

### U wilt efficiënt in uw archieven kunnen zoeken...

Naar alle waarschijnlijkheid beschikt u, als content provider, over grote archieven met digitale tekst, beeld of geluid. U hebt zo uw eigen methoden om efficiënt te zoeken in die data. Grote kans dat u daarbij gebruik maakt van TST: Taal- en Spraaktechnologie. Zoekt u met behulp van trefwoorden? Gebruikt u tools om lange documenten samen te vatten? Of gebruikt u geavanceerde zoektechnologie om (gesproken) documenten terug te vinden? In al deze gevallen komt de onderliggende techniek voort uit de TST.

### ... en TST helpt u daarbij.

Zoals dat bij iedere technologische ontwikkeling gaat, werkt ook in TST het mechanisme van voortschrijdend inzicht. Er wordt doorlopend onderzoek gedaan naar nieuwe en nóg betere toepassingen. Ideeën zijn er genoeg en aan kennis ontbreekt het niet. Nederland en Vlaanderen lopen wat TST betreft voorop in Europa. Maar van één ding is er nooit genoeg: grote hoeveelheden spraak en tekst.

### TST heeft veel data nodig...

Voor het ontwikkelen van betere classificatietechnieken voor gesproken en geschreven teksten, betere spraakherkenners, betere automatische samenvatters en betere zoekstrategieën is heel veel taalmateriaal nodig. De nieuwste generatie taal- en spraaksoftware leert van realistische voorbeelden: hoe schrijft een journalist voor de krant? Hoe spreekt een sportverslaggever bij Studio Sport of bij Sporza? Hoe meer voorbeeldmateriaal, hoe beter de resultaten. TST-resultaten drijven op statistische analyse en de grondstof daarvoor is veel, meer, en nog meer taal.

Teksten en spraakopnamen zijn doorgaans goed beschermd door auteursrechten. TST-onderzoekers maken alleen gebruik van rechtenvrij taalmateriaal, en soms van collecties die de rechthebbenden voor onderzoeksdoeleinden beschikbaar hebben gesteld.

### ... helpt u TST daarbij?

U kunt TST een impuls geven door taalcollecties te doneren. De Nederlandse Taalunie heeft een bronnencentrum, de TST-Centrale, ingesteld, die deze collecties zal beheren. U krijgt de garantie dat het materiaal alléén gebruikt zal worden voor het (verder) ontwikkelen en testen van TST. Teksten en geluidsopnamen mogen na gebruik voor statistische analyses niet verder verspreid worden door TST-onderzoekers. En ... doordat u de mogelijkheid krijgt de ontwikkelingen vanaf het eerste moment te volgen kunt u straks nóg sneller de weg vinden in uw digitale archieven.

De **Nederlandse Taalunie** is een internationale verdragsorganisatie waarin Nederland, België en Suriname samenwerken op het gebied van de Nederlandse taal en letteren en het onderwijs in en van het Nederlands.

**STEVIN** is een meerjarig onderzoeks- en stimuleringsprogramma met als doel het stimuleren van de taal- en spraaktechnologische sector in Vlaanderen en Nederland zodat de positie van het Nederlands in de informatie- en communicatiewereld wordt versterkt. STEVIN wordt gefinancierd en begeleid door de Nederlandse en Vlaamse overheden en uitgevoerd onder auspiciën van de Nederlandse Taalunie. Nederlandse partners zijn de Ministeries van Economische Zaken en Onderwijs, Cultuur en Wetenschappen, de Nederlandse Organisatie voor Wetenschappelijk Onderzoek en SenterNovem; Vlaamse partners zijn het Departement Economie, Wetenschap en Innovatie, het Instituut voor de aanmoediging van Innovatie door Wetenschap en Technologie in Vlaanderen en het Fonds Wetenschappelijk Onderzoek Vlaanderen.

De **TST-centrale** (Centrale voor Taal en Spraaktechnologie) is verantwoordelijk voor het verwerven, beheren, onderhouden en verspreiden van diverse spraak- en taalmaterialen die voor het Nederlandstalig taalgebied met overheidsmiddelen zijn ontwikkeld.

### Voor meer informatie over STEVIN:

I <http://www.taalunieversum.org/stevin>  
E [stevin@taalunie.org](mailto:stevin@taalunie.org)

### Voor meer informatie over de TST-centrale:

I <http://www.tst.inl.nl>  
E [tst@inl.nl](mailto:tst@inl.nl)

