatief om Nederlandstalige* taal- en spraaktechnologie te ontwikkelen als een gezonde publiek-private samenwerking: open, vertrouwd, inclusief, en soeverein.



teit) verliezen. Er ligt te veel macht bij partijen die enkel vanuit commerciële belangen opereren. Zij bepalen wat er aan diensten ter beschikking komt en hoe. Voor Nederlandse en Vlaamse bedrijven wordt het dan lastiger om hun data te delen via de kanalen van deze partijen.

Het gebrek aan overzicht en kunde van en regelgeving voor de toepassing van spraaktechnologie

De complexiteit van spraak-dataverwerking en privacy staat niet in verhouding tot de kennis bij de overheid, het bedrijfsleven en gebruikers, hetgeen regelmatig leidt tot gebrek aan bewaking en bescherming. Er is een te groot kennisgat tussen de wetenschap, publieke organisaties, het bedrijfsleven en een gebrek aan voldoende experts.

De positieve impact van spraaktechnologie wordt in de weg gestaan door deze problemen. De positieve impact in het Nederlandstalige taalgebied is er een van omzetgroei, banengroei en het versterken van de internationale concurrentiepositie door de technologieleveranciers en de dienstverleners. Zij maken het mogelijk dat organisaties (commercieel en publiek) effectiever en efficiënter zijn door middel van spraaktechnologie. Daarbij komt de groei aan inclusie voor bijvoorbeeld laaggeletterden, blinden en slechtzienden en oudere volwassenen die met de groei van spraaktoepassingen meer deelnemen aan de maatschappij.

De Nederlandstalige Spraak Coalitie en haar missie

Uiteindelijk werd na validatie van de problemen en erkenning van de potentie van het

spraakkanaal ook het nut van de coalitie bekrachtigd. Vertegenwoordigers van bedrijven uit een breed scala aan sectoren zoals KPN, Sevagram, DPG, Achmea en VGZ; aanbieders en dienstverleners zoals Deloitte, y.Digital en Telecats; vertegenwoordigers van universiteiten uit Nederland en Vlaanderen, Stichting Open Spraaktechnologie en vertegenwoordigers van instellingen zoals Waag en Visio stelden als missie het volgende:

Spraaktechnologie in het Nederlandse taalgebied te ontwikkelen als een gezonde publiek-private samenwerking; open, vertrouwd, inclusief en soeverein.

Vier werkstromen om de missie te realiseren

De Nederlandstalige Spraak Coalitie zal haar missie volbrengen, structureel en toekomstvast, door samen met haar partners en leden, activiteiten te ontwikkelen en uit te voeren. De volgende vier werkstromen geven richting aan deze activiteiten:

1. Faciliteren, ontwikkelen en onderhouden van open Nederlandstalige akoestische en taalmodellen;
2. Belangenbehartiging richting politiek en grote technologiepartijen;
3. Versterken en ontwikkelen van spraakinfrastructuur;
4. Bijeenbrengen van vraag en aanbod spraakdiensten.

Status en uitnodiging

De coalitie is nu (oktober 2021) een stichting in oprichting, met zes - betaalde - lidmaatschapstoezeggingen. Samen vertegenwoordigen deze, zoals beoogd, het hele spraaklandschap met een nadruk op toepassing door commerciële en publieke organisaties.

De moderne, open werkprincipes van de stichting zijn reeds operationeel. Onlangs is voor het eerste project een fondsaanvraag gerealiseerd en voor het einde van het jaar staan er meer op de planning. Met het aan Nederlandse AI Coalitie werkgroepen gelieerde Nederlandse AI voor het Nederlands wordt nauw samengewerkt om het pad naar subsidies te vergemakkelijken en de band met overheidsorganisaties te versterken.

Spraaktechnologie is een fundamentele technologie voor iedereen die alleen door samen te werken tot haar volste recht gaat komen. •

Doe mee: Resoneert de missie bij jou en je organisatie? Of zie jij problemen om op te pakken dan wel heb je ideeën voor oplossingen in het geschetste veld? Ik hoor het graag!

Maarten Lens-FitzGerald
Nederlandstalige Spraak Coalitie

www.spraakcoalitie.nl
maarten@sprakcoalitie.nl

5 Boeken over conversational AI die je gelezen moet hebben

Conversations with things - Rebecca Evanhoe en Diane Deibel

Hard op weg om hét standaardwerk voor conversation designers te worden. Een praktische, vlot geschreven en grondige gids over alle aspecten van het conversation-designproces.

Talk - Elizabeth Stokoe

Aan de hand van getranscribeerde gesprekken neemt conversation analist Elizabeth Stokoe ons mee in de wonderlijk complexe wereld van hoe wij mensen met elkaar praten.

Invisible Women - Caroline Criado Perez

Een ontluisterende eye-opener over hoe vrouwen stelselmatig niet worden meegenomen in data, onderzoek en productontwerp, en welke problemen dat oplevert in de praktijk.

It's better to be a good machine than a bad person - Bruce Ballentine

Doen we er wel goed aan om onze conversational interfaces een persoonlijkheid en menselijke trekjes te geven? Bruce Ballentine is daar heel duidelijk in: nee.

Voice user interface design - Michael H. Cohen, James Giangola, Jennifer Balogh

Een klassieker uit 2004, oorspronkelijk geschreven voor IVR-ontwerp, maar met ontwerpadviezen die nog steeds staan als een huis.

Meer lezen? Bezoek Maaikes Bibliotheek op www.convocat.nl/leesvoer voor 50 boeken over conversational AI, taalkunde, inclusive design, NLP/NLU en content management.

* Met Nederlandstalig worden alle vormen, variaties en dialecten van het Nederlands bedoeld. Ook wat buiten de landsgrenzen wordt gesproken zoals het Vlaams.

CONVERSATIONAL TTS

Tekst-naar-Spraaksynthese (text-to-speech of TTS) is de omzetting van tekst naar spraak door software. In het verleden werd TTS voornamelijk ingezet voor mensen met een leerprobleem of leesbeperking zoals dyslexie, voor mensen met een verminderd gezichtsvermogen en voor mensen die een taal wilden leren. Maar TTS is ook handig voor alle mensen die gewoon liever luisteren dan lezen. Met de opkomst van AI en krachtige computers is er tegenwoordig veel meer mogelijk; persoonlijke assistenten zoals Siri en Alexa, robots en chatbots worden steeds meer gebruikt. Om de spraak natuurlijk en geloofwaardig te laten klinken in verschillende applicaties, is het nodig om te kijken naar hoe een TTS-stem wordt gemaakt en hoe dat proces verder verbeterd kan worden.

Door:
Esther
Klabbers,
ReadSpeaker

AI in TTS

De huidige generatie TTS-stemmen wordt gemaakt met behulp van neurale netwerken die op basis van opnamen (gemaakt met een professionele stemacteur) leren om tekst (en de taalkundige inhoud daarvan) om te zetten naar spraak. In figuur 1 zie je de architectuur van zo'n systeem dat bestaat uit drie modules. Het eerste is een taalkundige module die tekst omzet naar een representatie van de uitspraak van de tekst, met symbolen voor de klanken (fonemen), klemtonen, woordgrenzen en leestekens. De tweede module is een akoestische module. Daarin wordt een neurale netwerkdetraind om deze symbolen te vertalen naar een representatie van de akoestische signalen. De derde module is de neurale vocoder die getraind wordt om de akoestische representaties weer te vertalen naar audiosignalen. Door de netwerken met veel voorbeelden te trainen leren ze te generaliseren en kunnen ze nieuwe tekst omzetten naar spraak van een hoogwaardige kwaliteit.

Ontwikkeling van een neurale TTS-stem

Om een TTS-stem te maken heb je als bronmateriaal opnamen nodig van een professionele stemacteur. Die opnamen worden geannoteerd met de taalkundige informatie, en de akoestische eigenschappen worden afgeleid van het audiosignaal. Het is gebruikelijk om meerdere uren spraak op te nemen en om de neurale netwerken net zo lang te trainen tot er een stabiel model is bereikt waarmee natuurlijk klinkende spraak gegenereerd kan worden.

In het verleden werden TTS-stemmen vooral gemaakt met een techniek die unit selection synthesis (USS) heet. Daarvoor was een database van tien tot dertig uur opgenomen audiomateriaal nodig die ook weer geannoteerd moest worden. Om tekst naar spraak om te zetten wordt bij USS in deze database gezocht naar korte audiofragmenten die aan elkaar geplakt worden. Om ervoor te zorgen dat dit natuurlijk klonk en er weinig hoorbare klikken en toonhoogtesprongen

te horen waren, moest de stemacteur tijdens de opnamen een neutrale voorleesstijl gebruiken. Dit maakte dat de sprekers niet met veel expressie konden spreken en niet konden worden ingezet in situaties waar andere spreekstijlen gewenst waren. Daartoe zou de database uitgebreid moeten worden met nog veel meer materiaal, wat kostbaar en tijdsintensief is. Bovendien werd door het opnemen van expressievere spraak de kans op grotere toonhoogtesprongen ook groter. USS was kortom niet ideaal voor de steeds hogere eisen die in de markt aan TTS-stemmen worden gesteld.

Met de nu gangbare neurale synthese is er meer flexibiliteit en kan men met minder data modellen trainen. Maar hoe goed de neurale modellen ook zijn, voor het trainen van een model dat natuurlijk klinkende spraak in verschillende spreekstijlen en/of met uiteenlopende emoties kan genereren, is het nodig dat er voldoende kwalitatief hoogwaardige audio wordt opgenomen in de gewenste spreekstijlen en emoties, en dat op basis van een opnamescript dat daarbij past.

Conversational AI

De laatste jaren hebben we een enorme opmars gezien van conversational AI, waarbij mens en machine met elkaar 'praten'. De TTS-stemmen die in het verleden ontwikkeld werden met USS, en die er vooral op gericht waren om langere teksten voor te lezen, zijn niet optimaal voor het gebruik in conversational AI. Daarom onderzoeken we bij ReadSpeaker in de projecten BLISS en COBRA hoe we TTS-stemmen kunnen maken met een spreekstijl geschikt voor dialogen met de bijbehorende variatie in prosodie (zinsmelodie) en emoties.

BLISS

Binnen het kader van het DATA2PERSON project van de NWO werken we aan het BLISS project (Behaviour-based Language-Interactive Speaking Systems), samen met onderzoekers van het

ReadSpeaker

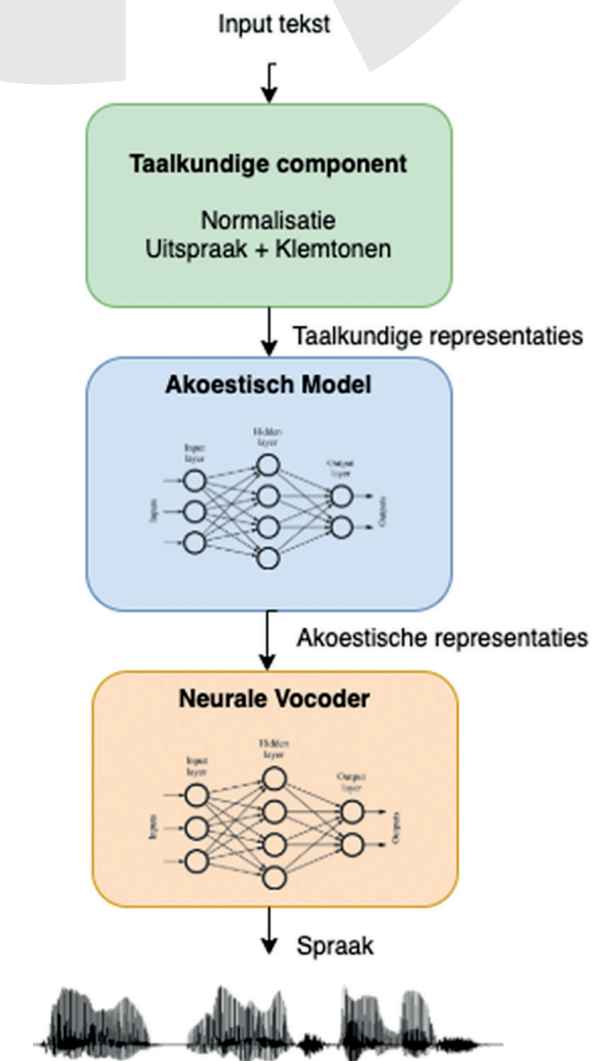
Centre for Speech and Language Technology (CLST) van de Radboud Universiteit, onderzoekers van Human Media Interaction (HMI) van de Universiteit Twente en het bedrijf Game Solutions Lab (GSL) (zie DIXIT December 2020, pp. 18-19, <https://bliss.ruhosting.nl>). Het doel van dit onderzoeksproject is het ontwikkelen van een spraakgestuurd interactief, Nederlandssprekend systeem dat oudere mensen helpt zelf hun gevoel van welzijn en geluk te vergroten. ReadSpeaker ontwikkelt daarvoor een neurale Nederlandse TTS-stem. Er worden momenteel experimenten gedaan met verschillende samenstellingen van het akoestische model om meer controle over de prosodie van de spraak te krijgen.

COBRA

In het kader van het EU COBRA project (Conversational BRAins, <https://www.cobra-network.eu/>) begeleiden we bij ReadSpeaker een PhD-student. Samen met een conversational designer onderzoekt zij hoe een optimaal opnamescript samen te stellen voor conversationale spraak. Er zijn verschillende vragen die we hiermee willen beantwoorden, zoals: klinkt de gegenereerde spraak beter (meer natuurlijk of plausibeler) wanneer met de stemacteur zinnen zijn opgenomen die in een conversationeel systeem gebruikt zouden worden? Klinkt de gegenereerde spraak beter als opnamen gemaakt worden in een conversationale simulatie met twee sprekers? Welke scenario's kom je tegen en zou je willen terugzien in het opnamescript?

Dit is nog maar het begin van dit onderzoekstraject, dat als doel heeft om in de toekomst expressievere TTS-stemmen te kunnen genereren die passen bij de applicaties waar ze voor gebruikt worden. •

Zie <https://www.readspeaker.ai/> voor meer informatie over onze neurale TTS-stemmen en demo's.



Figuur 1: Neurale TTS-architectuur die bestaat uit:
1. De taalkundige component die de input tekst vertaalt naar een taalkundige representatie met fonemen, klemtonen en leestekens;
2. Het akoestisch model dat een omzetting leert van deze representatie naar de akoestische eigenschappen;
3. De vocoder die een omzetting leert van de akoestische eigenschappen naar spraak.

EEN VOICE OF CHATBOT VOOR VLAANDEREN EN NEDERLAND? HARTSTIKKE NEIG!

Door:
Tess Tettelin,
Chatlayer by
Sinch

Nederlands en Vlaams, vaak hetzelfde maar soms ook helemaal anders. In binnen- en buitenland wil er al wel eens verwarring ontstaan over de Nederlandse taal die gesproken wordt in Vlaanderen en Nederland. Zo worden het Vlaams en het Nederlands nog te vaak beschouwd als twee aparte talen, of denkt men dat het Vlaams een dialect is binnen het Nederlands. Goe zot!

Ook voor ons Nederlandstaligen durft het al wel eens mislopen. Ik, Vlaamse, vertelde een Nederlandse vriendin over het nieuwe kledje dat ik onlangs gekocht had. Ik denk dat we vijf minuten lang aan het praten waren voordat ik doorhad dat zij dacht dat ik een nieuw tapijtje gekocht had. Voor alle duidelijkheid: ik had het over een nieuwe jurk. Dus ook al is onze grammatica nagenoeg hetzelfde, niet alle Nederlandstalige woorden betekenen hetzelfde in Vlaanderen, of worden daar überhaupt gebruikt. Verwarring zit in een klein hoekje. Dat je elkaar verstaat, wil nog niet zeggen dat je elkaar ook begrijpt.

Hetzelfde geldt voor een chatbot. Als wij mensen al worstelen met natuurlijke taal, iets wat we van kleins af aan leren, hoe moeilijk is het dan voor bots om nuances te kunnen begrijpen?

Ken jij bijvoorbeeld het verschil tussen lopen en stappen? Wat gebeurt er met je lichaam wanneer je je regels hebt? En wat verwacht je te krijgen wanneer je een chatbot om patat vraagt?

In dit artikel som ik de **vijf geboden** op voor als je een Nederlandstalige bot gaat bouwen die in zowel Vlaanderen als Nederland actief gaat zijn.

Gaan al deze tips nuttig zijn voor jou? Waarschijnlijk niet. Maar wel plezant!

1. Gij zult lokaliseren eer ge gaat publiceren

Er zijn taaldeskundigen die beweren dat 'hét Nederlands' niet bestaat en ik sluit me daarbij aan. Er is namelijk geen afgesloten groep woorden,

grammaticale regels en uitspraakrichtlijnen voor het hele taalgebied. Wij Vlamingen horen al van bij het eerste woord of er een Nederlander of Surinamer aan het praten is, en dat geldt ook omgekeerd. Ik hoor zelfs of iemand uit het noorden of het zuiden van ons land komt. Er is dus zeker variatie in de Nederlandse taal en het is net die variatie die een chatbot-ervaring pas echt persoonlijk kan maken.

Wanneer je een voicebot bouwt die zowel Vlaamse als Nederlandse gebruikers gaat helpen, is het belangrijk om een stem te selecteren die voor beide taalgebieden past. Mensen vormen binnen enkele seconden een beeld van een bot. Klinkt die Hollands? Dat werkt niet in Vlaanderen, net zoals een Vlaamse bot niet werkt in Nederland. Beide taalgebieden hebben hun eigen spreektaal en zijn daar ook trots op. Je bot vervlaamsen of vernederlandsen werkt dan ook in jouw voordeel. Op die manier spreek je pas écht je doelgroep aan en creëer je een persoonlijke gebruikerservaring.

De stem voor een voice assistent moet aansluiten bij de mensen die de bot gebruiken. Bij Chatlayer.ai werken we samen met ReadSpeaker om custom text-to-speech voices (TTS) te kunnen aanbieden. Zo'n text-to-speech voice zet geschreven tekst om naar gesproken tekst en wordt dus de stem van je voice assistant. Op het Chatlayer platform zijn er zowel Vlaamse als Nederlandse stemmen beschikbaar, waardoor je voice-ervaring pas echt lokaal wordt.

2. Gij zult uw persoonlijkheid laten zien

Normen zijn (ongeschreven) regeltjes die onze sociale interacties in goede banen leiden. Zo gedraag je je waarschijnlijk anders naar een collega toe, dan naar een goede vriendin. Bij mij vallen die twee gelukkig samen, dus ik kan deze regels gewoon aan mijn botten* lappen. Hoezeel!

Ook taal kent normen die bepalen hoe we communiceren met anderen. Wanneer ik iemand aanspreek op straat om de weg te vragen, zeg ik "Hallo, kunt u mij helpen?" maar wanneer ik hulp nodig heb van iemand die ik al jaren ken, zeg ik eerder iets als "Hey hallookes kom mij is redden!" Zoals je ongetwijfeld zelf kan aanvoelen, leidt een overtreding van zulke normen in gesprekken tot misverstanden omdat de gesprekspartners elkaar niet begrijpen, of nog erger, niet vertrouwen. Het voelt gewoon niet juist.

Bij het bouwen van een chatbot is het dan ook belangrijk om deze van een persona te voorzien. Zo bied je niet enkel een functionele, maar ook emotionele ervaring aan de gebruiker. Geen idee wat een persona is? Een bot-persona bepaalt hoe je bot praat, eruitziet en reageert in bepaalde situaties. Je kunt het een beetje vergelijken met een personage in een boek of film. Hoe zegt jouw bot bijvoorbeeld hallo? Hoe biedt het zijn excuses aan wanneer het een fout maakt? Hoe reageert het op spontane liefdesverklaringen? Een persona zorgt ervoor dat je gebruiker zich meer aangesproken voelt.

Denk hierbij ook aan je gebruiker. Wat verwacht deze van je chatbot? Een chatbot voor een gevestigde bank communiceert wellicht heel anders dan die van een trendy kledingmerk. Wendy's, de Amerikaanse hamburgerketen, durft al wel eens brutaal uit de hoek te komen op haar socialemediakanalen. Tot groot plezier van hun klanten!

Mochten zij een chatbot inschakelen, verwacht ik dat deze net zo grappig en gedurfd uit de hoek zal komen. Een chatbot kan je merk enorm versterken, maak er dus gebruik van!

* 'Botten' is Vlaams voor 'laarsjes' wat dan weer een soort schoeisel betekent in het Algemeen Nederlands.

3. Gij zult blijven evolueren

Op school probeert men de Vlaming nog steeds formeel Algemeen Nederlands (AN) aan te leren als enige correcte vorm van het Nederlands. Voor lange tijd was het tv-nieuws-Nederlands van Martine Tanghe dan ook de norm, alleen bleek dat niet door iedereen gebruikt te worden. En ondanks dat heel Vlaanderen massaal aan het beeld gekluisterd zat tijdens Martine's laatste nieuwsaflevering, spreken vele (jonge) Vlamingen haar Nederlands niet omdat ze zich in een heel andere context bevinden dan ons nieuwsdool.

Wanneer je een chatbot bouwt, is het belangrijk om voldoende gevarieerde woorden en expressies toe te voegen aan je NLP-model. Dit zorgt er namelijk voor dat je chatbot beter kan begrijpen wat de gebruiker aan het zeggen (of eerder typen) is. Wanneer ik een bot bouw voor zowel

een ouder als een jonger doel-publiek, voeg ik naast 'wat is de prijs?' ook 'wa kost da?' toe aan het NLP-model. Ik weet het, van die laatste zin krijg ook ik de wubbes* maar zo praat de jeugd nu eenmaal.

* Een Antwerpenaar krijgt het niet op zijn heupen, zij krijgen er "de wubbes" van.

4. Gij zult voor voldoende variatie zorgen

Goed, we spreken nagenoeg dezelfde taal en dit gaat meestal zonder problemen, maar wat wel voor misverstanden kan zorgen zijn de woorden die zowel in Nederland als België gangbaar zijn, maar die verschillende betekenissen hebben per taalgebied. Naast klassiekers als poep, bank en tas, zijn er nog tal van andere voorbeelden. Als een Vlaming bijvoorbeeld een telefoontje verwacht in de voormiddag, dan bel je hem of haar tussen 9 en 12 uur. In Nederland rinkelt de telefoon dan weer best pas tussen 12 en 14 uur.

Weten wat de gebruiker zegt én wat hij of zij daarmee bedoelt is dan ook een van de fundamenteën van een goede chatbot. Het geheim?

Dat schuilt in de Natural Language Understanding (NLU) die de bot gebruikt om onze natuurlijke taal te verstaan. NLU valt dan weer onder het bekende Natural Language Processing (NLP), wat de AI-technologie is die taalkundige modellen opbouwt en integreert waardoor een chatbot mensen ook echt kan begrijpen.

Een verzekeringsbot die klanten helpt bij het inplannen van reparaties moet begrijpen wat een Vlaming bedoelt wanneer hij of zij het over een 'autobatterij' heeft, net als wanneer een Nederlander het over de 'accu' heeft. Anders dikke pech! Train je NLP-model dus ook op synoniemen voor bepaalde woorden.

Ik denk hierbij ook aan sentiment analysis en het correct kunnen interpreteren van trefwoorden. Als je chatbot om feedback vraagt aan een Vlaming en deze het woord 'bangelijk' gebruikt, moet je bot begrijpen dat het om een zeer positieve ervaring gaat. Wanneer een Nederlander daarentegen hetzelfde woord gebruikt, zit er iets grondig mis. •



VAN INTENT-CLASSIFICATIE NAAR WEDERZIJDIG GESPREK MET EEN CONVERSATIONAL AGENT

De afgelopen jaren is de populariteit van conversational agents enorm toegenomen. Door nieuwe ontwikkelingen op het gebied van taalmodellen worden bots steeds taalvaardiger en zijn voor steeds meer toepassingen praktisch inzetbaar. In sommige systemen wordt gesproken over Natural Language Understanding (NLU), maar is er al daadwerkelijk sprake van begrip van taal en wat is daarvoor nodig?

Door:
Serge
Cornelissen,
Founder &
Chief
Technology
Officer,
Cornelists
B.V.

De complexiteit van Natural Language Understanding voor computers

Natural Language Understanding is een zeer complex vakgebied waar computers tot nu slechts in beperkte mate toe in staat zijn. Een belangrijke reden is de beperkte kennis van de werkelijkheid van onze wereld voor computers. Het begrijpen van taal begint namelijk in wezen in het begrijpen van de fysieke wereld.

Je ziet deze grondslag terug in het menselijk proces van het leren van taal. Als we geboren worden zullen we als baby eerst de fysieke wereld gaan ontdekken. Taal helpt ons later bijvoorbeeld voorwerpen en principes te verwoorden die van toepassing zijn in die fysieke wereld.

De kennis en ervaring die we individueel opbouwen van de fysieke wereld brengt twee andere belangrijke eigenschappen met zich mee. De eerste is een stukje gezond verstand dat maakt dat we inzicht hebben in wat waarheid is in onze wereld. Mensen kunnen bijvoorbeeld lopen, lampen kunnen dat niet. De tweede is dat we met relatief weinig woorden met elkaar kunnen communiceren. Een zin als "Ik kreeg een bal hard tegen mijn neus", zal de luisteraar doen begrijpen dat dit pijn kan doen, zonder dat dit letterlijk beschreven staat. De spreker gaat ervan uit dat deze kennis reeds aanwezig is bij de ontvanger. De ontvanger zal op basis van zijn eigen 'model' van kennis en ervaring de informatie extrapoleren naar de denkbeeldige situatie en daar (hopelijk) begripvol op reageren.

Natural Language Understanding gaat verder dan analyse van tekst alleen

Computers kunnen onze wereld niet zelf ontdekken en zijn afhankelijk van teksten en databases die door mensen zijn opgesteld. Ondanks dat er veel van onze wereld in beschreven staat is dit lang niet genoeg om daadwerkelijk de wereld te begrijpen. Meer kennis van de wereld zal sa-

mengebracht moeten worden en dat zal verder moeten gaan dan analyse van tekst alleen. Op Google I/O 2021 presenteerde Google zijn eerste stappen daarin met het Multitask Unified Model (MUM), een model dat informatie uit meerdere talen en media zoals tekst en beeld combineert.

De meeste chatbots werken met intents

Kijkend naar de huidige status van chatbots dan zien we dat deze vooral werken met intent-classificatie. Een intent is de bedoeling van een uiting en kan worden gezien als een groep voorbeeldzinnen die eenzelfde actie/reactie van een gebruiker vertegenwoordigen. Denk bijvoorbeeld aan zinnen als "Ik wil mijn adres wijzigen" en "Hoe verander ik mijn adres?". Deze hebben min of meer dezelfde intentie. De voorbeeldzinnen worden door de ontwerpers van de agent gekoppeld aan een antwoord. Dankzij machine-learning technieken kan een chatbot ook een iets anders gestelde vraag als "Kan ik mijn adres aanpassen?" herkennen en classificeren.

Met intents alleen leert men een chatbot niet spreken

Zoals intent-classificatie al in het woord aangeeft leert een computer zinnen te classificeren, waarmee (gelijkende) vragen aan antwoorden kunnen worden gekoppeld. Dit classificeren wordt in de chatbotwereld vaak benoemd als Natural Language Understanding, maar toch zal dit niet altijd gaan volgens de fundamentele die van toepassing zijn in taalbegrip. We geven immers een computer weliswaar de gewenste input en output, maar leggen niet uit welke grondslag van toepassing was om van vraag tot antwoord te komen. Zelfs voor mensen is dit een lastig leerproces en het laat ruimte voor eigen interpretatie. Machine-learning bijvoorbeeld, probeert een wiskundig verband tussen vraag en antwoord te vinden, dat anders kan zijn dan het door mensen bedoelde verband. Het gevolg is dat conversational agents op basis van intent-classificatie,



De werking van de wereld staat niet alleen in tekst beschreven.

soms de kleinste aanpassingen aan zinnen niet voorzien. Kleine aanpassingen die grote gevolgen kunnen hebben in de betekenis. Denk bijvoorbeeld aan het verschil tussen 'wel' en 'niet'.

Toevoegen van fundamentele principes van wederzijdse communicatie

Een belangrijk aspect in de ontwikkeling van NLU is dat er daadwerkelijk met chatbots een dialoog kan worden gevoerd die verder gaat dan intents en vooraf gedefinieerde gesprekspaden. Het toevoegen van fundamentele principes van een wederzijds gesprek is noodzakelijk om een conversational agent communicatieve vaardigheden te geven die niet door intent-classificatie verloren gaan. Zo passen wij hiervoor een zelf ontwikkeld voorwerp-actie dialoogmodel toe in onze chatbots.

In het voorwerp-actie dialoogmodel staan het voorwerp en de actie die hiermee moet gebeuren centraal. Het idee achter het model is dat we taal beschouwen als een communicatiemiddel om voorwerpen te manipuleren, te beschrijven of over te dragen. Bijvoorbeeld in een zin als "Ik wil mijn adres wijzigen" is *adres* het voorwerp wat de actie *wijzigen* moet ondergaan.

Voorwerpen en acties spelen een belangrijke rol in opdrachtgestuurde communicatie. Stellen we een vraag aan een zoekmachine dan passen we vaak voorwerpen en acties toe, denk aan

de zoekopdracht 'adres wijzigen'. Het verschil met een zoekmachine is dat een conversational agent een gesprekspartner is. In een gesprek worden voorwerpen en acties door zinnen omgeven door kernfactoren die van belang zijn in wederzijdse communicatie. Stelt men een vraag of is het een opmerking of stelling? In welke richting vindt communicatie plaats? Wil iemand het voorwerp wel of niet? Wie bezit het voorwerp? Wordt er verwezen naar een voorwerp uit een vorige zin? Door dergelijke factoren mee te nemen in het model, kan een geautomatiseerde vorm van wederzijdse communicatie plaatsvinden. In tegenstelling tot intents leren we een agent geen zinnen te classificeren voor bepaalde taken, maar met voorwerpen en acties om te gaan in een dialoog.

Het voorwerp-actie dialoogmodel wordt op dit moment al succesvol toegepast in een aantal van onze concepten en prototypes. Zo ontwikkelden we Willem de Robot, een prototype om relatief vrije gesprekken te voeren met kinderen over dieren en de ruimte.

Door in chatbots fundamentele principes van een dialoog te implementeren, ontwikkelen we een systeem dat meer communicatielogica gebruikt dan enkel intents. Hierdoor kunnen bots beter omgaan met bijvoorbeeld verschillen tussen 'jij' en 'ik' of 'wel' en 'niet'. En dat scheelt een hoop miscommunicatie. •

WAT JE ALS TAALKUNDIGE KUNT BIJDAGEN AAN EEN CONVERSATIONAL INTERFACE

Als taalkundige ben je soms geneigd te denken dat wat jij kunt niet zo bijzonder is, helemaal als je ook nog eens geen computerlinguïst bent, niet kunt programmeren en niet thuis bent in statistiek. Een taal spreken doet tenslotte iedereen. Tot je in de praktijk ontdekt dat je wel degelijk over specifieke vak-kennis en vaardigheden beschikt, kennis en vaardigheden die heel bruikbaar zijn voor conversational interfaces. Natuurlijk is er de feitelijke kennis die je verwerft tijdens je studie en werk. Daarnaast ontwikkel je een analytisch vermogen. Je bent als taalkundige gewend om te categoriseren en te denken in systemen: syntactisch, semantisch, pragmatisch, stijl, register, fonologisch.

Door:

Martha Hofman, zelfstandig taalkundige en conversational specialist

Taalhandelingen, functies en intents

Taalkundige kennis helpt bij zowel het schrijven voor een conversational interface als bij intent recognition: het herkennen van wat een gebruiker wil zodat je de software daarop kunt trainen. Je weet wat meer en minder makkelijk is om te lezen of naar te luisteren, en je bent gewend om onderscheid te maken in stijlregisters en daarmee gebruikersgroepen.

Je bent als taalkundige ook vertrouwd met de inhoudelijke basiselementen van een conversational interface: intents, utterances (uitingen) en entities (items). In de taalwetenschap wordt uitgegaan van taalhandelingen en in methoden voor het vreemde- en tweedetaalonderwijs wordt gewerkt met functies. Intents en functies zijn over het algemeen specifiekere dan taalhandelingen, maar het principe is hetzelfde. De gebruiker wil of bedoelt iets: de taalhandeling, functie of intent. Die kan hij/zij formuleren als verschillende uitingen (utterances), en in elke uiting kunnen verschillende items (entities) worden ingevuld.

Zoals gezegd zijn functies en intents vaak een meer specifieke vorm van een taalhandeling: een vraag, bewering, mededeling, verzoek, gericht op een bepaalde behoefte, product of dienst. Functies in het taalonderwijs zijn gebaseerd op dat wat een leerder van een tweede of vreemde taal in de praktijk moet kunnen uitdrukken in het land waarvan hij/zij de taal leert. Intents in een conversational interface zijn gebaseerd op dat wat een (mogelijke) klant of gebruiker van een bedrijf of dienst kan willen of nodig kan hebben. Veel van deze specifieke taalhandelingen kunnen zowel functie als intent zijn.

Taalhandelingen in het algemeen zijn onder andere een bewering, vraag, verzoek/bevel, belofte, dreigement. Die kunnen op verschillende manieren worden geuit, zoals een vraag: Mag ik iets vragen?; Ik heb een vraag; Ik zou graag willen weten; Weet jij misschien ...?; Hoe zit het eigenlijk met ...?; Waar ik nou benieuwd naar ben ...; Vertel mij eens ...

Functies in het taalonderwijs kunnen zijn: zich voorstellen, bestellen, de weg vragen. Mogelijke uitingen van bijvoorbeeld de functie bestellen zijn: Ik wil graag ...; Mag ik een ...?; Voor mij een ...; Doet u mij maar een ...;, alstublieft; Geef mij maar; Ik neem ...; Ik denk dat ik ga voor de; Kan ik bestellen?; Mogen we ...?; Ik heb wel zin in ...

Intents in een conversational interface kunnen zijn: [...] aanvragen, [...] bestellen, [...] retourneren, informatie vragen, [...] opzeggen, [rekening] inzien. Utterances van bijvoorbeeld de intent [...] aanvragen kunnen zijn: Ik wil graag ... aanvragen; Kan ik een ... bij u openen? Kunt u mij ... sturen?; Mag ik ...?

Bij een functie of intent [...] bestellen of aanvragen kunnen in de utterances items (entities) worden ingevuld als: een kop koffie, een biertje, het vegetarische hoofdgerecht, één patat met, een spaarrekening, een creditcard, een abonnement, een account.

Intent recognition

Bij intent recognition is ook het taalkundige inzicht van belang dat uitingen van taalhandelingen, en dus ook van functies en intents, niet altijd de vorm of inhoud hebben die je zou verwachten.

Een uiting heeft soms een andere dan de verwachte syntactische vorm. Een zin met de vorm van een mededeling kan een bevel zijn, een vraag kan een mededeling zijn, of een imperatief/bevel een vraag: Ik wil dat je morgen het rapport inlevert = een mededeling die een verzoek/bevel is; Weet je dat Karim morgen komt? = een vraag die een mededeling is; Vertel eens wie er op je feest komen = een bevel/imperatief die een vraag is.

Een uiting kan ook indirect zijn. Mensen zeggen niet altijd rechtstreeks wat ze bedoelen maar doen dat vaak op een impliciete manier. Bijvoorbeeld de mededeling: Ik zou gebeld worden, bedoeld als een verzoek: Ik wil dat iemand mij belt of te woord staat.

Soms lijkt een reactie inhoudelijk niet aan te sluiten op een uiting: Wilt u een creditcard aanvragen? – antwoord: Ik ben zzp 'er, waar je als antwoord zou verwachten: Ja, ik wil er één aanvragen, of: Nee, ik wil er geen aanvragen. Hier geldt het coöperatieve principe: je gaat ervan uit dat er een coherente conversatie wordt gevoerd en dat een reactie op de een of andere manier aansluit bij de vraag. En je gaat uit van kennis van de wereld. Die houdt in dit geval in dat er een regelmatig inkomen moet zijn met een bepaalde ondergrens en dat een zzp'er daar niet altijd aan voldoet. Het antwoord betekent dus: Ik weet niet of ik aan de voorwaarden voldoe om er één aan te vragen.



Nee, ik wil geen nieuwe rekening openen! Ik wil hem alleen maar inzien!

Ook gebruiken mensen regelmatig incomplete zinnen met alleen kerninformatie. Bij het trainen van de software is het van belang te onderscheiden welke uitingen (utterances) incompleet zijn en eenduidig en welke incompleet en niet eenduidig, en de eenduidige zinnen te selecteren. Een incomplete maar eenduidige utterance is: Rekening aanvragen = Ik wil een rekening aanvragen. Niet eenduidig is: Aanvraag spaarrekening = Ik wil weten hoe het met mijn aanvraag staat; Ik wil een spaarrekening aanvragen, of misschien: Ik wil iets wijzigen in mijn aanvraag.

Dit is allemaal taalkundige kennis die een conversational interface helpt om de bedoeling van de gebruiker te herkennen en de conversatie gemakkelijker en prettiger te maken. •

Een goed gesprek: 5 podcasts voor op je luisterlijst

VUX World - Kane Simms

Kane Simms is een autoriteit op het gebied van voice en conversational AI. Op geheel eigen wijze interviewt hij conversation designers, technologie-ontwikkelaars en visionairs op het gebied van AI.

Response-Ability.Tech

Een maandelijkse podcast die bruggen slaat tussen sociale wetenschappen en technologie om een dialoog op gang te brengen over verantwoorde, inclusieve en transparante technologie.

Lingthusiam - Gretchen McCulloch & Lauren Gawne

Een maandelijkse podcasts met wonderlijke en diepgaande gesprekken over taal en taalkunde. "Joyously nerdy" –Buzzfeed.

Voicebot Podcast - Brett Kinsella

Een wekelijkse podcast van thought leader Brett Kinsella over de laatste ontwikkelingen in voice, conversational en AI. Met interviews, nieuws en vooral: heel veel marktinzichten.

Voice pioniers - DDMA

In de podcastserie Voice Pioniers gaat de DDMA Commissie Voice in gesprek met een aantal koplopers op het gebied van Voice-technologie.

Al deze podcasts zijn vindbaar op de gebruikelijke kanalen, zoals Spotify.

DE UITDAGING VAN CONVERSATIONAL INTERFACES

Een gesprek via een conversational interface tot stand brengen is niet makkelijk. Als ontwerper loop je al gauw tegen een aantal uitdagingen aan. Mensen kunnen op elk moment namelijk doen en zeggen wat ze willen. En mensen hebben allemaal hun eigen manier van praten.

Door:
Hans van
Dam,
Conversation
Design
Institute

Conversational interfaces zijn complex. De computer heeft een kunstmatig brein en de mens heeft een biologisch brein. Beide breinen hebben ieder hun eigen mogelijkheden en beperkingen. De kunst is om de triggers van beide breinen te snappen en te respecteren bij de inrichting van de assistant.

Het kunstmatige brein heeft gestructureerde data nodig om te kunnen functioneren. Het gebruikt taaltechnologie om de input van de mens te analyseren. Het herkent de intent, koppelt die aan de juiste dialoog en kan dan wat terugzeggen. Als dit niet lukt, dan valt het gesprek stil.

En dan is er de mens. De mens heeft een biologisch brein en heeft hierdoor heel andere mogelijkheden en beperkingen. Hij wil empathie, hulp en af en toe een complimentje. Als hij dit niet krijgt, dan stopt hij met communiceren en houdt het gesprek op.

Als conversational designers moeten we beide breinen dus voeden met wat ze nodig hebben

om te functioneren. We moeten ervoor zorgen dat de mens praat op een manier die door de assistant begrepen wordt. En dat de assistant op een manier praat waardoor de mens zich begrepen voelt.

In dit artikel bespreken we een aantal stappen uit de workflow die ervoor zorgen dat assistants menselijke, inclusieve en prettige gesprekspartners zijn. We staan vooral stil bij de stappen die de inclusiviteit borgen.

De workflow voor het maken van menselijke AI Assistants

Een aantal stappen is cruciaal voor het maken van een inclusieve assistant. Het werkt als volgt: mentaliseer het gesprek, heb het gesprek, maak het inclusief en train het model. Laten we ze kort bespreken.

Mentaliseer het gesprek

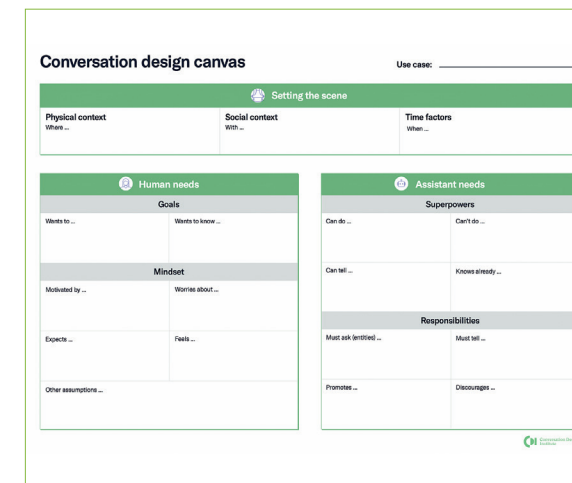
Om een goed gesprek te ontwerpen is het belangrijk je te realiseren dat je met mensen in gesprek bent. Deze mensen hebben niet alleen een

vraag of een doel, ze hebben ook motivaties, angsten en verwachtingen.

Als ontwerpers is het belangrijk ons hierin te verdiepen. Dit zorgt namelijk voor het menselijke element in de gesprekken. Door dit te begrijpen kan er empathie ontstaan.

We doen dit niet alleen voor de mens, maar we onderzoeken dit ook vanuit het perspectief van de assistant. Als de assistant deze persoon wil helpen, welke vragen moet de assistant dan stellen, wat moet hij of zij uitzoeken, en in welke mate is hij of zij in staat om te helpen?

Om hier een goed beeld van te krijgen gebruiken we onderstaand canvas. Dit canvas is een belangrijk hulpmiddel om empathie en menselijkheid te borgen in een conversational ervaring.



Heb het gesprek

Nu we de context van het gesprek begrijpen, houden we een rollenspel. Iemand speelt de mens en een ander de assistant. Het is belangrijk dat ze elkaar niet zien zodat ze alleen verbale communicatie hebben. Hiermee komen we tot de kern van het gesprek.

Nu twee mensen zich met een aantal iteraties door het gesprek worstelen, leren we wat ervoor nodig is om er een goed gesprek van te maken. Wat is de beste volgorde van de informatie, hoe geef ik vertrouwen, hoe haal je onzekerheid weg?

Door het gesprek een aantal keer te voeren begrijpt de conversation designer steeds beter hoe hij of zij de mens helpt zijn doel te bereiken. Er ontstaat empathie, dit zorgt ervoor dat de mens zich begrepen voelt en er vertrouwen in heeft dat de assistant gaat helpen.

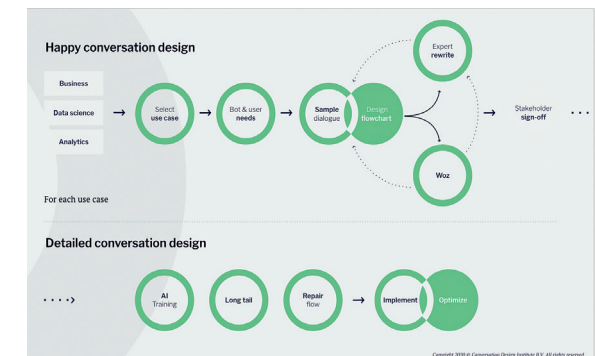
Deze oefening levert een dialoog op. Dit is een eerste versie die vervolgens gevalideerd kan worden. We hebben het gesprek nu ontworpen voor een persoon, in de volgende stap gaan we hem testen.

Maak het inclusief

Nu we een eerste versie van onze dialoog hebben, is het belangrijk om deze inclusief te maken. Waar hij ontworpen is met aandacht voor een persoon, gaan we er nu voor zorgen dat deze dialoog voor iedereen kan werken.

De manier om dit te doen? We testen de dialoog met een diverse groep mensen, het liefst een afspiegeling van de doelgroep. Zo wordt de dialoog duidelijk voor iedereen.

Een overheidsinstantie zal de dialoog met een volledige afspiegeling van de samenleving willen toetsen. Een merk dat zich op tieners richt, hoeft er natuurlijk niet voor te zorgen dat het op een manier praat die een bejaarde begrijpt – dat maakt het gesprek mogelijk minder krachtig voor de daadwerkelijke doelgroep.



Train je taalmodel

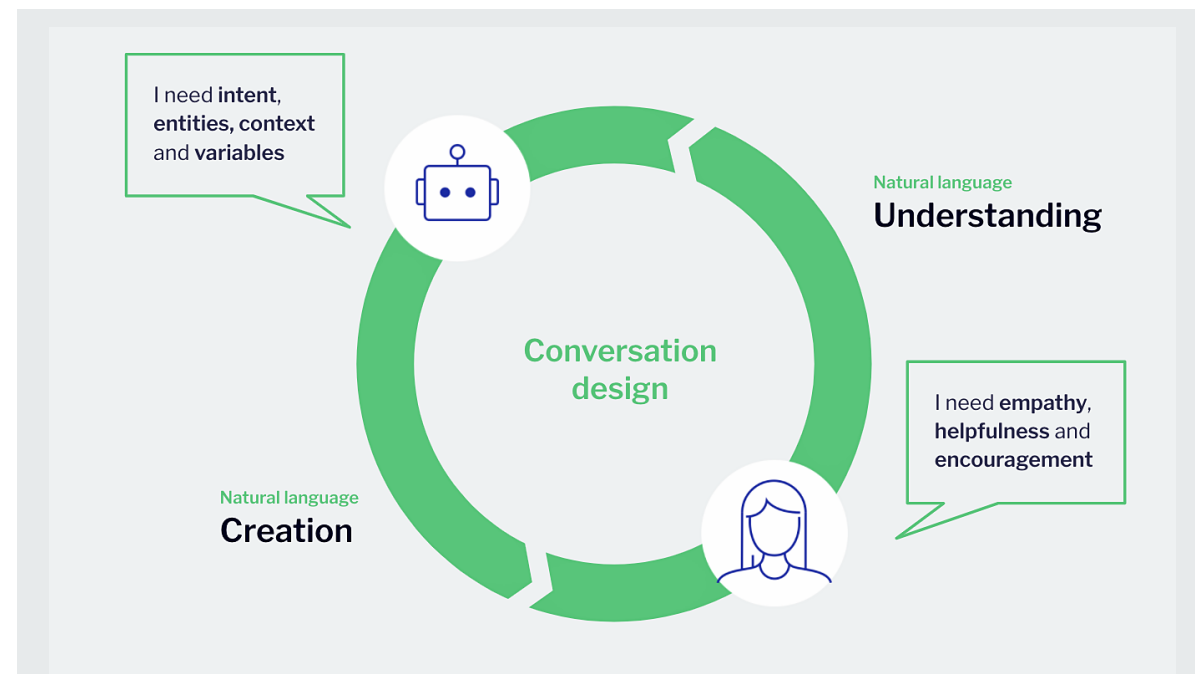
Nu gaan we de dialoog implementeren en het taalmodel trainen. We maken een intent aan en gaan trainingszinnen toevoegen. Oftewel, we gaan verschillende zinnen bedenken die dezelfde intentie hebben om mensen beter begrijpen.

Er zijn natuurlijk verschillende manieren om een kop koffie te bestellen. 'Doe maar een bakkie', 'kopje zwart goud alsjeblieft', of het normale 'mag ik een kop koffie'. Bij het toevoegen van deze trainingszinnen is het ook cruciaal om inclusief te zijn en je bewust te zijn van je eigen vooroordelen.

Als vier blanke mannen het taalmodel trainen dan snapt de assistant waarschijnlijk heel goed hoe blanke mannen praten. Met andere mensen heeft hij moeite. De trainingszinnen moeten daarom van een diverse groep mensen komen. Uiteindelijk willen we inclusief zijn op het gebied van taalvaardigheid en linguïstische voorkeur.

Inclusiviteit als speerpunt in de markt

Gelukkig is de markt zich hier steeds meer bewust van. In het Conversation Design Manifesto wordt nadrukkelijk opgeroepen om inclusief te zijn in hoe de assistant praat en in hoe hij de linguïstische voorkeuren van mensen respecteert. Dit manifest is ondertekend door organisaties als Salesforce, ServiceNow, Rasa en vele anderen.



GEPERSONALISEERDE SERVICE MET VOICEBOT SAM VAN VGZ

Door:
Sander
Hesselink,
Telecats

Om het innovatievermogen te vergroten heeft zorgverzekeraar VGZ een innovatieteam tussen IT en de business in gezet. Stefan Visser, data catalyst en innovatiearchitect bij VGZ vertelt over de werkwijze van het innovatieteam en het meest recente succesvolle innovatieproject: voicebot Sam.

Het innovatieteam van VGZ verkent toepassingsmogelijkheden van nieuwe technologie. Samen met de business zet het team experimenten op aan de hand van concrete en relevante businessvraagstukken. In eerste instantie werkte het innovatieteam van VGZ vooral met externe partijen, maar het team is inmiddels uitgebreid met een projectmanager, business-analisten en ontwikkelaars waardoor VGZ zelf de innovaties kan ontwikkelen. Afgelopen jaar is het team verder uitgebreid met datascientists naar 20 mensen. Voor de voicebot werkt VGZ samen met voicespecialisten Telecats en Readspeak.

Succes innovatieteam

Bij innovaties in klantcontact wordt vaak gestreefd naar kostenbesparing. Maar het aantoonbaar maken van toegevoegde waarde is soms lastig. Bij een niet-sluitende businesscase kan dan bijvoorbeeld de visie op de technologie de doorslag geven om toch door te gaan, aldus Stefan Visser. Daarnaast moet je de strijd aan met legacy, veroorzaakt door technologie-keuzen uit het verleden. Voor veel innovaties in klantcontact is bijvoorbeeld een goede ontsluiting van data een voorwaarde. Visser legt uit dat het innovatieteam van VGZ succesvol kan zijn omdat het IT-landschap over het algemeen goed op orde is: "Er is een uitgebreide laag middleware inclusief een goed API-platform waarmee we informatie uit backoffice-systemen kunnen ontsluiten. Daarnaast ontwikkelen we uitsluitend cloud native oplossingen, waarbij je in eerste instantie niet veel rekening hoeft te houden met de regels die gelden voor het reguliere IT-landschap."

Over naar de chatbot

De keuze voor een chatbot, nu alweer vijf jaar geleden, is mede het resultaat van twee andere overwegingen van VGZ. De ziektekostenverzekeraar wilde de bereikbaarheid buiten kantooruren vergroten én stoppen met het bewerkelijke e-mailkanaal. VGZ zocht eerst de samenwerking met een innovatiepartner. Uiteindelijk besloot VGZ om hiermee te stoppen omdat ontwikkeling, schaalbaarheid en beheer om

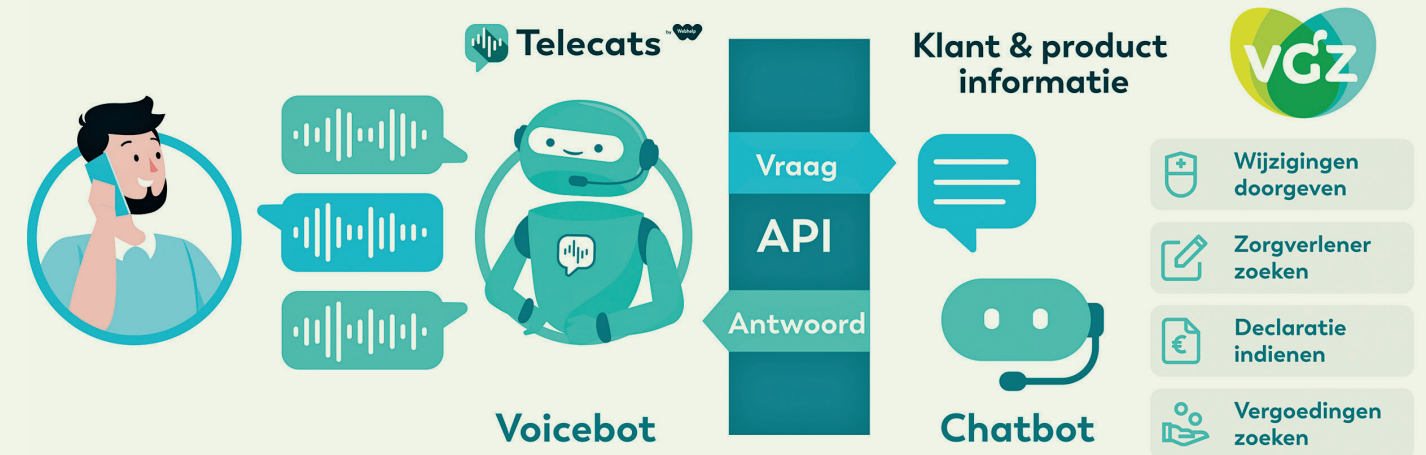
te grote investeringen in resources vroegen. Visser: "Onze ontwikkelaars hebben vervolgens zelf een volledig nieuw chatbotplatform gebouwd. Na een half jaar hebben we de chatbot in productie genomen en vervolgens overgedragen aan een nieuw team voor beheer en verdere ontwikkeling."

24x7 klanttevredenheid

Zoals gezegd: soms pakt de businesscase toch anders uit dan verwacht. De chatbot van VGZ leidde niet tot een reductie in het telefoniekanaal, aldus Visser. Toch wilde de directie wel doorzetten. Visser over het motief: "We hebben vier miljoen klanten en gemiddeld belt iedere klant eenmaal per jaar. Zo'n 60 procent van deze contacten is relatief eenvoudig van aard. De vragen die weinig tot geen waarde toevoegen, willen wij het liefst automatiseren en onderdeel laten uitmaken van 24x7 dienstverlening, vooropgesteld dat de klanttevredenheid op peil blijft. We hebben óók te maken met grote hoeveelheden ingewikkelde vragen – denk aan mensen die in meerdere ziekenhuizen zijn geweest en die vragen om een second opinion. In dit soort contacten is empathie noodzakelijk en automatisering geen oplossing. In de dialoog met de chatbot kan je rekening houden met twijfelgevallen: wanneer het lastig wordt voor de chatbot moet het gesprek door naar een medewerker. Verder moet je het toepassen van een chatbot beschouwen als een leerweg. Als klanten aan de telefoon vragen naar dingen die ook op de website zijn te vinden, kan er heel goed nog een andere vraag achter die ogenschijnlijk eenvoudige vraag zitten. Hier moet je doordacht mee omgaan."

Van chatbot naar voicebot

Inmiddels is de chatbot van VGZ wel een succes. In 2020 werden 582.000 klantgesprekken afgewikkeld, bijna driemaal zoveel als in 2019. En de chatbot krijgt een NPS (Net Promoter Score) van +46 van klanten. Het succes van de chatbot zette het innovatieteam aan het denken: waarom zetten we de bot niet in binnen het telefoniekanaal? Het idee voor de voicebot was geboren.



Telecats was al bij VGZ aan boord vanuit een initiatief van de business om een experiment uit te voeren rond het routeren van klantcontact met spraakherkenning. Die pilot viel samen met het anders organiseren van de klantenservice-teams, namelijk in zelforganiserende cirkels gericht op specifieke onderwerpen. Voor het laten werken van de voicebot was er behoefte aan transcripties van gesproken tekst én aan een spraakgenerator die tekst omzet in spraak. Hoe beantwoord je een telefonische klantvraag met kunstmatige spraak?

Voicebot Sam met spraaktechnologie van Telecats

We willen de klant zo snel mogelijk naar de juiste oplossing sturen. Met spraakrouteren wordt intelligent routeren mogelijk. Dat betekent dat je afhankelijk van klant en klantvraag kunt bepalen of je een gesprek door de voicebot, of beter door een medewerker kunt laten afwikkelen.

De voicebot met de naam Sam handelt nu zo'n duizend gesprekken per week af. Sam luistert op basis van de spraakherkenning van Telecats en spreekt op basis van (Text-To-Speech) technologie van Readspeak. Daarbij is Sam, net als

een medewerker, in staat om op een aantal onderwerpen gepersonaliseerde dienstverlening te bieden. We weten wie de klant is, want we vragen het klantnummer en we voeren een authenticatie uit door een code naar de telefoon van de klant te sturen. Als Sam, onze digitale assistent, het gesprek zelf gaat beantwoorden, dan weet hij al waar het over gaat en wie hij aan de lijn heeft.

API-platform

Visser legt uit dat de voicebot actief is op twee verschillende toepassingsgebieden: bij informatievoorziening en mutaties. "In die tweede categorie kunnen klanten nu vragen stellen en concreet zaken regelen rondom de zorgpas, betalingsregelingen en gespreid betalen van het eigen risico. Met het API-platform van VGZ kan Sam bij een vraag over een factuur in backoffice-systemen nagaan of er een betalingsachterstand is en wat de oorzaak is; maar hij kan ook meteen een betalingsregeling voorstellen en deze aan de klant bevestigen per e-mail. Spraakrouteren zal daarbij het bepalen van de intent ("vertel ons waarover u ons belt") en het routeren van het contact aanzienlijk verbeteren," aldus Visser. •

OPLEIDINGEN EN CURSUSSEN

Door:
Maaïke
Groenewege,
Convocat

De vraag naar Nederlandstalige conversation designers is op het moment van schrijven groter dan het aanbod. Met name banken, verzekeraars en energiebedrijven zien steeds meer toegevoegde waarde van conversational interfaces als onderdeel van hun customer experience.

Mooie kansen dus voor dit relatief onbekende vakgebied! Inmiddels hebben ook opleiders deze markt ontdekt en is er behoorlijk wat aanbod beschikbaar.

Conversation Design Institute

Het Amsterdamse Conversation Design Institute biedt een breed aanbod aan online cursussen en certificeringen. De cursussen zijn onderverdeeld in drie leerpaden:

- conversation designer: ontwerpen van flows, sample dialog, focus op design-aspecten en human centered design.
- conversational copywriter: het daadwerkelijk schrijven van conversational copy, ontwerpen van bot persona en persuasive design.
- AI trainer: kennis en vaardigheden voor het optimaliseren van de NLU en algehele botperformance.

Alle cursussen zijn Engelstalig, volledig online en te volgen in je eigen tempo.

www.conversationdesigninstitute.com

SRM: Schrijven voor chatbot en voice

Een tweedaagse cursus van praktijkexpert Susan van Gemeren waarin je aan de hand van actuele cases zelf het hele conversation-designproces doorloopt: van klantverwachtingen, plaats in de customer journey en use-casedefinitie tot het daadwerkelijk ontwerpen van happy flow en fallback en het optimaliseren door gebruikersonderzoek.

<https://www.srm.nl/data-en-tech/voice-en-ai-trainingen/schrijven-voor-chatbot-en-voice.html>

Voice Tech Global

Voice Tech Global (VTG) is een Canadese community en opleider die momenteel twee online klassikale cursussen aanbiedt. VTG werkt met drie docenten en een uitgebreid netwerk van mentoren (inclusief ondergetekende). Studenten wer-

ken tijdens een cursus aan een eigen use case die methodisch wordt uitgewerkt. Iedere cursus wordt afgesloten met een eindpresentatie van de projecten, en een certificering.

- Foundations in Conversational Experience Design: een 10-weekse cursus waar deelnemers de basisbeginselen leren van conversation design, conversational product design, persona design, testen & prototypen, deployment & onderhoud.
- Advanced Conversational Experience Design: een verdiepingscursus waarin deelnemers zich verder ontwikkelen op het gebied van concept en productdesign. User research, concepten uitwerken en valideren, user journey design.

www.voicetechglobal.com

Digital Assistant Academy

De Digital Assistant Academy is met 24 weken en drie capstone projects de meest tijdsintensieve training op het gebied van conversation design. Sterk op de Amerikaanse markt gericht, met aandacht voor carrière en portfolio. Hierdoor zijn adviezen over bijvoorbeeld sollicitatiegesprekken wellicht minder relevant.

www.digitalassistant.academy

Universiteiten en hogescholen

HvA: minor conversation design:

<https://conversational.amsterdam/>

Platformcursussen op Coursera

Contact center AI: conversation design fundamentals:

<https://www.coursera.org/learn/contact-center-ai-conversational-design-fundamentals#syllabus>

Create Your First Chatbot Using Google Dialogflow:

<https://www.coursera.org/projects/create-your-first-chatbot-using-google-dialogflow-9b1zb>

Create Your First Chatbot with Rasa and Python:

<https://www.coursera.org/projects/chatbot-rasa-python>

Build & Deploy AI Messenger Chatbot using IBM Watson:

<https://www.coursera.org/projects/build-deploy-ai-messenger-chatbot-ibm-watson>

NOTaS

STICHTING NOTAS

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

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

DEELNEMERS STICHTING NOTAS

Amberscript B.V.
Keizersgracht 6688
1017 ET Amsterdam
T: 06 13110023
W: www.amberscript.com
E: peterpaul@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.

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

Cornelists B.V.
Beeklaan 38,
1403 RE Bussum
T: 06 44330071
W: www.cornelists.nl
E: serge@cornelists.nl

Met eigen taal- en dialoogtechnologie helpt Cornelists bedrijven toepassingen te ontwikkelen op het gebied van Conversationele Interfaces, Data Science en Gamification. Met ons platform Botprof richten we ons specifiek op het Nederlandse taalgebied en ontwikkelen we vernieuwende technologische mogelijkheden en gespreksvaardigheden voor chatbots en voice-assistenten.

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

DANS is het Nederlands instituut voor permanente toegang tot onderzoeksgegevens. Onze activiteiten zijn geconcentreerd 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
Postbus 24,

5360 AA Grave
T: 0486 486 486
W: www.dedicon.nl
E: info@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.

Fluency
Prins Hendrikkade 159a, 1011 TB Amsterdam
T: 06 14502731
W: www.fluency.nl
E: info@fluency.nl

Fluency maakt tekst-naar-spraaksoftware voor het Nederlands en het Fries. De belangrijkste producten zijn Fluency TTS en Spika. Spika is een handig hulpmiddel bij dyslexie en andere leesproblemen, dat direct vanaf een usb-stick werkt. Fluency TTS is een uitgebreider product, dat via API ook gebruikt kan worden in combinatie met screen readers voor blinden, en andere toepassingen. Beide producten bieden de gebruiker een ruime keuze aan stemmen.

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

Instituut voor de Nederlandse Taal
Rapunberg 61,
2311 GJ Leiden
Postbus 9515,
2300 RA Leiden
T: 071 514 16 48
W: www.ivdnt.org
E: secretariaat@ivdnt.org

Het Instituut voor de Nederlandse Taal is dé plek voor iedereen die iets wil weten over het Nederlands door de eeuwen heen. 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
Postbus 9103,
6500 HD Nijmegen
T: 024 361 16 86
W: www.ru.nl/clst
E: clst@let.ru.nl

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 op ontwikkeling en curatie van data en tools tot onze expertise.

ReadSpeaker
Princenhof Park 13,
5372 NG Driebergen-Rijsenburg

gen, verkrijgen, classificeren, categoriseren en ontsluiten ten behoeve van intelligent (her)gebruik van informatie en kennis. Door innovatief gebruik van taaltechnologie te combineren met zoekmachines en document management systemen realiseren wij oplossingen voor zowel numerieke, tekstuele als engineering gerelateerde informatie. Wij doen dit voor zowel bedrijfsbrede data als Big Data oplossingen en voor zowel bedrijven als overheden.

Martha Hofman
Skaalvestrjitte 16,
9178 GR Wanswert
T: 06 19858338
W: www.marthahofman.nl
E: mahofman@hetnet.nl

Taalkundige werkzaamheden voor taal- en spraaktechnologie, conversation design: lexicons en datasets, part-of-speech tagging, conversation design, intent classification.

Nederlands Instituut voor Beeld en Geluid
Media Parkboulevard 1,
1217 WE Hilversum
T: 035 677 5555
W: www.beeldengeluid.nl
E: info@beeldengeluid.nl

Het Nederlands Instituut voor Beeld en Geluid is een van de grootste audiovisuele archieven van Europa en herbergt ruim 1 miljoen uur aan radio, televisie, film en muziek. Beeld en 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.

Oracle Cloud Service, Oracle Nederland
Hertogswetering 163,
3543 AS Utrecht
T: 030 6699 000
W: www.oracle.com/nl
E: fabrice.nauze@oracle.com

The Oracle Cloud Service Digital Assistant enables the most efficient way to engage with self-service consumers online, shaping next-generation customer experiences. The Digital Assistant provides a deeper understanding of customer intent and helps guide customers to high-value interactions to help improve customer satisfaction, increase customer loyalty, and drive higher conversion.

T: 030 692 4490
W: www.readspeaker.com
E: nederland@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 desktopoplossingen, spraakproductie en meer. Met meer dan 20 jaar ervaring is het team van ReadSpeaker leidend in tekst-naar-spraak.

Telecats
Colosseum 42,
7521 PT Enschede
Postbus 92,
7500 AB Enschede
T: 053 488 99 00
W: www.telecats.nl
E: info@telecats.nl

Telecats levert innovatieve klantcontactoplossingen met spraaktechnologie en kunstmatige intelligentie. Wij assisteren de klant door met spraakherkenning slim te routeren naar de meest geschikte medewerker. Voor en tijdens het gesprek tonen we de medewerker relevante informatie en antwoordsuggesties. We signaleren en rapporteren o.b.v. gesprekskarakteristieken en woordkeuze en leveren optimaal inzicht in klantcontact.

Tilburg University – Cognitive Science & Artificial Intelligence
Warandelaan 2,
5037 AB Tilburg
Postbus 90153,
5000 LE Tilburg
T: 013 466 81 18
W: www.uvt.nl
E: A.H.Verschuur@tilburguniversity.edu

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 semantiek, dialoogmodellen, en computationele modellen van verbale en non-verbale menselijke communicatie. Het departement verzorgt onder andere de bloeiende bachelor- en masteropleidingen Cognitive Science & Artificial Intelligence, en is verantwoordelijk voor de interfacultaire Master Data Science & Society.

Technische Universiteit Delft - Afdeling Intelligent Systems
Van Mourik Broekmanweg 6,
2628 XE Delft
W: www.tudelft.nl/ewi/over-de-faculteit/afdelingen/intelligent-systems
T: 015 2789818
E: o.e.scharenborg@tudelft.nl

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, en veiligheid en incidentbeheer.

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

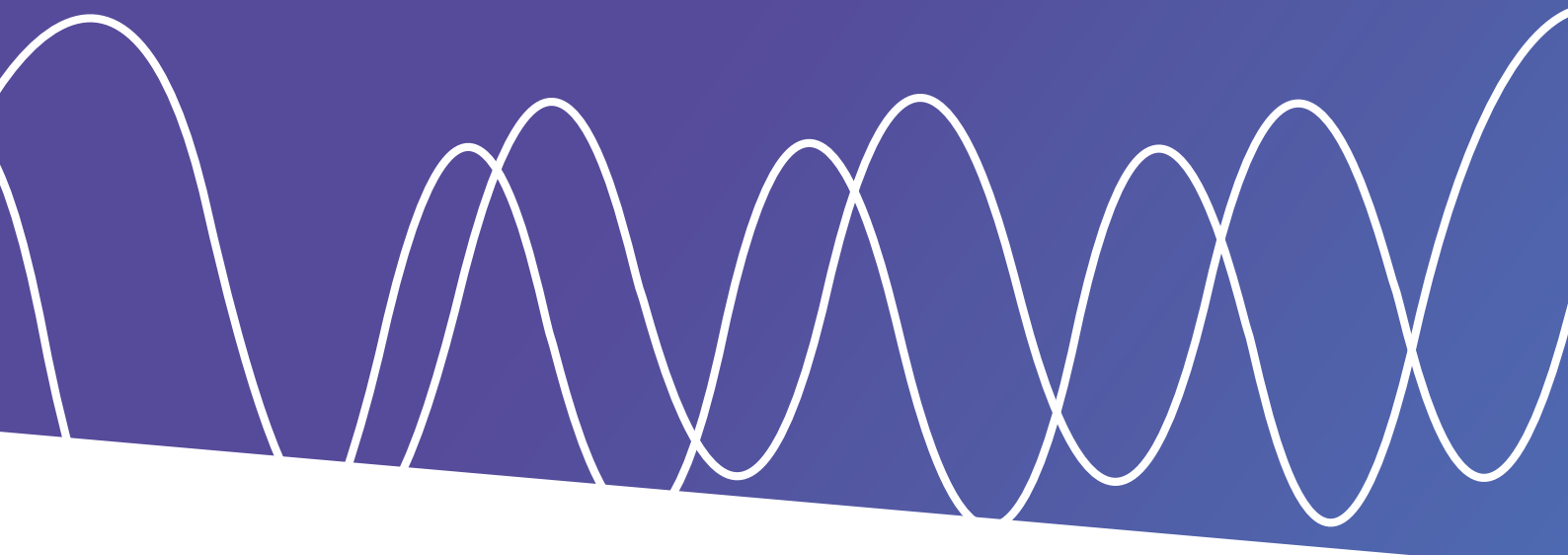
Het Leiden Institute of Advanced Computer Science (LIACS) is het instituut voor onderzoek en onderwijs op het gebied van informatica en kunstmatige intelligentie aan de Universiteit Leiden. Het LIACS werkt actief samen met overheden en bedrijven. Achter de oplossingen voor maatschappelijk relevante problemen 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
Drienerloaan 5,
7522 NB Enschede
Postbus 217,
7500 AE Enschede
T: 053 489 37 40
W: www.utwente.nl/en/eemcs/hmi/
E: HMI-SECR-L@lists.utwente.nl

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 (social) 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 belangrijke 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 (met cursussen zoals Natural Language Processing, Speech Processing, Conversational Agents, en Affective Computing) en de bacheloropleiding Creative Technology.

Universiteit Utrecht – Utrecht Institute of Linguistics OTS
Trans 10,
3512 JK Utrecht
T: 030 253 60 06
E: uil-ots@uu.nl
W: www.hum.uu.nl/uilots

Het Uil OTS is een onderzoeksinstituut binnen de Faculteit Geesteswetenschappen van de Universiteit Utrecht en stelt zich ten doel onderzoek te doen naar menselijke taal. De volgende drie dimensies bepalen het onderzoek: (i) de architectuur van het menselijke taalstelsel, (ii) de cognitieve systemen die ten grondslag liggen aan de verwerking en de verwerking van taal, en (iii) de wijze waarop taal wordt gebruikt in communicatie tussen mensen en tussen mens en machine.



Human Media Interaction aan
de Universiteit Twente doet onderzoek naar:

Spraak & Dialoog

Hoe gaan conversational agents een dialoog aan met mensen,
en hoe kunnen we dit verbeteren?

Sociale & Affectieve Interactie

Hoe herkennen we sociale en affectieve signalen
in gesproken mens-mens en mens-machine interactie?

Spraak & Gezondheid

Hoe kunnen we de fysieke en mentale gesteldheid
van mensen herkennen op basis van hun spraak?

Spoken Document Retrieval

Hoe maak je grote archieven
met gesproken data doorzoekbaar?

Vind ons op: <https://www.utwente.nl/en/eemcs/hmi>
of mail:

Khiet Truong

k.p.truong@utwente.nl

Roeland Ordelman

roeland.ordelman@utwente.nl

**UNIVERSITY
OF TWENTE.**

Human
Media
Interaction