

DIXIT

TIJDSCHRIFT OVER TAAL- EN SPRAAKTECHNOLOGIE

TAAL- EN SPRAAKTECHNOLOGIE voor de OVERHEID



**Een efficiëntere en meer klantvriendelijke
overheid dankzij taal- en spraaktechnologie**

ivWebcommunicator



Een revolutie in web- en spraakintegratie

Centrale en lokale overheden streven er naar burgers en bedrijven zo goed en effectief mogelijk te bedienen en voortdurend hun dienstverlening te verbeteren.

Websites wordt steeds meer toegepast om een zo breed mogelijk publiek 24 uur per dag, 7 dagen per week van informatie te kunnen voorzien. Websites moeten toegankelijk gemaakt worden voor slechtzienden en laaggeletterden. Tevens is er een groep mensen met een motorische of manuele beperking die de website wel kunnen zien, maar een PC niet met de normale invoermethoden zoals toetsenbord en muis kunnen bedienen.

Met de toepassing van ivWebcommunicator kunnen al deze groepen ongehinderd een website bedienen en de informatie in beeldscherm presentatie of in gesproken woord tot zich nemen.

Veel mensen die informatie opvragen, meldingen willen doen of afspraken willen maken, gebruiken de telefoon omdat ze dat makkelijker vinden of omdat het beter uitkomt. Door ivWebcommunicator kan de communicatie met de website ook telefonisch plaatsvinden.



Dialogs Unlimited biedt revolutionaire oplossingen, gebaseerd op het innovatieve ImmediateVoice™ platform, waaraan het TechnoPartner label is toegekend door SenterNovem.



ivWebcommunicator™



PC toegang



- Informatie uitwisselen met een website via de computer, door menukeuzes in te spreken in de microfoon of headset en teksten te beluisteren via de speakers of headset.
- Interactief door een website navigeren door het geven van spraakinstructies nadat de keuzemogelijkheden zijn voorgelezen (non-visible mode) of weergegeven (visible mode).
- Gegevens invullen in een website door gesproken tekst.

Telefonische toegang



ivWebcommunicator via telefoon biedt alle mogelijkheden van ivWebcommunicator PC toegang non-visible mode, door naar een opgegeven telefoonnummer te bellen. Hierbij is geen computer nodig. Op deze manier kan een deel of de gehele inhoud van een website telefonisch worden ontsloten en kunnen telefonische transacties (inclusief betaling) worden voltooid.

Smartphone toegang



Een website kan vanaf een smartphone naar keuze met spraak, toetsenbord en/of stylus bediend worden.

ivWebcommunicator via Smartphone is de nieuwste ontwikkeling van Dialogs Unlimited. De eerste release zal de volgende functionaliteit bevatten:

- Weergave van de gesproken informatie op het scherm van de smartphone.
- Simultane invoermogelijkheid via spraak en druktoetsen van de smartphone.

Via de ontwikkelomgeving van ivWebcommunicator kunnen nieuwe webgebaseerde telefoniediensten snel gecreëerd worden door gebruik te maken van bestaande websites. Door de internet gebaseerde architectuur kunnen nieuwe diensten direct worden geïntroduceerd bij abonnees zonder dat deze daarvoor een applicatie hoeft te downloaden en/of te installeren.

Voor meer informatie:

Dialogs Unlimited

Dialogs Unlimited BV
Takkebijsters 17 sub 11
4817 BL Breda
T 076 - 57 24 777 - F 076 - 57 12 022
E info@dialogsunlimited.com
I www.dialogsunlimited.nl

INHOUD

Algemeen

- TST voor de overheid 5
- Communiceren is een kerntaak van de overheid: TST kan daarbij helpen! 6

Heldere communicatie

- Woordenlijst.org en Taaladvies.net 8
- Begrijpelijke brieven met Klinkende Taal 10

Toegankelijke overheid

- Géén momentje geduld alstublieft! 13
- Onder burgers en professoren 14
- Duurzame opslag van en lange-termijntoegang tot onderzoeksdata 14
- Toegankelijker websites door TST 16

Wat wil de burger?

- Yournews: automatisch uw eigen nieuws 18
- TST voor een democratische en klantgerichte dienstverlening 19
- DuOMAn: Dutch Language Online Media Analysis 20

Instrumenten voor de organisatie

- TST voor de Rechtspraak, de politie en het NFI 23
- Automatisch vertalen en de overheid 25
- TDA: één wereldomvattend vertaalgeheugen 27
- Digitale bronnen: termextractie helpt bij het zoeken naar informatie. 28
- Spraak in bedrijf 30

En verder

- Voorwoord 3
- Colofon 3
- NOTaS-nieuws 4
- NOTaS-agenda 17
- Stijgend bezoek www.notas.nl 21
- Directory 31

COLOFON

DIXIT: Tijdschrift over toegepaste taal- en spraaktechnologie – 6e jaargang, editie TST voor de Overheid. DIXIT is een uitgave van Stichting NOTaS, Postbus 31070, 6503 CB Nijmegen, Tel. 024 - 352 88 88 Fax 024 -354 00 90 www.notas.nl **Redactieadres:** Stichting NOTaS, Postbus 31070, 6503 CB Nijmegen **Redactie:** Henk van den Heuvel: h.vandenheuvel@let.ru.nl, Arjan van Hessen: hessen@cs.utwente.nl, Remco van Veenendaal: remco.vanveenendaal@inl.nl, René Ouëndag: rene.ouendag@q-go.com, Anne Wijnen: notes@malta-online.nl **Gastredacteuren:** Folkert de Vriend: fdevriend@taalunie.org, Catia Cucchiari: ccucchiari@taalunie.org **Advertenties:** Stichting NOTaS, Anne Wijnen: notes@malta-online.nl, 024 -352 88 88 **Abonnementen:** Voor een gratis abonnement kunt u zich wenden tot een van de NOTaS-deelnemers **Druk:** Leonard Nijmegen bv **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.

Voorwoord

De branche Taal- en Spraaktechnologie blijft innoveren

De economische uitdagingen van dit moment hebben nauwelijks effect op enkele sterke eigenschappen van de TST-branche: innovatiekracht en doorzettingsvermogen. Dit werd duidelijk zichtbaar tijdens de afgelopen STEVIN-dag op 4 september jl. op de campus van de Universiteit van Tilburg waar onderzoekers en ontwikkelaars uit kennisinstellingen en het bedrijfsleven bij elkaar kwamen voor de jaarlijkse kennissuitwisselingsdag.

TST – Nice to have or need to have

Op dit moment kijken bedrijven vooral kritisch naar de efficiëntiewinst van mogelijke investeringen. Heldere communicatie, één van de hoekstenen van efficiënte bedrijfsvoering, speelt daarbij zowel in het bedrijfsleven als bij de overheid, een belangrijke rol. Ook de mogelijkheden van (gedeeltelijke) elektronische zelfbediening door burgers, worden gezien als een belangrijke stimulans voor een efficiëntere overheid. In deze DIXIT leest u meer over de wijze waarop TST op meerdere fronten een sleutelrol speelt binnen de overheid.

Is investeren nog mogelijk?

Mede dankzij diverse subsidiemogelijkheden en het STEVIN-stimuleringsprogramma, blijven we als TST-branche investeren in "spraakmakende" applicaties. Zo kunnen wij ervoor zorgen dat er ook in de toekomst toonaangevende taal- en spraakoplossingen zullen zijn voor het Nederlands.

NOTaS = Samen sterk

De bedrijven binnen de Nederlandse Organisatie voor Taal- en Spraaktechnologie (NOTaS) komen uit het MKB. Ze hebben in de regel beperkte middelen en menskracht beschikbaar voor in-house onderzoek- en ontwikkelingsactiviteiten en zijn dus sterk gebaat bij een nauwe samenwerking met (Nederlandse) kennisinstellingen. NOTaS faciliteert deze samenwerking en is alleen daarom al belangrijk. De samenwerking onderling, ook tussen bedrijven en kennisinstellingen die op andere fronten met elkaar concurreren, mag een voorbeeld zijn voor andere branches van hoe we samen veel sterker kunnen zijn, vooral in moeilijke tijden!

Gastredactie

Bij deze DIXIT over TST voor de overheid zijn we bijzonder blij met de input van onze gastredacteuren van de Nederlandse Taalunie; Folkert de Vriend en Catia Cucchiari. De artikelen die zij samen met de vaste redactie hebben gekozen, laten zien op hoeveel verschillende manieren TST de overheid kan ondersteunen bij het uitvoeren van haar taken.

Wij blijven innoveren zodat u in de toekomst dit ook kunt doen!

Namens de hele redactie en NOTaS wens ik u veel leesplezier!
Debbie Kenyon-Jackson, Voorzitter Stichting NOTaS

NOTaS-nieuws

Zorg, overheid en bedrijfsleven 2009

Op dit 5e congres, georganiseerd door Flevum (www.flevum.nl) en gesponsord door STEVIN hebben vertegenwoordigers van NOTaS gesproken met opinieleiders uit de gezondheidszorg. Tijdens een presentatie, gehouden door prof. dr. A.C.M. Rietveld, Radboud Universiteit Nijmegen & Sint Maartenskliniek Nijmegen en dr. A. van Helsen, Telecats & Universiteit Twente, zijn de mogelijkheden voor TeleHealth en spraak- en taaltherapie toegelicht.

Te gast bij de Nederlandse Taalunie

Eerder dit jaar waren NOTaS-deelnemers te gast in het prachtige pand van de Nederlandse Taalunie aan het Lange Voorhout in Den Haag. De Taalunie is de beleidsorganisatie waarin Nederland, België en Suriname samenwerken op het gebied van de Nederlandse taal, taalonderwijs en letteren. Voor NOTaS speelt de Taalunie een essentiële rol in het ondersteunen van diverse activiteiten en als coördinator van STEVIN, het meerjarige onderzoeks- en stimuleringsprogramma voor Nederlandstalige taal- en spraaktechnologie dat gezamenlijk door de Vlaamse en Nederlandse overheid wordt gefinancierd.

Taalunieversum

Op de website van de Nederlandse Taalunie, taalunieversum.org/taal/technologie, vinden bestaande en toekomstige afnemers van TST een uitgebreid beeld van de TST-sector in het Nederlands taalgebied. Studenten, onderzoekers en ontwikkelaars worden tevens op de hoogte gehouden van de laatste onderzoeksprojecten en ontwikkelingen in onze dynamische branche.

Tallose subsidiemogelijkheden

Aansluitend aan de deelnemersbijeenkomst bij de Taalunie ging een kleine delegatie namens NOTaS op bezoek bij SenterNovem, om subsidies en financieringsmogelijkheden voor TST-innovatie te bespreken. De conclusies zijn beschikbaar voor NOTaS-deelnemers via info@notas.nl.

Deelnemersbijeenkomst PCM 16 juni

(verslag door een van de deelnemers)
De opkomst viel een beetje tegen en dat was vooral jammer voor de mensen die niet gekomen zijn. Eerst was er een interessante lezing door de adjunct hoofdredacteur van Trouw over enerzijds de rol van de kranten in de huidige maatschappij en anderzijds de inzet van nieuwe technologie (samenwerken

van Internet, filmpjes en het oude vertrouwde papier) bij het maken van de moderne krant. Daarna volgde een heerlijke lunch en een spectaculaire rondleiding waarin je kon zien hoe 1500 kg zware rollen papier met een snelheid van 40 km/uur veranderden in de NRC van die dag. Bij het weggaan kregen we die pas gedrukte krant mee: echt vers van de pers.

Zichtbaar aanwezig

NOTaS wordt vertegenwoordigd op diverse evenementen, symposia en congressen. Nu hebben we ook onze eigen banner, die tevens als projectieschermbanner fungeert, om ervoor te zorgen dat we bij iedere gelegenheid het allerlaatste nieuws en demo's kunnen laten zien. Hebt u een bijeenkomst voor studenten, ontwikkelaars of klanten en wilt u meer weten over de laatste ontwikkelingen op het gebied van TST voor uw branche, neem dan contact op met het bestuur voor een toegespitste presentatie: info@notas.nl.

Nieuwe deelnemers

De afgelopen maanden hebben we twee nieuwe deelnemers mogen verwelkomen bij NOTaS: VoiceCorp Webservices B.V., een bedrijf dat zich bezig houdt met het beluisterbaar maken van websites (zie dit nummer), en het Utrecht Institute of Linguistics OTS (UIL-OTS) van de Universiteit van Utrecht, een kennisinstelling op het gebied van taal- en spraaktechnologie. Welkom in ons midden!

Tot slot: Voortzetting activiteiten Polderland

Soms vertrekt een deelnemer uit NOTaS omdat hun organisatie minder nadruk op TST gaat leggen. Het bestuur is echter geschrokken van het nieuws dat een van de NOTaS-oprichters, Polderland Language & Speech Technology, de huidige economische crisis niet heeft kunnen doorstaan. Voormalig directeur eigenaar Theo van den Heuvel is een van de founding fathers van de taaltechnologie in Nederland. Polderland werd opgericht in 1993 en groeide snel uit tot een toonaangevend IT-bedrijf, dat gespecialiseerd was in taaltechnologische oplossingen. We zijn zeer verheugd om te horen dat de activiteiten van dit pareltje in de Nederlandse TST-industrie door Knowledge Concepts, in samenwerking met voormalige medewerkers van Polderland, worden voortgezet. We wensen hun veel succes! Voor meer informatie: info@knowledge-concepts.com

**Debbie
Kenyon-Jackson
voorzitter
stichting NOTaS**

Taal- en spraaktechnologie voor de overheid

De overheid zet momenteel volop nieuwe ICT in voor een betere dienstverlening voor burgers en bedrijven en een hogere efficiëntie van interne werkprocessen. Onder meer in het verlengde van de EU voeren Nederland, België en Vlaanderen gericht beleid voor deze 'e-overheid' (digitale overheid). Deze DIXIT laat zien hoe taal- en spraaktechnologie (TST) kan bijdragen aan de verwezenlijking van de e-overheid. In verschillende regio's komen burgers momenteel al met TST in aanraking wanneer ze het nieuwe netnummer 14+ bellen. Na het inspreken van de naam van een gemeente wordt de beller met spraakrouting automatisch doorgeschakeld naar het juiste loket. Er is met TST echter nog veel meer mogelijk. Het eerste artikel van dit nummer vertrekt vanuit een breder beleidsperspectief. Veel van wat Johan Van Hoorde in zijn tekst naar voren brengt, wordt in de rest van dit nummer geïllustreerd met concrete bedrijfs- en projectresultaten, o.a. afkomstig uit het STEVIN-programma. Deze DIXIT is verder opgedeeld in vier thema's die voor de overheid van belang zijn.

**Folkert de Vriend
en
Catia Cucchiarini**
**Nederlandse
Taalunie**

Heldere communicatie

Als de overheid goed en efficiënt met haar burgers wil communiceren dan moet ze heldere taal hanteren. Dirk Caluwé gaat daarom in op online beschikbare taalhulpmiddelen voor de schrijvende ambtenaar. Tigran Spaan en Oele Koornwinder bespreken vervolgens een product dat geïntegreerd kan worden in een tekstverwerker en aanwijzingen geeft om de leesbaarheid van de tekst te verbeteren.

Toegankelijke overheid

De overheid moet verschillende communicatiekanalen faciliteren, ook voor mensen met een handicap. Its Kievits beschrijft hoe websites voor mensen met een visuele handicap automatisch kunnen worden voorgelezen met TST. Ook verwacht de burger steeds meer dat de overheid altijd bereikbaar is. Joop van Gent bespreekt daarom een prototype waarmee burgers 24/7, zonder wachttijden telefonisch te woord worden gestaan door een op TST gebaseerd systeem. De bijdragen van Henk van den Heuvel en René van Horik, ten slotte, gaan in op duurzame opslag van informatie waardoor toegang tot die informatie ook op de langere termijn gewaarborgd is.

Wat wil de burger?

Dankzij eParticipatie kunnen burgers een adviserende rol in het beleidsproces spelen. Hun 'probleem- en contextkennis' kan gebruikt worden voor betere en creatievere oplossingen. Zo bespreekt Vincent de Klerk een project waarin van individuele burgers feedback over de dienstverlening werd verkregen met gebruik van TST. De ondersteuning van media-analisten bij het monitoren van de reputatie van hun organisatie wordt

door Valentin Jijkoun en Maarten de Rijke besproken. Joop van Gent beschrijft vervolgens een systeem waarmee het alsmaar groeiende nieuwsaanbod beheersbaar wordt gehouden, bijvoorbeeld voor een knipselkrant.

Instrumenten voor de organisatie

Met het juiste instrument wordt ook de interne werking van de overheid efficiënter. Arjan van Hessen bespreekt het gebruik van TST bij Justitie en politie voor het doorzoeken van spraakopnames. Vincent Vandeghinste en Jaap van der Meer beschrijven de mogelijkheden van automatisch vertalen. Frieda Steurs gaat vervolgens in op betere ontsluiting van informatie door termextractie. Tot slot bespreken Christophe van Bael en Diana Binnenpoorte de inzet van spraakrouting.

Het is niet vanzelfsprekend dat TST ook met het Nederlands overweg kan. Daarom is enkele jaren geleden het STEVIN-programma opgezet, een meerjarig onderzoeks- en stimuleringsprogramma voor Nederlandstalige TST dat door de Vlaamse en Nederlandse overheid wordt gefinancierd (www.stevin-tst.org). Verschillende van de in dit nummer besproken TST-oplossingen zijn binnen dit programma ontwikkeld. Als coördinator van STEVIN ziet de Nederlandse Taalunie het als haar taak om er voor te zorgen dat de resultaten van het programma ook zo veel mogelijk hun weg vinden naar concrete producten. Deze doelstelling sluit uitstekend aan bij de huidige overheidsbrede initiatieven in Nederland, België en Vlaanderen voor de realisering van de e-overheid.

Communiceren is een kerntaak van de overheid: TST kan daarbij helpen!

Taal- en spraaktechnologie (TST) is voor steeds meer mensen een alledaagse zaak. Zelfs dat die met Nederlands overweg kan. We vinden het normaal dat het navigatiesysteem in de auto Nederlands spreekt, of dat we onze mobiele telefoon kunnen vragen om naar kantoor te bellen. En wie staat er stil bij wat er allemaal komt kijken bij de kringeltjes onder slecht gebouwde zinnen in een tekst?

Johan Van Hoorde
Nederlandse
Taalunie

Om dat soort dingen met een taal te kunnen doen, zijn basisvoorzieningen nodig, bijvoorbeeld uitspraakinformatie en programmatuur die zinnen in relevante delen kan opsplitsen. Die voorzieningen bleken voor het Nederlands vaak te ontbreken of van te bescheiden kwaliteit. Sinds meer dan tien jaar werkt de Nederlandse Taalunie hieraan, samen met vele partners in Nederland en Vlaanderen. Ontbrekende onderdelen worden ontwikkeld via het Vlaams-Nederlandse STEVIN-programma (www.stevin-tst.org), dat door de Vlaamse (EWI, IWT en FWO) en Nederlandse (EZ, NWO en OCW) overheid wordt gefinancierd en door de Taalunie wordt gecoördineerd. De resultaten (corpora, lexica, taalmodules) worden beheerd en up-to-date gehouden door de TST-Centrale van de Taalunie (www.tstcentrale.org). Via dat kanaal zijn ze ook beschikbaar voor gebruikers.

Veel ruwe grondstof als het ware, die zelf weinig in de kijker liep, maar er wel voor kon zorgen dat navigatiesystemen, telefoon-toestellen en tekstverwerkers natuurlijk ook met Nederlands overweg kunnen. Nu dat is bereikt kan directer worden ingespeeld op de behoeften van specifieke doelgroepen.

De Taalunie wil doelgroepen nu helpen om hun weg naar de voorzieningen te vinden. Centraal in die aanpak is de overheid zelf, niet alleen de centrale overheid maar ook provincies, gemeenten en instellingen binnen de publieke sector. Vrijwel alle overheden communiceren veel en vaak. In zekere zin kan men stellen dat communicatie een kernactiviteit van de overheid is. Ze is immers een organisator van maatschappelijke consensus en mediator tussen uiteenlopende inzichten en belangen in de samenleving. Hoe anders dan in communicatieve contacten met burgers en belangengroepen kan de overheid zorgen voor voldoende draagvlak voor haar maatregelen en besluiten?

Daarom wellicht dat de overheid zelf steeds meer belang is gaan hechten aan goede communicatieve vaardigheden en efficiënte communicatie-instrumenten.

Dat is ook het kader waarin Nederland, België en Vlaanderen e-governmentbeleid voeren, ook wel aangeduid als *eOverheid* of als *Digitale Overheid*. Door gebruik te maken van internet en ICT-toepassingen willen de overheden de kwaliteit en toegankelijkheid van hun informatie en dienstverlening verbeteren. Informatie komt sneller bij de burgers, wordt toegankelijk voor meer mensen en is minder afhankelijk van plaats en tijd. Bovendien kunnen burgers sneller en gemakkelijk hun zegje doen als het gaat om doelstellingen en inhoud van beleid. Anders gezegd: *eOverheid* zorgt voor meer interactiviteit. Er is niet alleen informatie van de overheid naar de burger, maar ook omgekeerd!

De Nederlandse Taalunie wil dat beleid ondersteunen met een actieplan Burger, Taal en Overheid. Het plan zal drie onderscheiden actielijnen omvatten:

- (a) het ondersteunen van e-government;
- (b) het ondersteunen van het bestaande overheidsbeleid voor *open standaarden* en *opensource software*;
- (c) het ondersteunen van overheidsinstanties bij het hanteren van correct, begrijpelijk en burgergericht Nederlands.

Eigenlijk is het niet meer dan vanzelfsprekend dat juist ook de overheid zelf mee kan profiteren van de resultaten van het TST-beleid. Toch vinden IT'ers uit de publieke sector niet vanzelfsprekend hun weg naar deze materialen. Men geeft zich zelfs onvoldoende rekenschap van de gebruiksmogelijkheden, terwijl de inzet van taal- en spraaktechnologie toch aanzienlijke kwaliteitsverbetering kan bieden, met name voor die doelgroepen die moeilijker bereikbaar

zijn, zoals voor mensen met communicatieve beperkingen. Denken we maar aan automatische voorleesmodules binnen overheidswebsites, die blinden en slechtzienden in staat moeten stellen schriftelijke – dus voor die groep moeilijk toegankelijke – informatie door de computer te laten voorlezen. Dit is slechts één voorbeeld maar het laat wel goed zien dat de overheid moet weten wat er op de TST-markt voorhanden is en hoe dit kan worden ingezet.

De TST-materialen kunnen ook het overheidsbeleid voor *open standaarden* en *opensource software* ondersteunen. Nederland, België en Vlaanderen willen het gebruik van open standaarden en open toepassingen bevorderen, omdat die verschillende systemen beter laten samenwerken en gebruikers minder product- en producentafhankelijk maken. Goede wil volstaat daartoe niet. Eén mogelijke hinderpaal is alvast de wijze waarop open toepassingen omgaan met het Nederlands, bijvoorbeeld of er spelling-, stijl- en grammaticacontrole is en hoe goed die werkt. Mensen die overstappen van gesloten naar open toepassingen zullen er zeker niet op achteruit willen gaan. Daarom wil de Taalunie de TST-materialen inzetten voor het ontwikkelen van taalmodules, zoals spelling- en stijlcontrole, synoniemen, maar ook tekst-naar-spraakmodules, terminologiebeheer en vertaalmodules. Onze basisvoorzieningen bieden voor Nederlandstaligen meer ondersteuningsmogelijkheden dan er nu zijn, terwijl bestaande functies aanzienlijk beter zouden kunnen!

Ter voorbereiding van het plan heeft de Taalunie tal van contacten gelegd, vooral over ambtelijk taalgebruik. De ambtelijke taal was zelfs het centrale werkthema in 2008, onder het motto 'Burger, Taal & Overheid'. Onze contacten met overheidsinstanties wijzen steeds weer uit dat er grote behoefte bestaat aan een centrale, bemiddelende en faciliterende instantie. Vaak wordt erover geklaagd dat er teveel in gespreide slagorde wordt gewerkt, waardoor het resultaat van de inspanningen niet altijd optimaal is.

Voorale lokale overheden ontbreekt het aan middelen en knowhow. "We willen wel", is een veelgehoorde reactie, "maar we weten niet waar we moeten beginnen". Die reacties lijken overeen te komen met een belangrijke vaststelling in het rapport 'Burgers en eOverheid' dat Ernst & Young in het voorjaar presenteerde. Daarin stelt het onderzoeksbureau:

Het verder centraliseren en inrichten van

uniforme ICT-voorzieningen voor overheidsinstanties kan de ontwikkeling van de digitale dienstverlening effectiever en efficiënter maken. Tot dusver worden ICT-projecten veelal decentraal uitgevoerd door afzonderlijke overheidsinstanties. Centrale, gezamenlijke, klantgerichte ICT-oplossingen voor meerdere instanties vormen nog de uitzondering. Bij verdere centralisatie zou de rijksoverheid meer de regie moeten nemen. Het denken vanuit de separate overheidsinstantie moet hierbij vervangen worden door het centraal stellen van de burger en de onderneming.

De Taalunie is natuurlijk geen kandidaat om zelf de regie voor de *eOverheid* te voeren. Wel kan ze een belangrijke schakel vormen in een grensoverschrijdende samenwerking, omdat ze goed zicht heeft op de betrokken instanties in de beide landen en dicht bij de beide regeringen opereert. Veel pleit voor een grensoverschrijdende aanpak. De ervaring van de Taalunie heeft geleerd dat die tot betere resultaten én tot lagere kosten leidt. Zoals het geval was met het ambitieuze STEVIN-programma. Waarschijnlijk hadden de landen apart nooit dezelfde resultaten bereikt, noch naar omvang noch naar kwaliteit.

In een centrale regie ziet de Taalunie verder uitstekende kansen voor het maximaal inzetten van taal- en spraaktechnologie in innovatieve projecten. De overheid zal gemakkelijker zijn weg vinden naar TST als ze dat vanuit een centrale regiekamer kan doen.

De komende jaren zal de Taalunie taal- en spraaktechnologie nadrukkelijk op de agenda plaatsen van de overheid als gebruiker van informatiesystemen. Ze zal symposia opzetten, masterclasses organiseren, best practices in de schijnwerper plaatsen en pilotprojecten opzetten. Deze DIXIT laat zien hoe veel mogelijkheden daartoe bestaan!

Meer informatie over TST is te vinden op het Taalunieversum (www.taalunieversum.org/taal/technologie)



e-OVERHEID



Woordenlijst.org en Taaladvies.net

De websites *Woordenlijst.org* en *Taaladvies.net* behoren voor het Nederlands tot de populairste online taalhulpmiddelen. *Woordenlijst.org* heeft in de eerste plaats een normerende functie voor de spelling. *Taaladvies.net* speelt vooral een faciliterende en stimulerende rol voor het gebruik van de Nederlandse standaardtaal. Door de integratie van de twee hulpmiddelen en de interactieve aanpak worden taalgebruikers op een efficiënte en klantgerichte manier geholpen.

Woordenlijst.org

Zowel in Nederland als in België zijn overheidsorganen verplicht om de officiële Nederlandse spelling te gebruiken. Hetzelfde geldt voor de onderwijsinstellingen die door de Nederlandse en de Vlaamse overheid georganiseerd worden of van die overheden subsidies krijgen voor de organisatie van het onderwijs. De *Woordenlijst Nederlandse Taal*, beter bekend als het *Groene Boekje*, bevat een overzicht van de spellingregels zoals ze vastgesteld zijn door het Comité van Ministers van de Nederlandse Taalunie. Voor de dagelijkse praktijk is vooral de lijst met de spelling van meer dan 110.000 woorden belangrijk.

Bij de publicatie van de *Woordenlijst* in oktober 2005 werden er zoals bij de vorige editie in 1995 een papieren versie en een cd-romversie te koop aangeboden. Bij de nieuwe editie werd ook met een gratis onlineversie van de *Woordenlijst* gestart. Net zoals het nieuwe *Groene Boekje* een bestseller op de boekenmarkt werd, is de onlineversie een succesnummer geworden. Het *Groene Boekje* op internet is een van de populairste diensten van de Nederlandse Taalunie. Sinds februari 2006 is de website meer dan 15 miljoen keer geraadpleegd. Per week worden circa honderdduizend woorden opgezocht.

Door de eenvoudige zoekfunctie op *Woordenlijst.org* kan elke taalgebruiker de correcte spelling van een woord gemakkelijk opzoeken. Wie een woord fout intikt (bijvoorbeeld *nivo* of *copie*), komt meestal toch uit bij de juiste spelling (*niveau*; *kopie*). Bovendien is *Woordenlijst.org* helemaal actueel: de errata die in de verschillende drukken van de papieren versie zijn opgenomen, zijn ook in *Woordenlijst.org* doorgevoerd. Als iemand een woord toch niet kan vinden, krijgt die via het antwoordscherm de suggestie om het

woord via *Taaladvies.net* aan een taaladviseur voor te leggen. Als een taaladviseur een spellingkwestie voorgelegd krijgt waarop hij geen eenduidig antwoord kan geven, vraagt hij advies aan het Instituut voor Nederlandse Lexicologie of de Spellingcommissie van de Nederlandse Taalunie. Op die manier kunnen taalgebruikers formeel uitsluitel krijgen over lastige spellingkwesties. Zo heeft *Woordenlijst.org* niet alleen een normerende maar ook een arbitrerende functie bij onduidelijke kwesties, bijvoorbeeld in de context van het taalgebruik van de overheid.

Taaladvies.net

Ook *Taaladvies.net*, de taaladviesite van de Nederlandse Taalunie, is een veelgebruikt hulpmiddel. De site bevat meer dan duizend teksten met veelgestelde taalvragen en de bijbehorende antwoorden. *Taaladvies.net* werd gelanceerd in maart 2001, maar bouwde eigenlijk voort op de cd-rom *Taaladviesbank*. Die was al in de jaren negentig verschenen en oorspronkelijk bedoeld als vaktechnisch hulpmiddel voor taaladviseurs en andere taalspecialisten. Door de onlineversie werd de bank met taaladviezen gratis ontsloten voor het grote publiek. Op die manier kan het corpus met taaladviezen ook

Dirk Caluwé
Vlaamse Over-
heid, Departement
Diensten voor het
Algemeen
Regeringsbeleid

op een eenvoudige manier verrijkt en actueel gehouden worden. *Taaladvies.net* kan doorzocht worden door het intikken van een (voorbeeld)woord, maar het is ook thematisch raadpleegbaar. Rubrieken als 'wetgeving en normen', 'namen van instellingen en organisaties' en 'leestekens' maken het mogelijk om kwesties in samenhang te raadplegen. Wanneer bezoekers het antwoord op een vraag onverhoopt niet vinden, kunnen ze hun vraag via een webformulier voorleggen aan een taaladviseur. Voor deze service doet de Nederlandse Taalunie in Vlaanderen een beroep op de Taaltelefoon, de openbare taaladviesdienst van de Vlaamse overheid. Voor Nederland werkt de Taalunie samen met de taaladviesdienst van het Genootschap Onze Taal.

Taaladvies.net moet ervoor zorgen dat vragen waarop het antwoord al in de databank te vinden is, zonder telefonisch of ander menselijk contact worden beantwoord. Dat is in de praktijk ook het geval. In 2008 was de website goed voor 4,4 miljoen geconsulteerde pagina's. De pagina 'Stel een vraag over taalgebruik' werd in dat jaar 12.372 keer ingevuld, waarna door de zoekmachine een of meer antwoordsuggesties werden gegeven. Van die vragenstellers stuurden 6.619 personen hun vraag uiteindelijk niet door. Daaruit mag afgeleid worden dat veel vragenstellers kennelijk tevreden zijn met de antwoordsuggesties die ze automatisch via de zoekfunctie van *Taaladvies.net* krijgen aangeboden. Hoewel het aantal bezoekers sinds de start in 2001 exponentieel is gegroeid, blijft het aantal via het formulier gestelde vragen op die manier onder controle.

Hoewel de Nederlandse Taalunie het niet als haar taak beschouwt om te bepalen wat wel of niet goed Nederlands is, heeft ze toch het initiatief genomen om *Taaladvies.net* op te zetten. Ze wil daarmee als overheidsorgaan een faciliterende en stimulerende rol spelen voor 'de gemeenschappelijke bevordering van de kennis en het verantwoorde gebruik van de Nederlandse taal' (artikel 3 van het Taalunieverdrag). Het veelvuldige gebruik toont aan dat *Taaladvies.net* in de behoefte van veel sprekers en schrijvers van het Nederlands voorziet. Dankzij het internet kan snel en gemakkelijk voldaan worden aan de vraag naar deskundig advies.

Normvorming

Alle adviezen van *Taaladvies.net* zijn geschreven door een team van ervaren taaladviseurs en -deskundigen. Voordat de adviezen op *Taaladvies.net* gezet worden, zijn ze

bovendien in het *Taaladviesoverleg* van de Nederlandse Taalunie besproken en goedgekeurd. Die goedkeuring garandeert de taalgebruiker dat de adviezen betrouwbaar zijn. Aan dit overleg nemen mensen deel uit de Nederlandse en Vlaamse taaladviespraktijk. De adviezen geven dus niet alleen maar het standpunt van een enkele auteur weer. Met uitzondering van de spellingadviezen hebben de adviezen geen 'officieel' of 'wettelijk' karakter. Toch bieden ze de gebruiker een grote garantie dat ze actueel, juist en volledig zijn. In die zin hebben ze een normvormende functie.

Taalvariatie

Taaladvies.net speelt ook een innoverende rol bij de visie op taalvariatie. Voor elk advies wordt een documentatiedossier opgesteld waarin per variant informatie wordt verzameld over het gebruik, de frequentie en de attitude van taalgebruikers. Door de nieuwe werkwijze is het mogelijk om taalvariatie in de standaardtaal genuanceerder te beschrijven, conform de aanbevelingen die de Raad van de Nederlandse Taal en Letteren in september 2003 betreffende taalvariatie heeft uitgebracht. In de taaladviezen wordt uitgegaan van een principiële taalpolitieke en taalkundige gelijkwaardigheid van de standaardtaal in België en de standaardtaal in Nederland. Het gebruik van labels zoals: 'standaardtaal in het hele taalgebied', 'standaardtaal in België', 'standaardtaal in Nederland' en 'geen standaardtaal' maakt het mogelijk voor taalgebruikers – onder wie heel wat medewerkers van overheidsdiensten – om de variatie in de standaardtaal vakkundig te beoordelen.

**Tigran Spaan
en
Oele Koornwin-
der
GridLine B.V.**

Begrijpelijke brieven met Klinkende Taal

Communiqueert uw organisatie klantgericht en open? Schrijven uw collega's mooie volzinnen... waar de burger niets van begrijpt? Volgt uw klant altijd wat u bedoelt? Gebruikt u heldere terminologie op uw website of vinden de lezers het onbegrijpelijk jargon?

Overheidsorganisaties en bedrijven willen helder communiceren. Een organisatie die helder communiceert laat zien dat zij open en klantgericht is. Heldere communicatie voorkomt problemen en bespaart kosten: minder klachten, minder telefoontjes en minder onduidelijkheid.

Veel organisaties schakelen inhoudelijke experts in voor het schrijven van brieven, webpagina's en brochures. Dat zijn niet altijd mensen die gewend zijn te schrijven voor het algemeen publiek. Om deze experts bij het schrijven te ondersteunen heeft GridLine, mede dankzij financiering vanuit het STEVIN-programma, het product Klinkende Taal ontwikkeld (www.klinkendetaal.nl).

Met Klinkende Taal kunnen professionals zich richten op de inhoud. Klinkende Taal is altijd aanwezig en met één druk op de knop aan te roepen tijdens het schrijven. Het programma geeft duidelijke aanwijzingen om de leesbaarheid van de tekst te verbeteren en af te stemmen op de doelgroep.

GridLine is een IT-bedrijf dat gespecialiseerd is in taaltechnologie voor het Nederlands. Doordat we gespecialiseerd zijn in het Nederlands, hebben wij alle mogelijke basis-modules en applicaties voor het Nederlands klaar liggen, zoals zoekmachine-optimalisatie, spelling, automatische classificatie, terminologiebeheer, opinion mining, spraakherkenning en ook leesbaarheids-beoordeling. Onze hardcore IT'ers integreren onze taalproducten in veelgebruikte software, zoals Microsoft Word, Sharepoint, FAST en Lucene.

Overheidsbeleid

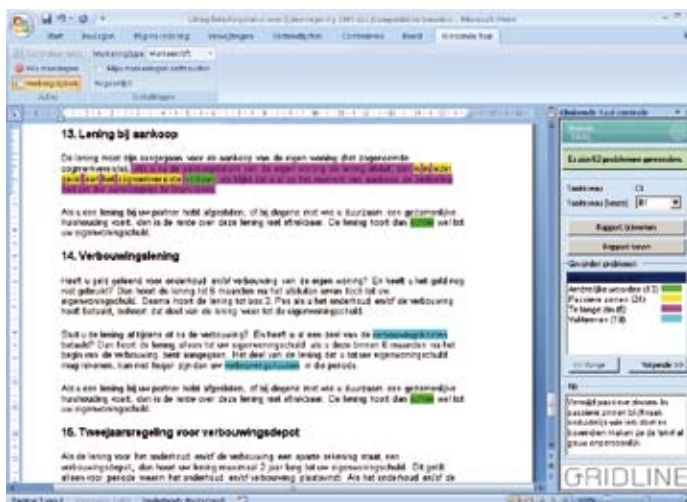
Klinkende Taal sluit aan bij een concrete overheidsbehoefte. De regering heeft namelijk bepaald dat de overheid zich moet inspannen om burgers begrijpelijk te informeren en binnen redelijke termijnen moet antwoorden op brieven, e-mails en bezwaarschriften. Deze doelstellingen zijn vastgelegd in verschillende documenten, waaronder de Webrichtlijnen-norm voor overheidswebsites, de richtlijn voor

Begrijpelijke Formulieren, de e-mailgedragslijn van burger@overheid en het manifest 'Overheid heeft Antwoord'. Het bevorderen van begrijpelijke taal staat bovendien in de Aanpak Top 10 van het Ministerie van Binnenlandse Zaken.

De Nationale ombudsman ziet scherp toe op de naleving van deze richtlijnen. In 2008 verscheen bijvoorbeeld een rapport met de titel 'Behandeling burgerbrieven kan behoorlijker'. Hierin concludeert de ombudsman dat de overheid in 2007 vooruitgang heeft geboekt met het beantwoorden van burgercorrespondentie (met name e-mails), maar dat de afhandeling van bezwaarschriften nog veel te wensen overlaat. De ombudsman heeft ook concrete brieven gecontroleerd op stijl en begrijpelijkheid. In het rapport valt te lezen hoe de onderzochte ministeries het resultaat kunnen verbeteren.

Klinkende Taal kan ambtenaren helpen bij het uitvoeren van deze richtlijnen en adviezen. Met Klinkende Taal kan een ambtenaar bijvoorbeeld snel nagaan of zijn brief het juiste taalniveau heeft. Bovendien vertelt de tool precies hoe hij de brief kan verbeteren. Dit is

Voorbeeld van een Klinkende Taalverbeteradvies bij een gecontroleerde tekst



niet alleen handig, maar werkt ook tijdbesparend. Niet onbelangrijk, gezien de eis dat de overheid binnen twee dagen antwoord moet geven op eenvoudige burgerbrieven en daarbij ook nog eens begrijpelijk moet zijn.

Klinkende Taal

Klinkende Taal beoordeelt teksten op basis van het Common European Framework of Reference (CEFR), dat teksten indeelt in niveaus van A1 (heel makkelijk) tot C2 (heel moeilijk). Voor nadere informatie, zie de sectie over het CEFR op de Taalunieversum-website (taalunieversum.org). De praktische kennis van de afdeling Communicatie van de Gemeente Den Haag vormde een goede basis voor de eerste versie van de tool. Voor het tunen van Klinkende Taal hebben we onderzoek laten doen door Henk Pander Maat van de afdeling Taalbeheersing van de Universiteit Utrecht.

In het vervolgtraject hebben we Klinkende Taal in Microsoft Office Word geïntegreerd en met nieuwe functies uitgebreid. Hierbij hebben we goed naar onze gebruikers geluisterd. Dit heeft uiteindelijk een product opgeleverd dat niet alleen het niveau van een tekst kan beoordelen, maar ook praktische aanwijzingen geeft om de tekst zo te veranderen dat de doelgroep hem begrijpt.

Taalcriteria

Klinkende Taal controleert zowel het woordgebruik als de zinsopbouw, de logische samenhang en de opmaak van een tekst. Bij de beoordeling van het woordgebruik markeert Klinkende Taal moeilijke woorden, specialistische taal, ambtelijk of ouderwets taalgebruik en abstract taalgebruik. Op zinsniveau kijkt Klinkende Taal onder meer naar zinslengte, bijzinnen, passiefconstructies, verbindingswoorden en tangconstructies. Op paragraafniveau beoordeelt Klinkende Taal de samenhang en de tekstopmaak.

Naast deze standaardcriteria kan Klinkende Taal ook aanvullende criteria meenemen, bijvoorbeeld uit de stijlwijzer van de klant. Een andere mogelijkheid is om Klinkende Taal te laten afstemmen op de eigen terminologie. Dit helpt om beter zicht te krijgen op het voorkomen van vaktermen en om de termkeuze te standaardiseren. Dit alles is mogelijk dankzij de Nederlandse taaltechnologie van GridLine.

Productvormen

Klinkende Taal is verkrijgbaar als plugin in Microsoft Office Word en als laagdrempelige web-plugin. De schrijver van een brief

of webtekst roept Klinkende Taal met één druk op de knop aan. Er verschijnt dan een beoordeling van de moeilijkheid van de tekst (bijvoorbeeld B2) en duidelijke aanwijzingen om de leesbaarheid te verhogen (bijvoorbeeld naar B1). Klinkende Taal levert ook een Quick Scan. Deze tool kan in één klap een verzameling documenten of een complete website testen. Hierbij kijkt de tool onder meer naar de begrijpelijkheid voor de doelgroep en het jargongebruik, criteria die ook onderdeel zijn van de Webrichtlijnen. Met de Quick Scan kan een redacteur snel nagaan welke documenten en webpagina's extra aandacht nodig hebben.

Klinkende Taal is eenvoudig te installeren. Alle taaltechnologie staat op de server. Na het aanroepen van Klinkende Taal legt Microsoft Office Word via een webservice verbinding met de centrale taalserver. Deze aanpak heeft als voordeel dat het heel eenvoudig is om nieuwe versies door te voeren: hiervoor is een update op de server voldoende. De webversie roept dezelfde server aan. Het beheer van woordenlijsten kan de organisatie zelf doen via een eenvoudige webmodule.

Gebruikers

Klinkende Taal is oorspronkelijk ontwikkeld voor gemeentes. De vier grootste Nederlandse Gemeentes hebben de tool als eerste in gebruik genomen. Andere overheidsdiensten zijn daarna gevolgd. Een nieuwe groep gebruikers zijn communicatie-bureaus. Zij vinden het handig om Klinkende Taal te gebruiken voor controledoeleinden, zodat de schrijvers zich kunnen richten op de inhoud en het creatieve proces. Bovendien biedt Klinkende Taal een meerwaarde voor hun klanten. Verzekerings-maatschappijen, woningcorporaties, banken en ziekenhuizen sluiten inmiddels aan bij de groeiende groep gebruikers van Klinkende Taal.

Het project Klinkende Taal werd mede gefinancierd door het STEVIN-programma. STEVIN is een meerjarig onderzoeks- en stimuleringsprogramma voor Nederlandstalige taal- en spraaktechnologie dat gezamenlijk door de Vlaamse en Nederlandse overheid wordt gefinancierd.

Instituut voor Nederlandse Lexicologie

Schatkamer van de Nederlandse taal



Het Instituut voor Nederlandse Lexicologie (INL) is een Vlaams-Nederlands instituut dat Nederlandse woorden verzamelt, bestudeert, opslaat in databases, voorziet van allerlei (taalkundige) gegevens en daarmee wetenschappelijke woordenboeken maakt. Of u nu taalkundige bent of simpelweg geïnteresseerd in taal: met al uw vragen over woorden kunt u bij ons terecht.

TST-Centrale

Vlaams-Nederlandse centrale voor beheer, onderhoud en distributie van digitale taalmaterialen

Bent u op zoek naar Nederlandstalige, digitale taalmaterialen?
Bij ons bent u aan het juiste adres!

De Centrale voor Taal- en Spraaktechnologie (TST-Centrale) stelt taalmaterialen beschikbaar die veelal met overheidsgeld zijn gefinancierd, waaronder materialen die op het INL ontwikkeld

worden en resultaten uit het STEVIN-programma (STEVIN: Spraak- en Taaltechnologische Essentiële voorzieningen voor het Nederlands). Daarnaast ondersteunt de TST-Centrale het gebruik van de materialen door uw vragen te beantwoorden en gastcolleges en workshops te organiseren.

De TST-Centrale is een initiatief van en wordt gefinancierd door de Nederlandse Taalunie. De TST-Centrale is ondergebracht bij het Instituut voor Nederlandse Lexicologie met vestigingen in Leiden en Antwerpen.

Beschikbare materialen:

- Historische en wetenschappelijke elektronische woordenboeken (o.a. Woordenboek der Nederlandsche Taal online)
- Gesproken, geschreven en multimediale corpora (o.a. Corpus Gesproken Nederlands)
- Mono- en bilinguale lexica (o.a. Referentiebestand (Belgisch-)Nederlands)
- Tools voor gesproken en geschreven teksten (o.a. AUTONOMATA-g2p-toolkit)

www.inl.nl
www.inl.nl/tst-centrale
www.inl.nl/producten

servicedesk@inl.nl

))) **nl | TST-Centrale**

“Géén momentje geduld alstublieft!”

"Er zijn nog 7 wachtenden voor u.", "Heeft u een vraag over onze aanbiedingen, toets dan 7, heeft u een vraag over uw rekening, toets 8, voor overige vragen, toets een 9. Een momentje geduld alstublieft." Wie kent ze niet, de ergerlijke telefoonbeantwoorders, die je simpelweg verwijzen naar een website of minutenlang – vaak tegen een stevig tarief – laten wachten met foute muziek, en je uiteindelijk doorverbinden met iemand die je niet verder kan helpen.

Dit soort 'kluitje-in-het-riet'-telefoontjes behoren binnenkort tot het verleden. De Delftse bedrijven Irion en Dutcheer hebben een prototype gebouwd, dat is gebaseerd op een spectaculaire en ingenieuze combinatie van taaltechnologie en spraakherkenning. Dit systeem – GemeenteConnect genaamd – levert de bellende burger meteen boter bij de vis: een bruikbaar antwoord op een concrete vraag, of automatisch doorverbinden met de juiste persoon bij de gemeente.

Wat is de clou?

Het idee is eigenlijk simpel: mensen hebben geen volledige, correct geformuleerde zinnen nodig om een betekenis of bedoeling te begrijpen. Sterker nog, mondelinge communicatie verloopt zelden grammaticaal correct. Denk aan zinnetjes als "al in je nieuwe tent geslapen?" of "jij koffie?". De zinnen zijn onvolledig, maar worden toch moeiteloos begrepen. Of denk eens aan een telefoongesprek via een belabberde, krakende verbinding ("sorry, ik zit in een tunneltje!"). Ook al versta je misschien maar een paar flarden, die flarden blijken vaak toch voldoende om de kern van de boodschap op te pikken. Dat komt doordat mensen in staat zijn deze flarden snel in een juiste context te plaatsen.

Wanneer een ontvanger al ongeveer weet binnen welke context hij iets moet begrijpen, is het interpreteren nóg gemakkelijker. Als u in een kamer zit, en iemand steekt zijn hoofd om de deur met de vraag "Koffie?", is de bedoeling duidelijk. Was de vraag "Krokodil?", dan zou u vermoedelijk niet begrijpend het hoofd schudden. In dit laatste voorbeeld zit de basisgedachte van het systeem opgesloten: als een beller maar een handjevol onderwerpen kan bedoelen, wordt het achterhalen van die bedoeling al snel een flink stuk makkelijker. Je belt een gemeente bijvoorbeeld niet voor een frietje satésaus of een huurauto. In een duidelijke context heeft een goed verstander maar een half woord nodig. En die goed verstander mag best een computer zijn.

Proefkonijn

De resultaten: in de 'proefkonijn'-gemeente Gilze-Rijen konden maar liefst 80% van de veel voorkomende vragen binnen 2 minuten door het systeem correct worden afgehandeld! Zelfs een test met Turkse inwoners leverde vergelijkbare resultaten op. De voordelen zijn duidelijk: bellers worden zonder wachttijden te woord gestaan, krijgen zonder op allerlei toetsen op hun toestel te hoeven drukken meteen een 'digitale deskundige' aan de lijn, en kunnen 24 uur per dag, 7 dagen in de week telefonisch met hun vragen bij de gemeente terecht! Ook voor de gemeente zelf biedt dit grote voordelen: er blijft meer tijd over om de complexere vragen af te handelen, de bereikbaarheid is drastisch verbeterd, en in veel gevallen zullen ook de kosten omlaag kunnen.

GemeenteConnect is een demonstratieproject uit de eerste ronde van het STEVIN-programma (zie taalunie-versum.org/taal/technologie/stevin/projecten/#gemeenteconnect)

Joop van Gent
Irion



Henk van den Heuvel
Centre for Language and Speech Technology (CLST)

Duurzame opslag van en langetermijntoegang tot onderzoeksdata: de missie van DANS

Onder burgers en professoren. De toegankelijke overheid.

De overheid produceert massa's aan informatie met publieke gelden. Een groot deel van die informatie is nuttig voor de burger, maar de toegang ertoe ontbreekt of komt slechts moeizaam tot stand. Althans, tot voor kort.

Dankzij taaltechnologie is het mogelijk om te zoeken in grote bossen informatie waar de arme burgergebruiker zonder deugdelijke hakbijlen hopeloos in verdwaalt. Een voorbeeld is het gemeentearchief van de gemeente Antwerpen dat als Digitaal Stadsarchief te doorzoeken is via de website www.felixarchief.be. De gebruiker kan via verschillende manieren in de archieven zoeken. Dat kan op woorden in de tekst maar ook op zogenaamde metadata, bijvoorbeeld: titel, taal, of rubrieken (zoals bijvoorbeeld huizen, families, straten). De gemeente Den Bosch biedt een vergelijkbare dienst, zie www.stadsarchief.nl.

Andere data die grotendeels door overheidsfinanciering tot stand komen zijn onderzoeksbestanden en wetenschappelijke publicaties. Een instelling die zich toelegt op het archiveren en doorzoekbaar maken van onderzoeksdata is DANS. Zie het artikel hiernaast.

Een recent initiatief dat van belang is voor wetenschappers is CLARIN-NL. Dit project streeft naar de vorming van een infrastructuur waarin onderzoeksdata (waaronder grote taal- en spraakdatabases) op een eenduidige en gestandaardiseerde manier gearhiveerd, beheerd en doorzocht kunnen worden. Voor meer informatie, zie www.clarin.nl.

Uiteindelijk moeten deze ontwikkelingen het gemak van de burger dienen, of deze nu wetenschapper is of niet: toegang tot relevante overheidsgegevens, 24 uur per dag, elke dag opnieuw, snel en accuraat!

Je staat er niet bij stil, maar de snelle vernieuwingen in de communicatie- en informatietechnologie, waarvan uiteraard ook de wetenschap profiteert, veroorzaken ook risico's die de vooruitgang van de wetenschap bedreigen. Databestanden kunnen ontoegankelijk worden, omdat het bestandsformaat uit de gratie raakt. Gegevens kunnen onbruikbaar worden, omdat niet duidelijk is of ze nog betrouwbaar zijn. Bestanden kunnen zo maar verdwijnen, omdat er geen plek is waar ze bewaard worden. Zeker omdat de creatie van onderzoeksdata kostbaar is, loont het de moeite om aandacht te besteden aan de duurzame archivering van datasets. Het toekomstig hergebruik van de data moet mogelijk blijven.

Te midden van diezelfde snelle vernieuwingen in de communicatie- en informatietechnologie is het noodzakelijk om goed na te denken welke inzichten en gereedschappen nodig zijn om investeringen in de wetenschappelijke data-infrastructuur voor de toekomst te behouden. DANS (Data Archiving and Networked Services), een initiatief van NWO en KNAW, streeft naar een duurzame opslag en blijvende toegankelijkheid van onderzoeksgegevens in de alfa- en gammawetenschappen. Samenwerking met onderzoekers, financiers van onderzoek, wetenschappelijke instellingen en het 'traditionele' archief- en bibliotheekwezen staat hierbij voorop.

De implementatie van duurzame opslag van en toegang tot onderzoeksdata dient volgens DANS te voldoen aan vijf kwaliteitscriteria. Op de eerste plaats dienen de gegevens vindbaar te zijn op het internet, ten tweede dienen ze toegankelijk te zijn (uiteraard rekening houdend met de wetgeving met betrekking tot persoonsgegevens en het intellectueel eigendom van de gegevens). Op de derde plaats dienen ze in een bruikbaar dataformaat beschikbaar te zijn. Vervolgens dienen de onderzoeksgegevens betrouwbaar te zijn. Het vijfde kwaliteitscriterium stelt dat onderzoeksdata refereerbaar moeten zijn. DANS heeft het initiatief genomen om bovenstaande criteria te concretiseren in een aantal richtlijnen die onderdeel zijn van het data seal of ap-

proval (www.dataealofapproval.org). Dit datakeurmerk wordt onderhouden door een internationale redactie, bestaande uit vertegenwoordigers van organisaties die actief zijn op het gebied van de wetenschappelijke data-infrastructuur. Het is de bedoeling dat in de toekomst op basis van het datakeurmerk diensten en systemen gecertificeerd gaan worden.

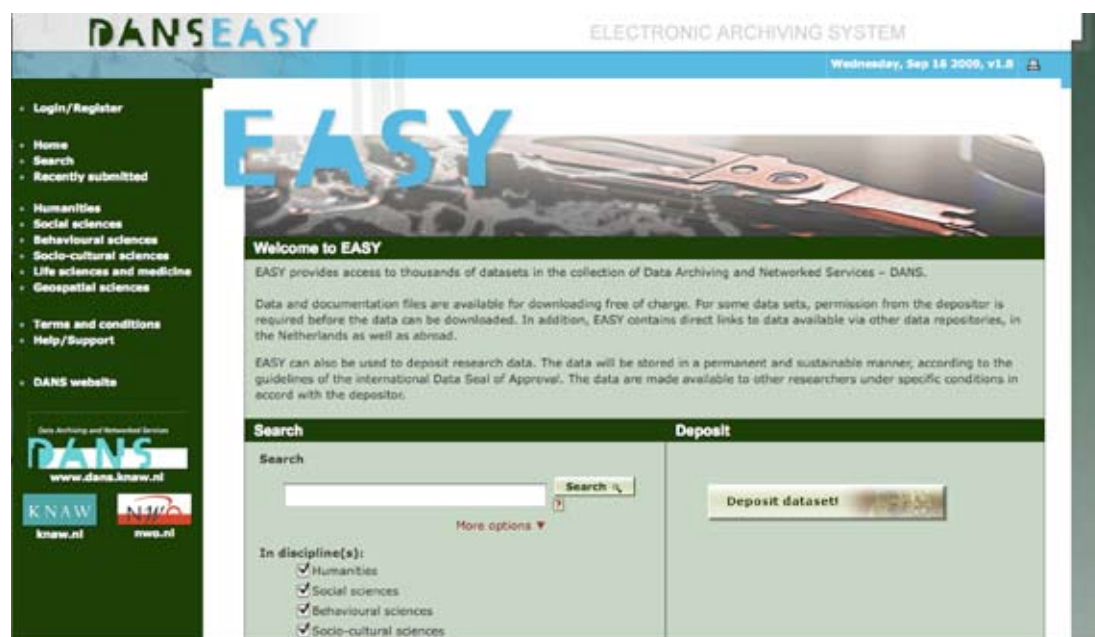
Hieronder enkele voorbeelden die aansluiten bij de bovengenoemde kwaliteitscriteria. Om de onderzoeksgegevens te kunnen vinden op internet is het data-archiveringssysteem EASY ontwikkeld (easy.dans.knaw.nl). Met EASY kan iedere onderzoeker zelfstandig datasets deponeren. DANS propageert zo veel mogelijk dat toegang tot onderzoeksdata Open Access is. Met betrekking tot de duurzaamheid van het dataformaat sluit DANS aan bij initiatieven op het gebied van preferred formats. Onder deze noemer valt ook het MIXED-project (mixed.dans.knaw.nl) waarin een webservice is ontwikkeld om verschillende soorten bestanden om te zetten naar een duurzaam geacht XML-dataformaat.

De betrouwbaarheid van datasets wordt vergroot door mechanismen toe te passen die de authenticiteit garanderen. Een goed voorbeeld hiervan is een systeem om een persistent identifier te koppelen aan een digitaal object, waardoor dit steeds onder dezelfde naam geïdentificeerd en gevonden kan worden, ook als het object wordt verplaatst naar een andere locatie. DANS werkt samen met andere belanghebbenden aan een infrastructuur voor duurzame identificatie en aan initiatieven waarin de persistent

identifier gebruikt wordt om naar datasets te refereren en om eruit te citeren.

Het belangrijkste DANS-initiatief op het gebied van de taal- en spraaktechnologie betreft het participeren in activiteiten die ontstaan in het kader van de CLARIN-onderzoeksinfrastructuur (www.clarin.nl). Inmiddels vinden veel lopende en nieuwe projecten op taal- en spraaktechnologisch gebied waar DANS in deelneemt hun weg naar CLARIN. Een voorbeeld hiervan is het project Veteran Tapes (www.dans.knaw.nl/nl/projectenpagina/veteran_tapes).

René van Horik
Data Archiving
and Networked
Services (DANS)



Toegankelijker websites door TST

Binnen de e-overheid is toegankelijkheid van websites voor de gehele bevolking een belangrijk politiek thema geworden. In een overleg met de Tweede Kamer heeft staatssecretaris Ank Bijleveld van Binnenlandse Zaken hierover gezegd: "Gemeenten, provincies en ministeries lopen ver achter wat betreft de toegankelijkheid van hun websites. Zij moeten hun sites uiterlijk eind volgend jaar wel helemaal op orde hebben...."

De richtlijnen voor toegankelijke websites houden onder meer in dat websites ook voor visueel gehandicapten goed toegankelijk zijn, dat ze vanaf mobiele telefoons te bereiken zijn en dat informatie makkelijk is te vinden.

Mede als gevolg van deze politieke druk worden steeds meer overheidswebsites voorzien van een voorleesfunctie waarmee synthetische spraak wordt ingezet om de inhoud van de pagina voor te lezen. De toegevoegde waarde van deze aanpak voor visueel gehandicapten is beperkt in die gevallen waar niet is nagedacht over het gebruik door de doelgroep. Veelal moet met een muisklik op een knop op het scherm de voorleesfunctie geactiveerd worden(!). Andere functies, zoals navigeren en invullen van velden worden niet ondersteund met TST. Koppelingen worden voorgelezen waarbij met een muisklik genavigeerd moet worden en invulvelden kunnen uitsluitend via het toetsenbord ingevuld worden.

Toegankelijkheid zoals bedoeld in bovenstaand citaat gaat verder dan de richtlijnen van W3C (www.w3.org/WAI/guid-tech.html) met betrekking tot toegankelijkheid van websites en het aanbrengen van een voorleesfunctie op de website.

Door TST in nieuwe multimodale en multi-channel technieken toe te passen kan deze problematiek in grote mate worden opgelost. Deze benadering is het best te realiseren in combinatie met het implementeren van de richtlijnen van W3C, die zich wel richten op het zeker stellen van de beschikbaarheid en de structurering van de informatie maar niet op de wijze van toegang tot deze informatie. Het is sinds kort mogelijk om webapplicaties en spraakapplicaties te koppelen. Het bieden van telefonische toegang tot websites is in de e-overheid onder meer succesvol ingezet bij het Politiekorps Rijnmond en UWV. Deze technologie is verder uitgebreid om multimodale toegang middels een multimediacomputer te bieden. Dit maakt spraakgestuurd browsen en communicatie in twee richtingen mogelijk waarbij met spraak invoervelden in de website kunnen worden ingevuld. Een bijkomend voordeel van deze aanpak is dat de overheid de functionaliteit aan het publiek ter beschikking kan stellen als dienst, waardoor gehandicapten niet op individuele basis oplossingen hoeven aan te schaffen en te onderhouden.

De nieuwste functionaliteit wordt momenteel ontwikkeld binnen het MIV-project (www.mivproject.nl). In dit project uit het Europese stimuleringsprogramma OP-Zuid wordt multimodale werking van smartphones gerealiseerd. Dit betekent dat de gebruiker vanaf een handheld device naar keuze de website met spraak en/of met de gebruikelijke invoermethodes (muis, keyboard) kan bedienen. Een nieuw platform voor mobiele applicaties waarmee gebruikers eenvoudig met spraak kunnen communiceren met mobiele webapplicaties voorziet in een behoefte van consumenten.

**Its Kievits
Dialogs
Unlimited BV**



NOTaS-agenda

Wat: Digitaal Bestuur e-overheid
 Wanneer: 26 november 2009
 Waar: Orpheus Theater Apeldoorn
 Meer weten: www.digitaalbestuurcongres.nl

Wat: Symposium Expertise, ambitie en toewijding: bindende factoren?
 Wanneer: 3 december
 Waar: Ede
 Meer weten: www2.viataal.nl

Wat: CLIN 2010 Computational Linguistics In The Netherlands
 Wanneer: 5 februari 2010
 Waar: Universiteit van Utrecht
 Meer weten: www.clin.nl/20

Heeft u ook een evenement dat u via de DIXIT of de NOTaS-website wilt publiceren? Stuur deze informatie dan naar webmaster@notas.nl

Deze DIXIT wordt u aangeboden door de deelnemers van NOTaS:

- o CLST
- o Data Archiving and Networked Services (DANS)
- o Dedicon
- o Dialogs Unlimited B.V.
- o Dutchear
- o GridLine B.V.
- o inTAAL B.V.
- o Knowledge Concepts
- o Logica
- o OSTT/ St. Maartenskliniek
- o PonTeM
- o Q-go.com B.V.
- o TeleCats
- o TNO Defensie en Veiligheid
- o TST-Centrale, Instituut voor Nederlandse Lexicologie
- o Universiteit van Tilburg - CIW
- o Universiteit Twente - HMI
- o Universiteit Utrecht - UIL OTS
- o Van Dale Uitgevers
- o VoiceCorp Webservices B.V.

En is in het bijzonder gemaakt door:
 het STEVIN-programma en de Nederlandse Taalunie



Yournews: automatisch uw eigen nieuws

Kort geleden zei iemand tegen me dat een mens tegenwoordig op één dag evenveel informatie moet verwerken als honderd jaar geleden in een heel jaar. Of dat waar is weet ik niet, maar zeker is dat de opkomst en kracht van de nieuwe media ervoor hebben gezorgd dat de tijd dat we voor het nieuws moesten wachten tot de krant op de mat viel voorgoed voorbij is.

**Joop van Gent
Irion**



Met de nieuwe vormen van nieuwsvoorziening is ook het aanbod exponentieel gegroeid. Het nieuws wordt ons door alle poriën naar binnen geduwd. We kunnen moeiteloos en gratis de digitale versies van honderden kranten op onze PC krijgen, of zelfs op ons mobieltje. Maar naarmate de stroom van informatie aanzwelt, groeit net als bij de rivier ook de behoefte aan indamming.

We verliezen het overzicht en worden soms wanhopig van de overstelpende hoeveelheid informatie. Dat geldt zeker ook voor ambtenaren bij de overheid die nieuwsberichten verwerken voor bijvoorbeeld knipselkranten voor hun departementen. Een geslaagde poging van het op de juiste plaats en tijd in de juiste hoeveelheid aanbieden van informatie is Yournews.

Het STEVIN-project Easyinfo

Yournews is een nieuwe nieuwssite, gebaseerd op de concepten uit het Stevin-project EasyInfo (taalunieversum.org/taal/technologie/stevin/projecten/#easyinfo). Dit project, uitgevoerd door twee Nederlandse taaltechnologiebedrijven, Irion en Carp, in samenwerking met MD Info, een multimediale aanbieder van gepersonaliseerde zakelijke informatie, had tot doel het nut van taaltechnologie te bewijzen voor selectieve nieuwsvoorziening in een werkende applicatie.

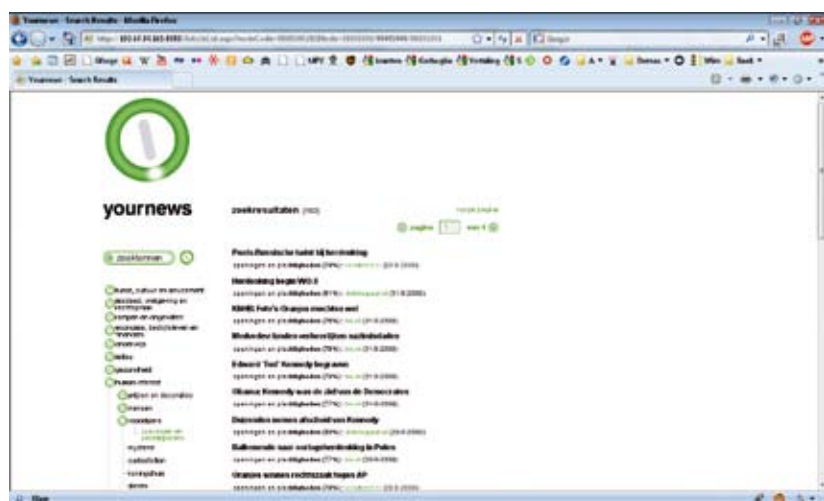
Yournews

Yournews sorteert nieuws en stemt krantenberichten af op het behoefteprofiel van de gebruiker. Bovendien vat Yournews de nieuwsberichten ook nog eens samen. Het hele proces is automatisch, en abonnees krijgen precies het pakketje dat voor hen interessant is. Zo kan het dagelijks nieuws uit honderden, zo mogelijk duizenden bronnen op een veel efficiëntere manier bij de mens worden gebracht dan tot nog toe mogelijk was.

Bestond dit dan niet allang? Wel, ja en nee. Online nieuwsdiensten zijn inderdaad niets nieuws, er zijn er duizenden van. Maar bij deze diensten wordt het nieuws doorgaans met de hand uitgesorteerd, en worden de behoefteprofielen voor klanten ook met de hand gemaakt. En dat is meteen ook hun achilleshiel. Want dergelijk arbeidsintensief handwerk heeft alleen zin voor gebruikers die daarvoor goed willen betalen. Een alternatief model waarbij de inkomsten niet uit betalende klanten, maar uit advertentieinkomsten komen, levert onvoldoende op om het handwerk te bekostigen. Dat is dan ook de reden dat de meeste van deze diensten zich richten op de zakelijke markt. Voor veel overheidsinstanties is abonneren niet aantrekkelijk, temeer omdat ook de profielen meer zijn toegespitst op de zakelijke markt, en in veel mindere mate op de maatschappelijke thema's.

Kostenbesparing

Doordat Yournews gebruik maakt van taaltechnologie, een technologie die het mogelijk maakt om tekstuele informatie automatisch door een computer te laten interpreteren, komt op maat gesneden nieuwsvoorziening beschikbaar voor een veel groter publiek. De technologie betekent een dermate grote kostenbesparing, dat Nederlandstalige nieuwsberichten niet alleen voor de zakelijke markt, maar ook voor burgers en overheid gemakkelijk binnen handbereik komen. Naar verwachting zal Yournews.nl aan het eind van dit jaar worden gelanceerd.



TST voor een democratische en klantgerichte dienstverlening

TST is inmiddels zo ver dat klanten 24 uur per dag toegang kunnen hebben tot bedrijven en instellingen. Afspraken maken of verzetten dan wel informatie opvragen, kan de klok rond. Deze toegankelijkheid klinkt eenvoudig, snel en ook klantgericht. Maar hoe weten instellingen en bedrijven nu werkelijk of ze klantgericht zijn en wat kan TST daar nog in betekenen?

Bij een publieke dienstverlener in Nederland is een pilot gehouden onder 700 medewerkers en hun klanten. Het doel van de pilot was om op individueel niveau feedback te genereren van de klant over de dienstverlening, zodat er van geleerd kon worden. Zo kon elke klant een vragenlijst anoniem inspreken middels TST. Hij kon dit ook doen via een email die linkte naar een website. Als een klant bijvoorbeeld gebeld had met een medewerker, dan kon deze klant direct worden doorgeschakeld met een telefonische enquête. De klant kon zijn waardering in cijfers inspreken en kon ook een persoonlijke tekst achterlaten voor de medewerker. De medewerker kon de gegeven antwoorden vervolgens direct terugzien op een persoonlijk dashboard op zijn computerscherm zodat hij er van kon leren.

Al deze 'data' die zijn ingesproken, werden opgeteld, automatisch samengevat en vervolgens teruggekoppeld aan twee hogere echelons. Zo kon er op drie organisatieniveaus van geleerd worden, terwijl het voor de klant een minimale inspanning was. Deze hoefde alleen maar zijn mening in te spreken of in te vullen op een website, en dat kostte niet meer dan drie minuten. Ook kunnen klachten op deze manier gemakkelijk worden voorkomen. Als de klant dat wil, dan kan hij namelijk worden teruggebeld. Dat kan een klacht, met alle imagoschade van dien, voorkomen.

Burger-feedbackplatform

Op deze manier zou de gehele (publieke) dienstverlening zich kunnen openstellen voor feedback van klanten. De burger kan aangeven wat hij vindt van de dienstverlening en er op vertrouwen dat zijn feedback echt aankomt bij de individuele medewerker, inclusief de complimenten. Als alle publieke dienstverleners hieraan mee gaan doen, kunnen er interessante benchmarks ontstaan en kunnen de dienstverleners zich ook vanuit het perspectief van de burger verantwoorden aan de maatschappij. Dit noemen we een Burger-feedbackplatform. Doordat de burger zijn stem op deze manier weer terugkrijgt in de uitvoering van publieke taken, ontstaat

er weer een waardevolle dialoog en kan er zelfs gesproken worden van het democratischer maken van de dienstverlening. Door het inzetten van TST worden twee belangrijke winstpunten geboekt:

- De toegankelijkheid van het platform is maximaal door het gebruik van zowel internet als telefonie, en door procesautomatisering van grote hoeveelheden transacties. Hierdoor kan bij ieder contactmoment de mogelijkheid geboden worden om feedback te geven, ook voor mensen met beperkingen en laaggeletterden.
- Door automatisering van het feedbackproces, de rapportage en de managementstructuur kan een Burger-feedbackplatform voor een haalbaar kostenniveau gerealiseerd worden.

Het is de opdracht van de dienstverlener om daadwerkelijk lering te trekken uit de gegeven feedback. Alleen als het permanent als input wordt gebruikt om van te leren, en als iedereen de mogelijkheid heeft feedback te geven, heeft het waarde voor de dialoog tussen burger en dienstverlener.

Reflexie



Vincent de Klerk
Comm & Sens

DuOMAn: Dutch Language Online Media Analysis

Het STEVIN-project DuOMAn is gericht op het ontwikkelen van taaltechnologie ter ondersteuning van media-analisten, die zich richten op onderzoek naar de mediareputatie van organisaties. Een belangrijk doel daarbij is inzichtelijk maken hoe er door media, journalisten en in sociale media over hen gesproken wordt.

Valentin Jijkoun
en
Maarten de Rijke
Universiteit van
Amsterdam

De technologie die binnen DuOMAn ontwikkeld wordt kan ook ingezet worden bij informatievoorziening ten behoeve van de overheid enerzijds - bijvoorbeeld om antwoord te krijgen op de vraag welke meningen er over een gegeven onderwerp leven onder de bevolking - en ten behoeve van de burgers anderzijds - bijvoorbeeld om inzicht te krijgen in de belanghebbenden rondom een onderwerp.

Waar het gaat om overheidsinformatie komt er steeds meer online materiaal, zowel geredigeerd, in de vorm van bijvoorbeeld persberichten, officiële websites, interviews of krantenberichten, als user generated in de vorm van blogs, tweets of bijdragen aan discussiefora van politici, vertegenwoordigers van de overheid en burgers. Analyses van dergelijk materiaal zijn nuttig in vele scenario's waarin overheid en burger elkaar treffen.

Hieronder noemen we er enkele.

- Een overheidsorganisatie wil graag een overzicht van de belangrijkste standpunten en argumenten rondom een gegeven onderwerp, bijvoorbeeld 'fietsverlichting' of 'OV-chipkaart'.
- Een burger wil achterhalen wat de recente discussies in het parlement waren rondom bijvoorbeeld de kosten van immigratie, en wil daarbij inzicht krijgen in de belangrijkste experts en belanghebbenden (plus hun standpunten).
- Het voorspellen van de mogelijke reacties en impact van beslissingen rondom gevoelige onderwerpen ('het verhogen van de pensioengerechtigde leeftijd').
- Het meten van publieke opinies in sociale media is duidelijk relevant voor politieke partijen, de overheid en de burger.

Binnen het DuOMAn-project wordt voornamelijk gewerkt aan technologie waarmee de

Peilend.nl: de publieke demonstrator van het DuOMAn-project

The screenshot shows the Peilend.nl website interface. At the top, the header reads 'Peilend.nl' with the tagline 'Analyseer wat er speelt rond het nieuws'. Below the header is a search bar with the text 'afgelopen 30 dagen' and 'amsterdam chipkaart'. To the right of the search bar are buttons for 'Zoeken' and 'Opties'. Below the search bar, a message states: 'Gezocht in de periode 08/15/2009 tot en met 09/15/2009 naar amsterdam chipkaart, en niet in nulij'.

The main content area displays a list of search results. Each result includes a date, a source icon, and a snippet of the article text. The results are sorted by date, with the most recent at the top. The snippets mention topics like 'D-day voor ov-chip in Amsterdamse metro', 'Invoeren chipkaart Amsterdam verloopt soepel', 'Alleen nog OV-chip in metro Amsterdam', 'Rover tevreden over einde strippenkaart Amsterdam', 'Invoering ov-chipkaart in Amsterdam verloopt soepel', 'GVB: invoering chipkaart gaat soepel', 'Rover positief over invoering ov-chipkaart', 'Amsterdamse metro stapt over op ov-chipkaart', 'GVB: invoering chipkaart gaat soepel', and 'Einde aan strippenkaart Amsterdamse metro'.

On the right side of the page, there is a 'Woordenwolk' (word cloud) section. It contains a list of words and phrases related to the search results, such as 'Amsterdam', 'GVB', 'OV-Chipkaart', 'Rb', 'Rotterdam', 'VS Volgens aantal afschaffen alle', 'alleen apparaat automaten', 'basistarief bijna bus', 'chipkaart', 'dag druk eerste', 'enkele extra gaat goed goedkoper', 'grote honderd Inchecken ingezet inz', 'jaar januari juli kaart', 'keer komen', 'kopen kost kwartaal laat lange maanden', 'maken medewerkers mensen', 'metro', 'mobiele mogelijk', 'ondergrondse openbaar', 'openbaarvervoerbedrijf OV', 'passagiers poortje poortjes', 'problemen procent reis', 'reizen reizigers retourtje retourtjes rijen', 'soepel stations', 'steden stonden', 'storingen', and 'strippenkaart'.

At the bottom of the page, there is a pagination bar showing 'Pagina 1 van 1' and a 'Feed' button. The bottom right corner indicates 'Toon onderwerpen 1 t/m 39 van 39'.

eerste twee scenario's ondersteund kunnen worden. (In het door NWO gefinancierde project Tracking News Events and their Impact (TNT), dat binnen dezelfde onderzoeksgroep wordt uitgevoerd als DuOMAN, ligt de nadruk juist op het derde scenario.)

Binnen DuOMAN kunnen we verschillende ontwikkelingslijnen onderscheiden die van belang zijn waar het gaat om analysemiddelen voor zowel overheid als burgers. Om te beginnen werkt DuOMAN aan het verzamelen en linken van geredigeerd en user generated materiaal. Voor deze links wordt gebruik gemaakt van named-entity-herkenning en named-entity-normalisatie, waarmee belangrijke belanghebbenden kunnen worden geïdentificeerd. Daarnaast kunnen sentimenten en opinies geanalyseerd en geaggregeerd worden. Op de linker pagina is de publieke demonstrator van het DuOMAN-project te zien, met daarin het resultaatsscherm voor de zoekvraag 'Amsterdam chipkaart', werkend op artikelen vergaard van nieuwssites. In dit scherm zien we links een overzicht van de resultaten, met de duplicaten gegroepeerd, met vermelding van het aantal reacties waartoe een bericht aanleiding heeft gegeven. Rechts is een kleur-gecodeerde woordenwolk waarin de kleuren overeenkomen met verschillende soorten van named entities (personen, organisaties, etc.).

Na doorklikken op een zoekresultaat heeft men toegang tot (een samenvatting van) het desbetreffende nieuwsartikel, met een woordenwolk waarin de voornaamste entiteiten in het artikel worden getoond. In de nabije toekomst zullen hier entiteiten en opinies te zien

zijn welke in de tekst gemarkeerd worden. Naast peilend.nl omvatten de te verwachten resultaten van het DuOMAN-project verschillende taalkundige bronnen (lexicons, geannoteerde corpora), software en taaltechnologische webservices die beschikbaar zullen zijn voor zowel commercieel als niet-commercieel gebruik. In peilend.nl maakt DuOMAN gebruik van online beschikbare nieuwsbronnen, discussiefora en binnenkort ook blogs. Hiermee kunnen aan zowel media-analisten als het brede publiek zeer gerichte publieke opinies geleverd worden over mensen, producten en onderwerpen.

DuOMAN is een samenwerkingsverband tussen de Universiteit van Amsterdam, Trend-Light Netherlands B.V., GridLine B.V., de Rijksuniversiteit Groningen en de Hogeschool Gent. Het project wordt gefinancierd door het STEVIN-programma. STEVIN is een meerjarig onderzoeks- en stimuleringsprogramma voor Nederlandstalige taal- en spraaktechnologie dat gezamenlijk door de Vlaamse en Nederlandse overheid wordt gefinancierd.



Gedetailleerde blik op een zoekresultaat

Stijgend bezoek www.notas.nl

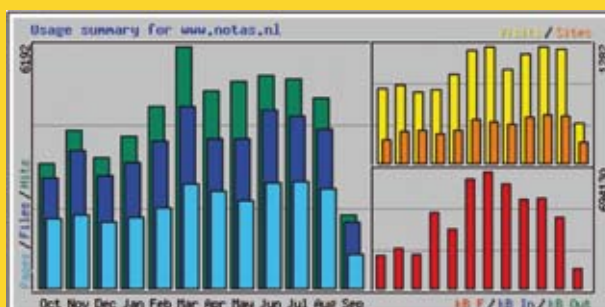
Zowel mens als machine hebben hun weg naar de NOTaS-website weten te vinden, gezien het stijgend aantal bezoekers. Via de vernieuwde website kunnen deelnemers van NOTaS zelf nieuws posten, deelnemerspagina's bijhouden, bestanden uploaden en vacatures plaatsen.

Op de korte termijn vindt klein onderhoud plaats aan de website van NOTaS. Zo zullen bijvoorbeeld de layout en inhoud worden opgefrist. Op de iets langere termijn wordt onderzocht of de deelnemers additionele wensen hebben voor interactieve componenten op de website.

Naast de website van NOTaS is sinds augustus 2009 ook een NOTaS-groep op de sociale netwerksite LinkedIn (www.linkedin.com) beschikbaar. Via deze

groep kunnen NOTaS-deelnemers, belanghebbenden en andere geïnteresseerden contact houden en ideeën uitwisselen.

U wordt van harte uitgenodigd feedback te geven op de NOTaS-website en deel te nemen aan de NOTaS-groep op LinkedIn.



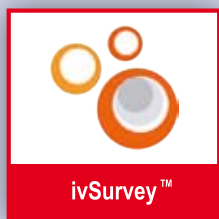
ImmediateDialogs



Een revolutie in web- en spraakintegratie

Dialogs Unlimited presenteert een nieuwe mogelijkheid om volledig geïntegreerde multi-channel (web en telefonie) self service diensten te ontwikkelen en aan te bieden, gebaseerd op het ImmediateDialogs service delivery platform. Met ImmediateDialogs kan ieder business proces voorzien worden van een geïntegreerde web- en spraakapplicatie, of een bestaande webapplicatie kan worden voorzien van een spraakapplicatie met volledig hergebruik van de back-office koppeling. Hierdoor kan in ongekend korte tijd en voor zeer lage onderhoudskosten multi-channel self service geïntroduceerd worden.

Voorbeelden van ImmediateDialogs diensten zijn:



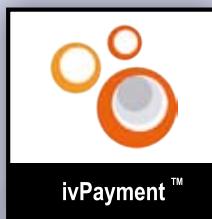
ivSurvey™, een end-to-end oplossing voor surveys, zoals klanttevredenheids-onderzoek of opiniepeilingen. De vragenlijsten worden zowel telefonisch als via internet eenvoudig ingebracht en geheel automatisch afgenomen en verwerkt. Een hoogwaardige implementatie is gedaan bij een overheidsinstelling, waar een klanttevredenheidmeting

is gecombineerd met rechtstreekse terugkoppeling naar medewerkers inclusief een leer- en verbetermethodiek.

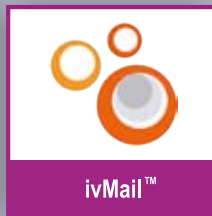


ivHelpdesk™, een multi-channel self service oplossing voor het doen van helpdeskmeldingen opvragen van status ongeacht tijd en plaats. ivHelpdesk kan de operatorfunctie van de helpdesk automatiseren, waarmee grote besparingen gerealiseerd kunnen worden. Een succesvolle implementatie is gedaan bij een overheidsinstelling, waarbij niet

alleen 70% van alle events met self service worden afgehandeld maar tegelijkertijd de gebruikertevredenheid is toegenomen.



ivPayment™ is de oplossing om de op de website aanwezige betaalprocessen telefonisch te ondersteunen.



ivMail™ is een volledig multi-channel messaging systeem waarbij gebruikers zowel per telefoon als via het web hun berichten kunnen bekijken/beluisteren, managen en opvolgen.



ivAnywebsite™ is de oplossing om iedere website van een spraakkanaal te voorzien. Veel websites hebben self service functies geïmplementeerd voor internetgebruikers. Met ImmediateVoice™ kunnen deze websites worden geconverteerd naar spraakapplicaties waar klanten naar toe kunnen bellen per telefoon of per multi-media PC.



Content management kan op eenvoudige en economische wijze via ImmediateDialogs gedaan worden. ImmediateDialogs diensten kunnen gehost worden waardoor geen investering in systemen nodig is. Zo kunnen tegen ongekend lage kosten multi channel self service diensten worden aangeboden.

Dialogs Unlimited is uw partner voor:

- het verbeteren van klanttevredenheidsprocessen,
- het efficiënt afhandelen van piekbelasting of
- het realiseren van productiviteitsverbetering bij het afhandelen van routinematige calls

Voor meer informatie:

Dialogs Unlimited

Dialogs Unlimited BV
Takkebijsters 17 sub 11
4817 BL Breda
T 076 - 57 24 777 - F 076 - 57 12 022
E info@dialogsunlimited.com
I www.dialogsunlimited.nl

TST voor de Rechtspraak, de politie en het NFI

Net als de rest van de maatschappij, ontkomen ook Justitie en politiediensten niet aan de algemene tendens om sneller en efficiënter met de beschikbare geluidsgegevens om te gaan. Om iets met de geluidsopnamen te kunnen doen, is het noodzakelijk dat de opnamen voorzien worden van beschrijvende, tijdsynchrone data (metadata) die iets vertellen over de opnamen op een bepaald moment. Nog mooier zou het zijn wanneer de spraak in zo'n opname automatisch omgezet kan worden in tekst zodat je een vorm van ondertiteling krijgt.

Naarmate de TST beter gaat werken, komt dit ideale einddoel beter in zicht. Bovendien is het zo, dat de mogelijkheid om spraak automatisch om te zetten in tekst, nieuwe toepassingen mogelijk maakt om data met spraak in te voeren.

Politie en Justitie

Het gesproken woord vormt nog steeds een van de belangrijkste bronnen van informatie voor Justitie. Vooral voor de politie en rechtbanken vormt het mondelinge verhoor de kern van het proces van waarheidsvinding. Maar ook op andere niveaus binnen deze organisaties is het gesproken woord een reële bron van informatie en dus is het zinvol te inventariseren waar TST, en dan vooral spraakherkenning, een zinvolle bijdrage kan leveren aan het ontsluiten van deze informatie.

Rechtbanken

In het STEVIN-demonstratieproject RechtSpraakHerkenning (2008-2009) is onderzocht in hoeverre het herkennen van alle spraak tijdens de rechtszitting een zinvolle bijdrage kan leveren in het sneller en beter verwerken van de rechtszitting. Hoewel de evaluatie van het project op dit moment nog loopt, is duidelijk dat er een aantal significante voordelen zitten aan het gebruik van TST. Tijdens de rechtszitting wordt de spraak van iedere spreker op een apart kanaal opgenomen waardoor precies duidelijk is wie wanneer aan het woord is. Vervolgens wordt alle spraak herkend en omgezet in een tijdsynchrone tekst zodat ook duidelijk is wanneer wat gezegd werd. Omdat per kanaal bekend is wie er sprak, kan voor de bekende sprekers een eigen akoestisch profiel gebruikt worden zodat de herkenning optimaal is. Uiteraard wordt de spraak niet foutloos herkend en is het zeker niet zo dat er geheel automatisch een ondertiteling van de rechtszitting gemaakt kan worden.

Het grote voordeel van het systeem zit in het kunnen zoeken in de opname op spreker,

gesproken woord en deel van de rechtszitting. Gevonden passages kunnen dan met één druk op de knop worden beluisterd. Vooralsnog is het systeem bedoeld om het werk van de griffier te ontlasten.

Arjan van Hessen
Universiteit Twente
en Telecats



Politie

Net zoals bij de rechtbank, worden ook bij de politie veel verhoren afgenomen. Anders dan bij de rechtbank, is het verhoor bij de politie meer één-op-één (verdachte en ondervrager) en is het taalgebruik minder formeel. Wanneer het om een incidentele ondervraging gaat waarbij een verdachte, getuige of expert maar één of tweemaal wordt verhoord, is het weinig zinvol om veel tijd te spenderen aan het creëren van een eigen akoestisch model. Anders ligt dat bij mensen die gedurende lange tijd vaak worden ondervraagd. In die gevallen levert de initiële inspanning van het maken van een eigen akoestisch profiel voldoende voordeel op (betere herkenning) om die inspanning ook te rechtvaardigen.

NFI

De meest minimale vorm van verhoren is die waarbij men zichzelf bevraagt middels een beschrijvend verslag. Dit is wat de experts van het NFI doen wanneer ze sporenonderzoek doen op een plaats delict. Om vervuiling van de sporen te voorkomen, mogen alleen bepaalde vooraf gescreende mensen

naar binnen. Zij doen het sporenonderzoek en nemen met een camera op het hoofd alles op wat ze doen. Soms maken ze schriftelijke aantekeningen en na afloop wordt er een verslag van het geheel gemaakt. Het zou zoveel makkelijker zijn wanneer ze commentaar konden geven tijdens het sporenonderzoek en wanneer deze spraak dan realtime, of na afloop, door de herkenner zou worden gehaald. Ook hier is het resultaat een synchrone tekst en spraak die het mogelijk maakt om in de opnamen te gaan zoeken. Op dit moment zijn er gesprekken met het NFI gaande om te zien of een proof of concept mogelijk is.

Tapgesprekken

Met enige fantasie zou men het af luisteren van een telefoongesprek óók als een vorm van verhoor kunnen definiëren. Hoewel er door de politie geen vragen gesteld worden, wordt de opgenomen spraak wel gebruikt bij de waarheidsbevinding en uiteindelijk ook gebruikt in het proces. Echter, de spraak van de meeste opgenomen telefoongesprekken is meestal zo slecht, dat automatische herkenning niet mogelijk is. Toch kan spraakherkenning hier zinvol gebruikt worden! Niet de telefoonspraak wordt door de herkenner gehaald, maar de door de politiemans of -vrouw nagesproken conversatie. Dit herkennen van nagesproken spraak heet *respeaking* en wordt al met veel succes door de omroepen gebruikt bij het live ondertitelen van tv-programma's.

Op dit moment wordt een tapgesprek binnen 14 dagen nadat het is opgenomen door de politiebeambte zo letterlijk mogelijk uitgeschreven, iets dat de meeste beambten niet heel erg plezierig vinden. Dit uitschrijven kost ongeveer 6 tot 8 keer meer tijd dan het gesprek zelf. Door het gesprek gewoon na te spreken en deze nagesproken tekst te herkennen, wordt de benodigde tijd voor het uitschrijven sterk bekort. Een bijkomend voordeel is dat er vervolgens ook in de tekst gezocht kan worden omdat de nagesproken spraak en de herkende tekst op millisecondenniveau aan elkaar gekoppeld zijn. Omdat bij het naspreken ook de originele spraak beschikbaar is, krijgen we uiteindelijk drie aan elkaar gekoppelde 'informatielagen': de originele opgenomen spraak, de nagesproken spraak en de herkende/uitgeschreven tekst. De drie lagen zijn op tijdsniveau aan elkaar gekoppeld. Hierdoor kan er gezocht worden in de tekst en kan naar believen een van de twee audiolagen worden beluisterd.

Maar als dit mogelijk is, dan moet het ook mogelijk zijn om dit te doen bij tapgesprekken waarbij de originele taal geen Nederlands

is. Op dit moment wordt er voor niet-Nederlandse tapgesprekken een beëdigde tolk-vertaler gebruikt die de gesprekken moet uitschrijven. Door deze tolk-vertaler de opgenomen spraak direct in het Nederlands te laten naspreken en dan de Nederlandse spraak door de herkenner te halen, krijgen we net als hiervoor een gekoppeld drielaagsmodel. Het enige verschil is dat de brontaal in dit geval geen Nederlands is.

Kentekenlijn

Omdat spraakherkenning, zeker bij beperkt woordgebruik, ook goed mogelijk is met mobiele toepassingen, werd in 2007 in het kader van het STEVIN-demonstratieproject Kentekenlijn door de politie in Utrecht een experiment uitgevoerd met het herkennen van nummerborden van geparkeerde auto's (taalunieversum.org/taal/technologie/stevin/projecten/#SNRT)

In plaats van het intoetsen of scannen van de nummerborden (waarvoor altijd een relatief zwaar apparaat nodig is) werd onderzocht of het mogelijk is om de nummerborden gewoon via de mobiele telefoon in te spreken en te herkennen. De proef slaagde geheel en momenteel wordt onderzocht of de pilot verder wordt uitgerold.

Conclusie

De afgelopen jaren werd er op verschillende niveaus door Justitie en politie onderzoek gedaan naar de mogelijkheden om taal- en spraaktechnologie in te zetten in het dagelijks werkproces. Behalve bij de Kentekenlijn waar het eigenlijk ging om het mondeling bevragen van een database, zijn alle applicaties gericht op het doorzoeken en vervolgens af luisteren van opgenomen spraak. We begonnen deze bijdrage bij de ietwat formele spraak in de rechtbank. Vervolgens werden de mogelijkheden onderzocht om dezelfde technologie in te zetten op verschillende niveaus van het politiewerk. De toepassing voor het NFI is hier weer direct aan gekoppeld.

Automatisch vertalen en de overheid

Elke overheid heeft nu en dan nood aan vertalingen van documenten. Het kan hierbij gaan om documenten van die overheid die ter beschikking moeten komen van anderstaligen of het kan gaan om documenten van anderstaligen (eventueel andere overheidsinstanties) die vertaald moeten worden naar de taal van de overheid zelf. De vertalers van de Belgische federale overheid maken geen of nauwelijks gebruik van automatische vertaling om eigen documenten in andere talen beschikbaar te maken. Wel wordt er gebruik gemaakt van vertaalgeheugens en tools zoals Trados of EURAMIS.

Vertaalgeheugens

Een vertaalgeheugen is een collectie reeds manueel vertaalde fragmenten. Via softwaretools wordt het vertaalgeheugen geïntegreerd in tekstverwerkers en in de werkomgeving van de vertaler, zodat deze snel en gemakkelijk al vertaalde fragmenten kan opsporen en zelf kan beslissen welke hij gebruikt en of er aanpassingen moeten gebeuren. Aangezien in veel documenten bepaalde zinnen of fragmenten vaker terugkomen, is het goed dat deze niet telkens opnieuw manueel vertaald worden, maar door middel van een vertaalgeheugen een consistente vertaling krijgen.

Vertaalbureaus

Volgens een recente studie van Forbes maken vertaalbureaus wel gebruik van automatische vertaling, maar hier wordt niet bij vermeld hoe. Wordt de automatische vertaling systematisch gebruikt in het vertaalproces in combinatie met post-editing? Wordt automatische vertaling slechts gebruikt om af en toe dingen op te zoeken? Of is de workflow geautomatiseerd zonder dat de feitelijke vertaling geautomatiseerd is?

NL-Translex

In 1999 zette de Nederlandse Taalunie een project op om het gebruik van automatische vertaling door de overheid van en naar het Nederlands mogelijk te maken, en zocht ze daarvoor een private partner. NL-Translex werd ontwikkeld door Systran, maar lijkt achteraf bekeken niet vaak gebruikt te worden door overheden, onder andere door de hoge kosten voor aanpassing van de woordenboeken en integratie in de workflow. Er zijn wel een aantal grote bedrijven, zoals Fortis AG, die claimen tot 18% productiviteitswinst gehaald te hebben door het gebruik ervan.

Europa

De Europese overheid is bij ons de enige overheid die echt zelf automatische vertaling gebruikt. De Europese Commissie heeft

jarenlang Systran gebruikt, van midden jaren 70 tot nu, en ontwikkelde zelf haar eigen variant van het systeem: ECMT (European Commission Machine Translation). NL-Translex werd nooit aangepast voor of door de Commissie. In 2006 werd besloten de verdere ontwikkeling en verbetering van ECMT stop te zetten, omdat het niet vaak genoeg gebruikt werd door de vertalers van de Commissie. Woordenboeken kunnen wel nog steeds door de vertalers zelf worden aangevuld.

Ook al gebruiken de vertalers het systeem niet of niet vaak, het blijkt wel regelmatig gebruikt te worden door niet-vertalers om een idee te krijgen waarover een bepaalde anderstalige tekst gaat ("gisting" genoemd). De Commissie heeft het ECMT via een webservice opengesteld voor de overheidsinstanties van de lidstaten en de andere gebruikers van hun websites. Op die manier wordt het systeem dagelijks gebruikt.

Ondertussen voert het Euromatrixplus-project in opdracht van de Commissie experimenten uit met statistische vertaalsystemen, die getraind worden op de gigantische vertaalgeheugens van de Commissie. Aangezien meertaligheid het officieel beleid van de Commissie is, zal automatische vertaling alsmat belangrijker worden, misschien meer nog bij andere diensten dan bij de vertaalafdeling zelf.

PaCo-MT

Binnen het STEVIN-project PaCo-MT (Parse and Corpus-based Machine Translation), bouwen we verder op het succes van vertaalgeheugens binnen de vertaalsector. We bouwen een machinevertaalsysteem dat gebruik maakt van dezelfde principes als een vertaalgeheugen, maar dat zelf beslist welke vertalingen er gekozen worden en welke aanpassingen er nog moeten gebeuren.

Hierbij gaan we als volgt te werk: we nemen de verslagen van het Europees parlement

Vincent Vandeghinste
Centrum voor
Computerlinguïstiek,
K.U. Leuven

van 1996 tot 2006 in de verschillende talen waarmee we willen werken, in ons geval Nederlands, Frans en Engels. Deze verslagen worden heel vaak gebruikt in datagedreven automatische vertaalsystemen (zoals statistische vertaalsystemen), omdat ze gratis beschikbaar zijn in alle officiële talen van de lidstaten die op dat moment lid waren van de EU. In deze verslagen worden de zinnen gealigneerd: er wordt aangeduid welke zinnen vertalingen van elkaar zijn. Zinsalignering is eveneens een belangrijke component in een vertaalgeheugen. Daarnaast wordt ook woordalignering uitgevoerd: welke woorden zijn vertalingen van elkaar? Uit deze gegevens kunnen woordenboeken afgeleid worden als we weten hoe vaak welk woord als wat vertaald wordt. Let wel, al deze aligneringen gebeuren machinaal, zonder menselijke interventie, en zijn dus verre van perfect. We hopen deze imperfectie op z'n minst gedeeltelijk te compenseren door de grote hoeveelheden data die door het systeem in rekening gebracht worden.

Op basis van al de aligneringen wordt een eerste vertaalmodel opgesteld, dat het vertaalsysteem toestaat beslissingen te nemen omtrent welke vertaalde fragmenten er gekozen worden en hoe deze deelvertalingen gecombineerd worden tot een nieuwe zin. Als basissysteem gebruiken we de teksten van het Europees Parlement, omdat het hierin gehanteerde taalgebruik breed en gevarieerd is. Een mogelijk probleem hierbij is dat het niet per se om rechtstreekse vertalingen van elkaar gaat. Zowel bron- als doeltaaldocument kunnen vertaald zijn uit een derde taal.

Omdat niet elke overheid hetzelfde taalgebruik hanteert als het Europees Parlement is het belangrijk PaCo-MT aan te passen aan het domein waarvoor het gebruikt zal worden. Dit kan gebeuren door een reeds bestaand vertaalgeheugen in te laden en te bewerken zoals hierboven beschreven. Zo past het systeem zijn taalgebruik automatisch aan het type vertalingen aan dat ingeladen wordt.

Omdat we er niet onmiddellijk van uit gaan dat PaCo-MT altijd de gewenste vertaling oplevert, wordt er toch nog een fase van nabewerking voorzien, waarbinnen de vertaler de door de computer genomen beslissingen kan corrigeren. Deze correcties worden teruggevoerd naar het systeem zodat, na verloop van tijd, PaCo-MT zichzelf aanpast aan de vertaalgewoontes van de vertaler in kwestie.

De laatste paar jaar is er een zeer grote toename van het wetenschappelijke onderzoek naar automatisch vertalen, en er wordt dan ook een aanzienlijke vooruitgang verwacht op kwalitatief gebied. PaCo-MT hoopt daaraan bij te dragen door taalparen met het Nederlands te ontwikkelen. Eenmaal die taalparen kwalitatief goed genoeg zijn zal dat hopelijk leiden tot meer gebruik van automatische vertaling bij de overheid, en dus ook tot een efficiëntere overheid.



TDA: één wereldomvattend vertaalgeheugen

Stel je voor dat er één centrale database is waar alle vertaalgeheugens van de hele wereld kunnen worden opgeslagen en teruggevonden voor hergebruik door vertalers en ontwikkelaars van vertaalsystemen.

Vanaf juni 2009 bestaat die 'super-cloud of language data' en hij heet TDA, oftewel TAUS Data Association. TDA werd in juli 2008 opgericht door een groep van 45 bedrijven, waaronder grote IT-bedrijven en vertaal- en technologiebedrijven. Enkele maanden geleden is het platform gelanceerd met drie eenvoudige services: TM Sharing, Data Pooling en Language Search.

TM Sharing stelt leden in staat om vertaalgeheugens in de database te plaatsen. Data Pooling laat gebruikers selecties maken in de database en vertaalgeheugens ophalen. Alle vertaalgeheugens kunnen worden geraadpleegd met een zeer geavanceerde Language Search Engine, die de vertalingen van zoekwoorden weergeeft in context en met verwijzingen naar industrie, eigenaar van de vertalingen en linguïstische attributen. De Language Search Engine – als een vrij toegankelijke service van TDA – zoekt in betrouwbare bronnen van kwaliteitsvertalingen. TDA heeft een widget beschikbaar gesteld die vertalers op hun desktop kunnen installeren om zo-doende snel en gemakkelijk van deze nieuwe tool gebruik te kunnen maken. TDA is een in-

dustrievereniging zonder winstoogmerk. TDA verwelkomt alle bedrijven, overheidsorganisaties, universiteiten en vertalers als leden.

De beschikbaarheid van grote volumes vertaalgeheugens stelt een ieder in staat om machinevertaalsystemen te trainen en aan te passen aan verticale sectoren. De statistische of corpusgebaseerde vertaalsystemen, zoals PaCo-MT, LanguageLens, PageAMatic en Translated.net zullen direct kunnen profiteren van de data van TDA-leden. Sommige van de TDA-leden gebruiken de data al om het open source machinevertaalsysteem Moses te trainen voor hun eigen doeleinden.

De TDA-vereniging is gevestigd in Nederland en wordt geleid door Jaap van der Meer. Bij het ter perse gaan van dit nummer bevatte de database al bijna 1 miljard woorden in 84 taalparen.

Voor meer informatie over TDA en een verslag van de testprojecten, zie www.tausdata.org of contacteer customerservice@tausdata.org.

Jaap van der Meer
TAUS Data Association



Digitale bronnen: termextractie helpt bij het zoeken naar informatie

Vaktaal, terminologie en termextractie

De communicatie tussen burger en overheid wordt alsmaar complexer: dossiers worden technischer, informatie wordt ingewikkelder. Vaak wordt informatie uit andere talen vertaald, of in een meertalig land zoals België wordt de informatie in meerdere talen beschikbaar gesteld. Bij digitale teksten wordt meestal een zoekfunctie aangeboden, die het mogelijk maakt om via trefwoorden informatie terug te vinden. Als we begrippen, termen of afkortingen willen terugvinden in een digitale tekst, dan kunnen we daarvoor beroep doen op een specifieke software-ontwikkeling: een termextractor.

Frieda Steurs
Lessius /
K.U. Leuven

Termextractie is een semi-automatisch proces dat kandidaattermen detecteert in een-talige teksten of in parallelle teksten (we bedoelen hier een tekst met zijn vertaling). Dit proces is nuttig als hulp bij het aanleggen van terminologiecollecties op basis van grote hoeveelheden tekst. Termextractie is dat onderdeel van terminologiewerk dat betrekking heeft op het overlopen van een tekst of een groep teksten om begrippen en hun aanduidingen (termen, afkortingen) te identificeren, het verzamelen van alle relevante informatie over een begrip (zoals definities en contexten) en het gebruik van termen. Termextractie is een belangrijke stap in het systematisch onderzoeken van de terminologie in een vakgebied. Termextractie kan ook zeer nuttig zijn om op de hoogte te blijven van recente terminologische ontwikkelingen in het vakgebied. Nadat kandidaattermen in een tekst geselecteerd zijn en de nodige conceptuele en gebruiksinformatie is verzameld, is verder onderzoek meestal vereist om te bepalen of de kandidaattermen daadwerkelijk deel uitmaken van een vaktaal en of de begrippen waarnaar ze verwijzen deel uitmaken van het vakgebied dat bestudeerd wordt. Nadat de begrippenanalyse werd uitgevoerd, is het mogelijk dat een aantal kandidaattermen aan de kant geschoven worden.

Programma's voor termextractie zijn al geruime tijd door de academische wereld ontwikkeld, op basis van statistische en/of taalkundige informatie. Termextractie is ook door een aantal commerciële bedrijven ontwikkeld, maar dan meestal via een statistische methode, niet taalgebonden en zonder toevoeging van taalkundige kennis.

Termextractie is mogelijk voor één of meerdere talen. De eentalige termextractie tracht een tekst of document te analyseren om

kandidaattermen te identificeren, terwijl de meertalige termextractie bestaande bron-teksten en hun vertalingen analyseert om kandidaattermen en hun vertalingen te identificeren. Zelfs al wordt de oorspronkelijke extractie uitgevoerd door een computerprogramma, dan nog moet de resulterende lijst van kandidaattermen nagekeken worden door een terminoloog of vertaler. Hierdoor is het extractieproces meer computerondersteund dan volledig automatisch.

Er zijn twee voorname systemen voor termextractie: linguïstisch of statistisch. Termextractietools die gebruik maken van een linguïstische aanpak proberen woordcombinaties te identificeren die overeenstemmen met sommige zinsconstructies (v.b. 'adjectief + zelfstandig naamwoord' of 'zelfstandig naamwoord + zelfstandig naamwoord'). Het is duidelijk dat een dergelijke aanpak sterk taalgebonden is, omdat de constructies verschillen van taal tot taal. Hierdoor worden termextractietools die gebruik maken van een linguïstische aanpak gecreëerd om in één enkele taal (of sterk gelijkende talen) te werken, en kunnen ze dus niet zonder moeite tot andere talen uitgebreid worden.

Termextractietools die gebruik maken van een statistische aanpak gaan op zoek naar herhaalde reeksen van lexicale eenheden. Om als een mogelijke kandidaat erkend te worden, moet een term een zekere frequentie hebben. Vaak kan de gebruiker deze frequentie specificeren. Het grootste voordeel van een statistische aanpak is zijn taalafhankelijkheid. Alhoewel, de hoeveelheid 'ruis' (d.w.z., het aantal onwaarschijnlijke termen) en 'stilte' (d.w.z., het aantal termen dat niet geïdentificeerd is) is relatief hoog. Om dit te vermijden worden vaak verschillende aspecten van de twee aanpakken gecombineerd in hybride termextractietools.

De voornaamste opdracht van de Commissie Terminologie van de Nederlandse Taalunie is het (doen) ontwikkelen van algemene hulpmiddelen en applicaties ten behoeve van het Nederlandstalige terminologieveld. Binnen die doelstellingen past de ontwikkeling van een goede, algemeen inzetbare en speciaal op het Nederlands gerichte terminologie-extractor, omdat een dergelijk hulpmiddel tot nu ontbreekt. Er bestaan commerciële producten, maar die zijn onvoldoende gericht op de verschillende vormen van gebruik, te kleinschalig en te zeer gericht op de onmiddellijke praktijkbehoeftes van vertalers, aldus de Commissie. De beschikbaarheid van een goede, op Nederlands tekstmateriaal toegespitste, extractietool is een noodzakelijke voorwaarde om de opbouw van terminologiecollecties zo kostenefficiënt mogelijk te maken, door beperking van de hoeveelheid handmatig werk.

Terminologieverzamelingen worden gebruikt als lexicale component in uiteenlopende taalverwerkende systemen en processen, bv. automatische vertaalsystemen (in de vorm van zgn. gebruikerswoordenboeken), vertaalhulpmiddelen zoals vertaalgeheugens, automatische informatiediensten, spraakgestuurde apparaten, etc.

Potentiële gebruikers van termextractiesoftware zijn dan ook: uitgevers van terminologie, beheerders van terminologie (o.a. de TST-Centrale), ontwikkelaars van taalverwerkende systemen, vertalers en/of terminologen binnen (ver)taaldiensten van overheid of bedrijfsleven.

De Taalunie en de Commissie Terminologie streven niet alleen naar de beschikbaarheid van een extractiehulpmiddel ten behoeve van het terminologische veld in zijn algemeenheid, maar hebben ook zelf behoefte aan zo'n voorziening. Krachtens het Verdrag tussen België en Nederland heeft de Taalunie tot doel te zorgen voor een gelijke terminologie ten behoeve van wetgeving en officiële publicaties. Omdat die opdracht een verborgen opdracht inhoudt tot vergelijking en recht-trekking van bestuurs- en rechtssystemen, is ervoor gekozen om de officiële terminologie uit Nederland en België/Vlaanderen te verzamelen en in relatie tot elkaar te beschrijven. Door het heel uitgebreide karakter van de overheidsterminologie en door het feit dat nieuwe wetgeving leidt tot voortdurende wijzigingen, is die taak onmogelijk zonder goede infrastructurele voorzieningen uit te voeren.

Bijna alle beschikbare applicaties die gebruik maken van taal- en spraaktechnologie zijn gericht op professionele gebruikers in

specifieke domeinen, bijvoorbeeld juristen of medici. Zulke applicaties kunnen slechts voldoende onmiddellijk rendement opleveren als zij op domeinspecifieke woordenschat zijn toegesneden. Daarom is speciale aandacht voor vaktalen onontbeerlijk.

In een adviesstudie in opdracht van Coterm werden verschillende commerciële termextractieproducten aan nauwgezette testen onderworpen, om bestaande software te beoordelen, en om criteria voor een performante termextractor voor het Nederlands te bepalen. Vervolgens heeft een wetenschappelijk comité het lastencahiers voor de gunning voorbereid. Na gunning van het project aan een consortium kan de termextractor voor het Nederlands worden ontwikkeld. De gewenste termextractor dient accuraat, compleet en gebruiksvriendelijk te zijn. Het termextractieproject is daarom ambitieus, maar zal een belangrijke bijdrage leveren aan de ontsluiting van de Nederlandstalige bronnen. Dit kan er voor zorgen dat het Nederlands een prominente plaats blijft houden in de lijst van de talen met digitale informatievoorzieningen.



Spraakrouteren biedt voor overheidsdiensten aantrekkelijke perspectieven om bellende burgers op een efficiënte manier bij het juiste loket te brengen. Onderstaand voorbeeld illustreert de werking van spraakrouteren bij een bank, en het belang van een goede betrokkenheid van de medewerkers van het callcenter.

Spraak in bedrijf

"Welkom bij ABN AMRO, waarmee kan ik u van dienst zijn?". Voor de klanten van ABN AMRO is dit geen toekomstmuziek meer. Sinds de zomer van 2008 horen bellende klanten deze open vraag in plaats van het conventionele toetsgestuurde keuzemenu. Klanten kunnen in hun eigen woorden vertellen waarvoor ze bellen, waarna ze direct doorverbonden worden met een relevante selfservice of de juiste medewerker in het callcenter. Deze nieuwe dienst, ook wel 'spraakrouteren' genoemd, is gebaseerd op 'open spraakherkenning' en is geïmplementeerd door het team Customer Contact Solutions van Logica.

**Diana
Binnenpoorte
en
Christophe
Van Bael**
Logica

Voor de bellende klanten zijn de voordelen duidelijk: spraakrouteren is snel, gemakkelijk en zeer klantvriendelijk. Maar wat betekent het voor het bedrijf, meer specifiek, wat betekent de inzet van spraakrouteren voor de medewerkers van het callcenter?

Een eerste voordeel is dat medewerkers van callcenters minder vaak klanten hoeven door te verbinden naar een collega. Met de inzet van spraakrouteren zullen klanten vaker direct naar de juiste medewerker worden gerouteerd en bovendien heeft het 'nuldrukken' van klanten bij wijze van escape geen effect meer.

klantvraag op de desktop van de medewerker. Uit beide sessies wordt duidelijk wat voor medewerkers de grootste veranderingen zijn en waarop in de training de nadruk moet liggen. Eén van de belangrijkste punten die in de gesprekken met de medewerkers van ABN AMRO naar voren kwamen, was dat de medewerkers zich ambassadeurs van hun bedrijf voelen, en dat ze daarom graag weten wat spraakrouteren inhoudt, hoe het werkt en wat de voordelen voor de klanten zijn. Naast deze inhoudelijke punten, is vooral ook aandacht besteed aan de interpretatie van de herkende klantvragen. Aangezien open spraakherkenning klantvragen niet noodzakelijk 100% cor-



Een tweede voordeel volgt uit het feit dat open spraakherkenning het toelaat om de herkende klantvragen op de desktop van de medewerker af te beelden. De medewerker weet dan voor aanvang van het gesprek met de klant waarvoor die klant belt. Dit heeft tot gevolg dat de gespreksvoering tussen de medewerker en de klant verandert. De klant heeft het systeem immers al verteld waarvoor hij belt en de medewerker kan direct doorvragen of de vraag beantwoorden.

De ervaring leert dat deze nieuwe gespreksvoering om een goede voorbereiding van de callcentermedewerkers vraagt. Het is van belang om bij de medewerkers te informeren naar hun visie op de aanstaande veranderingen. Dit kan middels een brainstormsessie met een kleine groep medewerkers, gevolgd door een gebruikerstest met betrekking tot de weergave en de interpretatie van de herkende

rect weergeeft, dient de medewerker gevoel te ontwikkelen om te kunnen beslissen welke woorden relevant zijn en welke niet. Uit deze trainingssessies kwam al snel naar voren dat er behoefte is aan een leeshulp voor een gemakkelijkere interpretatie van herkende klantvragen. Deze leeshulp is eenvoudig te implementeren als een visuele weergave van de betrouwbaarheidsmaten van de spraakherkenner.

Het trainen van callcentermedewerkers is een dure aangelegenheid; over het algemeen is er binnen callcenters een groot verloop van personeel. Echter, het succes van een spraakrouteringsapplicatie is niet alleen afhankelijk van de prestaties van het systeem. Ook de inzet van goed opgeleide en enthousiaste medewerkers is belangrijk voor een succesvolle implementatie en klantacceptatie van spraaktechnologie voor het grote publiek.

NOTaS

Stichting NOTaS

Secretariaat: Postbus 31070
6503 CB NIJMEGEN
T: 024-352 88 88
W: www.notas.nl
E: notes@malta-online.nl

NOTaS behartigt de belangen van bedrijven en kennisinstellingen die actief zijn op het terrein van Taal- en Spraaktechnologie (TST). Dit doet zij o.a. door middel van bijeenkomsten, lobbyactiviteiten en het tijdschrift DIXIT.

Deelnemers Stichting NOTaS

CLST

Postbus 9103, 6500 HD NIJMEGEN
Bezoekadres: Erasmusplein 1,
6525 HT 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. Belangrijke speerpunten zijn robuuste ASR en de inzet van TST ten behoeve van het onderwijs en voor mensen met communicatieve beperkingen. Tevens valt onder CLST het Speech Processing Expertise Center (SPEX), dat is gespecialiseerd in de productie en validatie van spraakdatabases.

Data Archiving and Networked Services (DANS)

Postbus 93067, 2509 AB Den Haag
T: 070-349 44 50
W: www.dans.knaw.nl
E: info@dans.knaw.nl

DANS is een instituut van de Koninklijke Nederlandse Akademie van Wetenschappen (KNAW), dat mede wordt ondersteund door de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). DANS zorgt voor de opslag en blijvende toegankelijkheid van onderzoeksgegevens in de alfa- en gammawetenschappen. Daartoe ontwikkelt DANS zelf duurzame archiveringsdiensten, bevordert het dat anderen dat doen, en werkt samen met databeheerders om ervoor te zorgen dat zo veel mogelijk data vrij beschikbaar komen voor gebruik in het wetenschappelijk onderzoek.

Dedicon

Postbus 24, 5360 AA GRAVE
Bezoekadres: Traverse 175 5361
TD GRAVE
T: 0486-48 64 86
W: www.dedicon.nl
E: info@dedicon.nl

Dedicon is in Nederland dé organisatie die lectuur en informatie toegankelijk maakt. Zo kan iedereen lezen wat hij wil in de vorm die hem past. Binnen de diensten van Dedicon speelt de moderne taal- en spraaktechnologie een belangrijke rol.

Dialogs Unlimited BV

Adres: Takbeijsters 17 sub 11,
4817 BL BREDA
T: 076-572 47 77
W: www.dialogsunlimited.com
E: info@dialogsunlimited.com

Dialogs Unlimited is de ontwikkelaar van de unieke ImmediateDialogs™ software suite, een complete product familie voor een geïntegreerde Self Service Delivery voor websites en spraakapplicaties, waarvan o.a. ImmediateVoice™ en de ivWebcom-municator™ deel uit maken. Dankzij de online ImmediateDialogs™ ontwikkelomgeving (SaaS) kunnen, zonder kostbaar specialisme in te huren of op te bouwen, in eigen beheer applicaties gebouwd en onderhouden worden.

Dutchear

Postbus 5050, 2600 GB DELFT
Bezoekadres: Brassersplein 2,
2612 CT DELFT
T: 015-219 11 11
W: www.dutchear.nl
E: info@dutchear.nl

Dutchear past spraaktechnologie toe voor telefonische selfservice, open vraag spraakherkenning en audio mining. Dutchear ontwerpt en realiseert de best-of-breed oplossing

op basis van deze technologieën, die voor elke organisatie de beste prijs-kwaliteits oplevert.

GridLine B.V.

Adres: Keizersgracht 520,
1017 EK AMSTERDAM
T: 020-616 20 50
W: www.gridline.nl
E: info@gridline.nl

GridLine is een succesvol IT-bedrijf dat gespecialiseerd is in taaltechnologie voor het Nederlands. Een aantal van onze producten en diensten: Klinkende Taal - plugin die ambtenaren helpt begrijpelijke brieven te schrijven; FAST; Share-Point; Lucene; NedLex - informatiplatform voor advocaten; de TaalServer - voor alle taalopgaven in uw bedrijf; opinion mining, information retrieval, kennismanagement. GridLine, voor professionals die het Nederlands gebruiken.

intaal B.V.

Adres: Winthontlaan 198,
3526 KV UTRECHT
T: 030-750 89 60
W: www.intaal.nl
E: info@intaal.nl

intaal past spraak- en taaltechnologie toe binnen de zorg. inTAAL ontwikkelt en levert hulpmiddelen ter ondersteuning van de communicatie en ter vergroting van de zelfredzaamheid van de gebruikers. inTAAL is tevens gespecialiseerd in de toepassing van spraakherkenning voor bijzondere doelgroepen zoals mensen met functiebeperkingen en leerlingen met dyslexie.

Knowledge Concepts

Adres: De Handboog 9,
5283 WR BOXTEL
T: 0411- 61 08 02
W: www.knowledge-concepts.com
E: sales@knowledge-concepts.com

Gereven door passie en ervaring in dit vakgebied ontwikkeld en realiseert Knowledge Concepts vooruitstrevende systemen ten behoeve van informatieverwerking. Onze oplossingen onderscheiden zich in aspecten als: verzamelen, samenvoegen, uitlijnen, opslaan, ontsluiten en verspreiden van informatie en kennis. Dit realiseren wij met name door innovatief gebruik van taaltechnologie in combinatie met o.a. zoekmachines, document management systemen en portals.

Logica

Customer Contact Solutions
Postbus 1360, 3430 BJ Nieuwegein
Bezoekadres: Merweplein 23,
3432 GN Nieuwegein
T: 030-602 77 00
W: www.logica.nl
E: ccsnl@logica.nl

De ontwikkelingen op het gebied van klantcontacten volgen elkaar in hoog tempo op. Staat binnen uw organisatie de klant centraal en wilt u de klanttevredenheid verhogen? En wilt u tegelijkertijd de operationele kosten verlagen? Dan bent u bij het Customer Contact Solutions Center van Logica aan het juiste adres. Het Customer Contact Solutions Center van Logica staat voor brede kennis en ervaring op het gebied van klantcontacten, contact center oplossingen en spraaktechnologie.

OSTT/Sint Maartenskliniek

Adres: Hengstdal 3,
6522 JV NIJMEGEN
T: 024-365 97 18
W: www.ostt.eu
W: www.maartenskliniek.nl
E: ostt@maartenskliniek.nl

Het "Ontwikkelcentrum voor Spraak- en Taaltechnologie ten behoeve van spraak- en Taalpathologie en revalidatie algemeen" is erkend als expertisecentrum door het Ministerie van het VWS aan de Sint Maartenskliniek. In het kader van het OSTT werkt de Sint Maartenskliniek samen met de afdeling Taalwetenschap van de Radboud Universiteit en met het UMC St. Radboud in Nijmegen.

PonTeM

Adres: Petrus Dondersplein 1,
5271 AA Sint-Michielsgestel

Postbus 7,
5270 BA Sint-Michielsgestel
T: 073-558 84 84
F: 073-551 78 97
E: info@pontem.nl
W: www.pontem.nl

PonTeM R&D maakt onderdeel uit van de sector Expertise & Innovatie

van Koninklijke Kentalis. PonTeM realiseert onderzoek, ontwikkeling en innovatie met als doel het kennisbestand en de product - dienst-combinaties in zorg, onderwijs, diagnostiek en dienstverlening voor mensen met auditieve en communicatieve beperkingen te vergroten, te vernieuwen en te valideren.

Q-go.com B.V.

Adres: 'Diemercircle' Eekholt 40,
1112 XH DIEMEN
T: 020-531 38 00
W: www.q-go.nl
E: info@q-go.com

Q-go is een internationale aanbieder van Natural Language Search technologie. De natuurlijke taaltechnologie wordt ingezet in applicaties waarmee onder andere banken, verzekeraars, telecombedrijven en logistieke dienstverleners hun online klantenservice verbeteren, hun kosten verlagen, en hun inzicht in de klant vergroten. Q-go werkt 40% van haar opbrengsten aan voor R&D.

Telecats

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

Telecats ontwikkelt producten en diensten voor spraakinteractie en -analyse met een focus op telefonie. Met onze oplos-singen voor Inter-actieve Voice Response en (open vraag) spraakherkenning worden telefoongesprekken (deels) geautomatiseerd. Met spraakanalyse kunnen inhoud en toon van live gesprekken en audioarchieven worden ontsloten. De taal- en spraaktechnologieën zijn als webservices beschikbaar voor derden die deze willen integreren in hun eigen propositie.

TNO Defensie en Veiligheid

Postbus 23, 3769 ZG SOESTERBERG
T: 0346-356 205/211
W: www.tno.nl
E: rob.drullman@tno.nl

De business unit Human Factors van TNO Defensie en Veiligheid in Soesterberg benadert spraaktechnologie op een integrale manier. TNO past de technologie niet alleen toe, maar helpt ook deze verder te ontwikkelen. Daarbij richten we ons op automatische spraak-, emotie- en taalherkenning en werken we ook aan evaluatie en implementatie van spraaksynthese en gesproken dialoogsyste-men.

TST-Centrale, p/a Instituut voor Nederlandse Lexicologie

Postbus 9515, 2300 RA LEIDEN
Bezoekadres: Matthias de Vrieshof
2-3 2311 BZ LEIDEN
T: 071-527 24 95
W: www.inl.nl/tst-centrale
E: servicedesk@inl.nl

Wanneer u op zoek bent naar (informatie over) digitale taalkundige bronnen dan bent u bij ons aan het juiste adres. Of u voor een kennisinstelling werkt of voor het bedrijfsleven, of u geïnteresseerd bent in taal of in spraak, of u een taalkundige bent of een spraaktechnoloog, wij zijn u graag van dienst.

Universiteit van Tilburg – CIW

Postbus 90153, 5000 LE TILBURG
Bezoekadres: Warandelaan 2, 5037
AB TILBURG
T: 013-466 91 11
W: www.uvt.nl
E: uvt@uvt.nl

In onderwijs en onderzoek van het Departement Communicatie- en Informatiewetenschappen (CIW) ligt het accent op aspecten van menselijke communicatie, zowel talige als niet-talige, inclusief mens-machine interactie, kennisrepresentatie, text mining, machine learning, kunstmatige intelligentie, computationele semantiek, semantische annotatie, en dialoogmodellen en -systemen. Het departement omvat sinds September 2008 het Tilburg Center for Creative Computing (TICC). Het departement verzorgt onder andere de bloeiende opleiding Bedrijfscommunicatie en Digitale Media, de masteropleiding Human Aspects of Information Technology, en samen met de Radboud Universiteit Nijmegen de research masteropleiding Language and Communication.

Universiteit Twente - HMI

Postbus 217, 7500 AE ENSCHEDE

Bezoekadres: Drienerlolaan 5,
7522 NB ENSCHEDE
T: 053-489 91 11
W: www.utwente.nl
E: info@utwente.nl

De HMI-groep van de Universiteit Twente (UT) is een ondernemende onderzoeksgroep die zich op het gebied van mens-machine interactie zowel met fundamenteel als met meer toepassingsgericht onderzoek bezig houdt. Rondom de campus van de Universiteit Twente bevinden zich een groot aantal spin-off bedrijven waarvan enkele ook op enigerlei wijze bezig zijn met Taal & Spraaktechnologie en Mens-Machine Interactie.

Universiteit Utrecht – Utrecht Institute of Linguistics OTS

Adres: Janskerkhof 13,
3512 BL UTRECHT
T: 030-253 60 06
W: www.uilots.let.uu.nl
E: uil-ots@let.uu.nl

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 taalsysteem, (ii) de cognitieve systemen die ten grondslag liggen aan de verwerving en de verwerking van taal, en (iii) de wijze waarop taal wordt gebruikt in communicatie tussen mensen en tussen mens en machine.

Van Dale Uitgevers

Postbus 19232, 3501 DE UTRECHT
Bezoekadres: St. Jacobsstraat 127
3511 BP UTRECHT
T: 030-232 47 11
W: www.vandale.nl
E: info@vandale.nl

Iedereen kent de Grote of "Dikke" Van Dale. De eerste uitgave van dit woordenboek stamt uit 1864. Bijna anderhalve eeuw later is Van Dale Uitgevers nog steeds bij de tijd en heeft het bedrijf zich ontwikkeld tot de meest toonaangevende en gezag hebbende uitgeverij van woordenboeken in Nederland en België. Met zorgvuldigheid en respect beschrijft Van Dale de taal zo compleet mogelijk, niet alleen het Nederlands, maar ook vreemde talen.

VoiceCorp Webservices B.V.

Adres: Dolderseweg 2A,
3712 BP HUIS TER HEIDE
T: 030-692 44 90
W: www.voice-corp.com
E: infoNL@voice-corp.com

VoiceCorp is een internationale organisatie opgericht in Zweden, met inmiddels kantoren in Duitsland, Engeland, Frankrijk, Nederland, de VS en Zweden. VoiceCorp is marktleider op het gebied van het omzetten van tekst op het internet naar spraak via TTS, en richt zich met haar diensten op toegankelijkheid en gemak. Onder ons hoofdmerk ReadSpeaker leveren wij innovatieve producten voor het voorlezen van websites (de eerste ooit), het toepassen van spraak voor andere web-based applicaties en RSS feeds en applicaties voor het gebruik op mobiele devices.

Sponsors/Samenwerking

Nederlandse Taalunie

Postbus 10595, 2501 HN DEN HAAG
Bezoekadres: Lange Voorhout 19
2514 EB DEN HAAG
T: 070-346 95 48
W: www.taalunie.org
E: info@taalunie.org

De Nederlandse Taalunie is een beleidsorganisatie waarin Nederland, België en Suriname samenwerken op het gebied van de Nederlandse taal, onderwijs en letteren.

STEVIN

Postbus 10595, 2501 HN DEN HAAG
(Ter attentie van Peter Spyns)
Bezoekadres: Lange Voorhout 19
2514 EB DEN HAAG
T: 070-346 95 48
W: www.stevin-tst.org
E: pspyns@taalunie.org

STEVIN is een meerjarig onderzoeks- en stimuleringsprogramma voor Nederlandstalige taal- en spraaktechnologie dat gezamenlijk door de Vlaamse en Nederlandse overheid wordt gefinancierd. STEVIN wordt gecoördineerd en financieel beheerd door de Nederlandse Taalunie.

NWO

Postbus 93138,
2509 AC DEN HAAG
Bezoekadres: Laan van Nieuw Oost-Indië 300 2593 CE DEN HAAG
T: 070-344 06 40
W: www.nwo.nl
E: nwo@nwo.nl
Contactpersoon NOTaS:
E: dijkstra@nwo.nl

De Nederlandse Organisatie voor Wetenschappelijk Onderzoek heeft tot taak het bevorderen van de kwaliteit en vernieuwing van wetenschappelijk onderzoek, het initiëren en stimuleren van nieuwe ontwikkelingen in het wetenschappelijk onderzoek alsmede de bevordering van de overdracht van kennis van de resultaten van door haar geïnitieerd en gestimuleerd onderzoek ten behoeve van de maatschappij.

Ontwikkelingsmaatschappij Oost Nederland N.V.

Postbus 5518, 7500 GM ENSCHEDE
Bezoekadres: Hengelostraat 585
7521 AG ENSCHEDE
T: 053- 851 68 51
W: www.oostnv.nl
E: info@oostnv.nl

Ontwikkelingsmaatschappij Oost Nederland is een NV die door middel van allerlei activiteiten en projecten de economie van Oost-Nederland versterkt en daarmee de werkgelegenheid bevordert. Wij werken voor het Gelderse en Overijsselse bedrijfsleven in opdracht van het Ministerie van Economische Zaken en de Provincies Gelderland en Overijssel.

SenterNovem, voor innovatie en duurzaamheid.

Een sterk innovatief bedrijfsleven in een leefbare, duurzame samenleving. SenterNovem stimuleert duurzame economische groei door een brug te slaan tussen markt en overheid, nationaal en internationaal. Bedrijven, (kennis)instellingen en overheden kunnen bij SenterNovem terecht voor advies, kennis en financiële ondersteuning.

Algemene informatie & advies:
Informatiepunt SenterNovem
T: 030- 239 35 33
E: informatiepunt@senternovem.nl
W: www.senternovem.nl

Ondersteuning

Cumlingua Language Services

Adres: Bronckhorstweg 48,
5363 TZ Velp (N.B.)
T: 0486 - 471 554
W: www.cumlingua.com
E: info@cumlingua.com
• Schrijf-, vertaal en correctie-diensten
• Meer dan 20 talen
• Proefschrijftvertaling en -correctie

Leonard Nijmegen bv

Groenestraat 294
6531 JC Nijmegen
T: 024- 378 26 43
W: www.leonard.nl
E: info@leonard.nl

Whole in one!
Leonard = communicatie+vormgeving+internet+drukwerk

Deykerhoff /Accountants

Postbus 2325, 5202 CH,
's-Hertogenbosch
Bezoekadres: Rompertsebaan 70,
5231 GT, 's-Hertogenbosch
T: 073-623 2450
W: www.deykerhoff.nl
E: denbosch@deykerhoff.nl

Deykerhoff/Accountants is een dynamisch accountantskantoor dat eigenzinnigheid paart aan betrouwbaarheid. Onze proactieve instelling is gericht op een grote betrokkenheid en een op maat gesneden persoonlijk contact met de klanten.

Malta & de Keyzer

Adres: Toernooiveld 300,
6525 EC Nijmegen
T: 024-351 21 08
W: www.malta-online.nl
E: welkom@malta-online.nl

Een probleem op het gebied van secretariaat of office management? Wij lossen het op. Snel, professioneel en helemaal volgens uw persoonlijke wensen. Kijk op onze website voor meer informatie.

U wilt graag weten waarom uw klanten u bellen.

Open vraag Spraakherkenning is de oplossing. De spraakherkenner stelt uw bellers een open vraag: "Geef kort en bondig aan waarvoor u belt".

Met behulp van geavanceerde spraaktechnologie, zoekfuncties en het klantprofiel kan elk klantcontact optimaal worden afgehandeld.

Luister naar de stem van uw klanten met Open vraag spraakherkenning

Telefonische dienstverlening is gebaseerd op:

- Identificatie van de beller: wie belt er?
- Classificatie van de wens: waarvoor wordt gebeld?
- Serviceverlening: welke oplossing bied ik deze beller?

In het streven naar kostenreductie, omzetvergroting en toename van de klanttevredenheid, is het essentieel om elke klant de best passende service te verlenen. Hierbij is het van belang om de bellers bij de juiste selfservice applicatie te krijgen of door te verbinden met de juiste gekwalificeerde medewerker.

Bij veel organisaties is classificatie door middel van hiërarchische keuzemenu's niet eenvoudig. De IVR menu's worden al gauw te breed (te veel opties in één vraag) of te diep (te veel vragen achter elkaar).

Bellers hebben vaak moeite met het maken van de juiste keuze en kiezen daarom voor de categorie "overig".

Wanneer de classificatie wordt gecombineerd met de identificatie van de beller (d.m.v. een klantnummer of via postcode en huisnummer) kan de juiste dienstverlening voor de desbetreffende klant worden bepaald.



Voor meer informatie over open vraag spraakherkenning:

Bel met Telecats 053 488 99 00
mail naar info@telecats.nl
of kijk op www.telecats.nl

Open vraag spraakherkenning geeft bellers de mogelijkheid om hun vraag of probleem in hun eigen woorden kenbaar te maken. Hiertoe wordt de beller een open vraag gesteld:

"Welkom bij de bank. Spreek alstublieft kort en bondig in waarvoor u belt".

De voordelen:

- Snellere en betere classificatie dan bij IVR keuzemenu's;
- Verbetering van de (eerste) routing en vermindering van herrotering;
- Positieve invloed op de klanttevredenheid en werknemerstevredenheid;
- en Kostenbesparing.