

Bellen met een pratende computer

Als je de sprekende computer van GemeenteConnect (gemeente Gilze en Rijen) aan de lijn krijgt, heb je binnen een mum van tijd alle informatie die je zoekt. Maar begrijpt deze computer je vraag wel echt? Over spraakherkenning en tekstclassificatie.

De meeste mensen weten wel wat leukers te doen dan naar een gemeenteloket te gaan. Lange rijen en veel gedoe met formulieren en administratie en je zult altijd zien dat je niet alle papieren bij je hebt, dus moet je ook nog een keer terugkomen. Helemaal leuk als je overdag werkt en je in de avonduren moet gaan. Om aan je informatie te komen kun je natuurlijk ook bellen. Maar hoe gaat dat met informatienummers? Je moet door een eindeloos menu heen waarbij je met je cijfertoetsen steeds een keuze moet maken: “Voor openingstijden toets een 1, voor identiteitszaken toets 2”, etc.

Het zou veel makkelijker zijn als je je vraag in je eigen woorden kunt zeggen en dat de computer dan gewoon antwoord geeft. Dat klinkt een beetje als Star-Trek maar toch is er zo'n systeem, zelfs voor het Nederlands. Twee taaltechnologiebedrijven uit Delft, Irion Technologies en DutchEar, hebben een programma gemaakt dat een heus gesprek met je aangaat om je naar de juiste informatie te leiden. Dit programma kan gebruikt worden door alle gemeentes in Nederland en geeft antwoord op 80% van de meest gestelde vragen.

Hieronder staat een voorbeeld van zo'n gesprek:

```
Computer: U spreekt met GemeenteConnect, de automatische
          vraagbaak van de gemeente Gilze en Rijen. Wat kan ik
          voor U doen?
Mens:    Mijn buurman wil een boom omhakken, mag dat zomaar?
Computer: OK, ik even voor U kijken.
          (pauze, de computer verwerkt de opdracht)
Computer: Ik geloof dat ik wat voor u heb, wat vindt U van
          "kapvergunning" of "groenvoorziening"? Welke kiest u?
Mens:    Ehh, doe maar vergunningen.
Computer: Uitstekende keus, ik lees U nu de samenvatting van
          "kapvergunning" op.
          (de computer leest de samenvatting voor)
```

Je kunt ook onderstaand geluidsfragment van een opgenomen gesprek beluisteren of zelf bellen met 015-2191180. Hier kun je GemeenteConnect zelf uitproberen.

[Geluidsfragment GemeenteConnect](#)

Hoe werkt zoiets nu? Een computer begrijpt namelijk niet echt wat je zegt. Bovendien verstaat de computer ongeveer 40% van de woorden die je zegt verkeerd, zeker door een telefoon. Een korte uitleg is op zijn plaats.

Als je een boom in je tuin om wilt hakken, zul je eerst een kapvergunning bij de gemeente aan moeten vragen. Vaak moet je dan lang in de rij staan bij het gemeentehuis, maar met GemeenteConnect heb je in no-time je informatie.

Een bom bakken in mei tuin

Eerst wordt het telefoongeluid van je stem omgezet naar woorden. Dat wordt gedaan door een spraakherkenner. Dat is een programma dat weet welke geluidsgolven ongeveer bij Nederlandse woorden horen. Een spraakherkenner is getraind met voorbeelden van vragen die mensen zouden kunnen stellen. Die vragen worden eerst opgenomen uit telefoongesprekken en vervolgens door iemand uitgetypt in woorden, met alle “uhs” and “ahs” erbij. Daarna wordt het geluid gekoppeld aan de losse woorden van de vraag en kan de computer ook andere vragen proberen te herkennen.

Een spraakherkenner is niet foutloos en als er veel ruis op de lijn zit of veel achtergrondlawaai dan wil het wel eens misgaan. Bovendien geeft de spraakherkenner de woorden vaak in willekeurige volgorde terug en voegt woorden toe die niet gezegd zijn. Een zin als “ik wil een boom omhakken in mijn tuin”, komt er dan uit als “ik wil een bom bakken in mei tuin” of als “hij heeft een boom takken in mijn duin”.

Tekstclassificatie

Hoe moet je daar wijs uit worden? De computer doet dat door middel van *tekstclassificatie*. Classificatie van tekst is ook een vorm van computerleren. Je geeft de computer voorbeeldteksten die horen bij bijvoorbeeld een *kapvergunning*. Dat kan de informatie zijn van de gemeente zelf, maar ook vragen van mensen waarvan je weet dat die daarbij horen. De computer kijkt dan naar de woorden die in die tekst staan en kijkt ook in welke andere teksten dat woord gebruikt wordt. Woorden die uitsluitend of voornamelijk in teksten over *kapvergunning* voorkomen zijn namelijk belangrijk voor *kapvergunning*, maar woorden die je ook bij andere soorten teksten vindt zijn minder belangrijk, ook al komen ze vaker voor. De woorden “boom”, “kappen”, “omhakken” verwacht je wel bij *kapvergunning* maar niet bij *identiteitszaken*. Woorden zoals “vergunning” en “aanvragen”, en zeker woorden als “de”, “het”, “zijn”, “worden”, “gemeente” komen in veel of misschien wel alle artikelen voor. Een woord is dus belangrijk voor *kapvergunning* als het veel voorkomt in teksten over *kapvergunning* maar niet in teksten over andere onderwerpen.

Een eenvoudige formule geeft je dan de waarde van een woord voor iedere soort tekst. Je telt het aantal keren dat een woord voorkomt in teksten over *kapvergunning* en deelt dat door het totaal aantal teksten waarin het voorkomt van alle onderwerpen samen. Als “boom” 3 keer voorkomt in een tekst die hoort bij *kapvergunning* en niet bij andere onderwerpen dan krijgt dat een hoge waarde, namelijk: $3:1=3$. Het woordje “de” komt misschien wel 20 keer voor in dezelfde tekst, maar komt ook voor in alle andere teksten, zeg 500 keer. Het krijgt dan slechts een waarde van $20:500=0,004$. Woorden als “de”, “moeten”, “kunnen” zijn dus zo goed als betekenisloos voor het systeem.

Nadat de computer op deze manier de tekst die de spraakherkenner heeft herkend geclassificeerd heeft, geeft hij het best passende onderwerp aan de vragensteller terug. Als deze het met de keuze eens is, leest de computer de samenvatting voor. Hoe meer woorden de klant geeft, hoe beter de tekstclassificatie werkt. Als de computer zich toch niet zeker voelt vraagt hij om verdere uitleg of toelichting en classificeert dan de uitgebreide tekst (de vorige vraag+de toelichting). De computer is onzeker als de score van resultaten dicht bij elkaar zit, hij is zeker als er een groot verschil is tussen het beste resultaat en de andere resultaten. Als de klant zegt dat hij het niet eens is met de classificatie, dan kijkt de computer naar de volgende mogelijkheid, tenzij die niet hoog genoeg scoorde. In dat laatste geval vraagt de computer ook om meer informatie.

Geen grammatica, maar een zak vol woorden

Het leuke van tekstclassificatie is dat het alleen de woorden gebruikt. De computer kijkt niet naar de grammatica of de volgorde van de woorden in de tekst. Dit noemen ze daarom ook wel de *bag-of-words*-methode: je stopt de woorden in een zak, flink schudden en je krijgt het juiste onderwerp eruit. Er is nog een voordeel van tekstclassificatie: onzinwoorden of verkeerd herkende woorden spelen haast geen rol. Kijk maar naar het voorbeeld van verkeerd herkende spraak. Als de computer drie inhoudswoorden van de spraakherkenner krijgt waarvan er twee goed zijn herkend (boom, tuin) en een woord fout (bakken), dan is het meestal zo dat het foute woord niet gebruikt wordt in combinatie met de andere twee woorden. Bij het berekenen van het beste onderwerp telt het programma de score van ieder woord voor ieder onderwerp bij elkaar op. Dus als twee woorden hoog scoren voor *kapvergunning* en een foutief woord hoog voor *openingstijden van de gemeente*, dan zal het systeem toch voor de eerste kiezen. De twee scores worden gewoon samengenomen. Het zal duidelijk zijn dat bij een of twee woorden moeilijk te kiezen valt. In dat geval zal de computer dus meerdere mogelijkheden moeten aanbieden. Daarom vraagt de computer in die situaties ook om meer informatie om een keuze te maken.

In de wereld van GemeenteConnect bestaat geen grammatica. Hij gebruikt alleen de woorden die hij van de spraakherkenner doorkrijgt, de volgorde waarin deze staan doet er niet toe.

Ik zie, ik zie wat jij niet ziet...

Wat voor gesprek wordt er nu eigenlijk gevoerd? Zoals al eerder gezegd: de computer begrijpt de vraag niet echt. Hij geeft ook geen direct antwoord maar leidt je naar de plek waar het antwoord staat. Dat doet hij door de vraag te koppelen aan antwoorden en goed te kijken of de woorden uit de vraag je in de goede richting te sturen. Wat dat betreft gedraagt het programma zich als een spelletje als *ik-zie-ik-zie-wat-jij-niet-ziet*, waarbij de suggesties als *koud*, *warm*, *warmer* en *heet* de computer in de goede richting sturen. Daarnaast wordt ook gekeken naar het hele proces van de communicatie: hoe lang ben je al bezig, zit er schot in of blijft het resultaat onzeker, zijn er veel onbekende woorden, is de spreker aardig, ongeduldig of boos, etc.

Het programma is dus heel goed in het vinden van informatie zonder die informatie te snappen. Dat heeft ook voordelen. Je kunt het systeem namelijk heel makkelijk gebruiken voor andere domeinen dan gemeentelijke informatie. Hetzelfde systeem kan bijvoorbeeld worden gebruikt voor het zoeken van informatie over ziektes (stel je eigen

diagnose), of problemen bij apparatuur of je auto. Het enige dat moet gebeuren is een nieuwe training voor de tekstclassificatie met voorbeelden uit het nieuwe onderwerp en een nieuwe training voor de spraakherkenner met voorbeeldvragen. Dat kan in een paar weken worden gedaan en is vele malen sneller en goedkoper dan het opleiden van een persoon in een nieuw onderwerp.

Gaat een computer ooit een mens begrijpen? Dat is een interessante vraag. Sommige wetenschappers beweren dat het fundamenteel onmogelijk is, omdat een computer niet dezelfde ervaringen en gevoelens heeft als een mens. Sterker nog, ze beweren zelfs dat mensen elkaar niet echt kunnen begrijpen, omdat die gevoelens en ervaringen persoonlijk zijn. Dat wil echter niet zeggen dat er heel nuttig met een computer gepraat kan worden, zoals we hebben gezien. Ook al gaat dat dan op een heel andere manier dan tussen mensen.