

Niet moe door virus

Nijmeegse wetenschappers van de Radboud Universiteit laten weten dat Chronisch Vermoedheid Syndroom (CVS) niet veroorzaakt wordt door een virus. Onderzoek naar de echte oorzaak is hiermee weer terug bij af. Vorig jaar oktober publiceerden Amerikaanse wetenschappers over een verband tussen het virus XMRV en CVS. De Nijmegenaren trokken het onderzoek in twijfel omdat alle testpersonen uit hetzelfde dorpje kwamen. (Kennislink.nl)



Hoofdpijn van 3D

Twintig procent van de bevolking kampt met hoofdpijn na een 3D-film vanwege verminderde oogspanning. Onderzoekers van de TU Eindhoven bedachten een leestest om te achterhalen hoe gevoelig iemand hiervoor is. De test bestaat uit het voorlezen van 3D-teksten. De onderzoekers hopen dat bedrijven als Sony en Philips de test in zullen bouwen in hun 3D-tv's, samen met een 'volumeknop voor diepte' voor optimaal kijkcomfort. (Kennislink.nl)

Wat er vandaag gebeurde in het jaar

1558
Francisco Fernandes brengt tabak naar Europa

1824
Eerste Birma-oorlog: Het Verenigd Koninkrijk verklaart officieel de oorlog aan Birma

1836
Samuel Colt maakt de eerste revolver voor massaproductie (.34-kaliber)



1966
Een BOAC Boeing 707 jet stort neer op de berg Fuji, Japan, 124 mensen komen om

1940
Jozef Stalin ondertekent het bevel om duizenden Poolse krijgsgevangenen om te brengen (Bloodbad van Katyn)

2003
16 personen, waaronder de dader, vinden de dood bij een zelfmoordaanval op een autobus in Haifa



Samengevat door de computer

Online zijn enorme hoeveelheden tekst te vinden. Het zou je soms maanden kosten als je alles over een bepaald onderwerp wilt lezen. Dan zou het handig zijn als de computer het even voor je kon samenvatten. DOOR ERICA RENCKENS

“Om een goede samenvatting te kunnen maken moet een computer de tekst volledig begrijpen en daar een globale, ingekorte representatie van geven”, aldus Iris Hendrickx. Taaltechnologen zoals Hendrickx werken al jaren aan automatische samenvattingssystemen.

„Computergegenereerde samenvattingen zijn nog niet zo goed als samenvattingen die door mensen zijn geschreven, maar ze kunnen je veel tijd en geld besparen als je grote hoeveelheden informatie moet doorzoeken.”

De methode die Hendrickx hierboven beschrijft heet abstractie. De computer moet dan de tekst daadwerkelijk begrijpen en zelf correcte zinnen formuleren. Dat is erg moeilijk voor een computer, waarvoor een tekstbestand niks meer is dan wat enen en nullen



Tekst QR code: Scan deze QR code met je mobiele telefoon voor het hele verhaal.

zonder verdere betekenis. De meeste onderzoekers gebruiken daarom de meer succesvolle extractie-methode. Bij deze methode kiest de computer de belangrijkste zinnen van een tekst en plakt deze vervolgens samen tot een nieuwe, kortere tekst.

Maar hoe weet de computer nou welke zinnen belangrijk zijn? Het samenvattingssysteem analyseert de tekst eerst op verschillende punten en geeft op basis hiervan elke zin een score. De zinnen met de hoogste score zijn het belangrijkste en maken het meeste kans om in de uiteindelijke samenvatting terecht te komen.

Analyse

Om te beginnen kan het systeem kijken naar de positie van een zin in een alinea of tekst. De eerste en laatste zin zijn meestal het belangrijkste. Vervolgens kan de computer ook nog kijken naar zogenaamde 'cue phrases'. Dit zijn zinnen die erop wijzen dat er iets belangrijks gaat volgen. Als je bijvoorbeeld ergens leest "Hiermee wil ik aantonen dat...", dan weet je dat je op moet letten: nu komt iets belangrijks!

Ook woordfrequenties kunnen een aanwijzing vormen voor een samenvattingssysteem. Woorden als 'de' en 'in' komen meestal vaak voor in een tekst, maar als het woord 'muziek' vaker dan normaal voorkomt, is de kans groot dat de tekst daarover gaat. De compu-

ter kan speciale frequentielijsten van het Nederlands gebruiken om de tekst mee te vergelijken.

Tenslotte kan de computer ook nog de tekst en de zinnen taalkundig ontleden. Uit deze ontleding ontstaat een boomstructuur. De kleinste takjes die het verst van de oorsprong afstaan zijn de zinsdelen die het makkelijkst achterwege te laten zijn. Zo kan de computer dus makkelijker de hoofdzaken van de bijzaken onderscheiden. Deze informatie over de tekst- en zinsstructuur kan de computer goed gebruiken bij het bepalen welke zinnen in de samenvatting mogen komen.

Compleet nieuws

De samenvatting die op deze manier uit je computer rolt is dus nog niet zo goed als wanneer je hem zelf had ge-

Computergegenereerde samenvattingen zijn nog niet zo goed als samenvattingen die door mensen zijn geschreven

In samenwerking met
kennislink.nl
maakt nuwring

schreven. „Maar samenvattingen schrijven is kostbaar in tijd en inspanning”, legt Iris Hendrickx uit. „Door de digitalisering zijn steeds meer informatiebronnen beschikbaar in elektronische vorm, en automatisch gegenereerde samenvattingen kunnen goed helpen met het doorzoeken en aanbieden van grote hoeveelheden informatie.”

En er zijn meer toepassingen voor deze techniek te bedenken. „Zelf vind ik multi-document summarization erg bruikbaar”, vertelt Hendrickx. „Dit is het creëren van een samenvatting voor een groep teksten over hetzelfde onderwerp.” Denk bijvoorbeeld aan verschillende nieuwsartikelen over hetzelfde onderwerp. Het samenvattingssysteem kan alle dubbele informatie eruit filteren en de artikelen samenvoegen tot één tekst. Zo hoef je niet langer talloze sites af te speuren naar de laatste details van een actualiteit, maar krijg je alles in één keer voorgeschoteld in een complete tekst.

BeterWeten: Chinese middenklasse optimistisch over de toekomst

Door de trek naar de stad en toenemende welvaart wordt de opkomende Chinese middenklasse een steeds belangrijker economische factor. Het is de kerk waar de Chinese economie op drijft. Maar hoe ziet zij haar persoonlijke toekomst op sociaal-economisch gebied? Het Europe China Institute van Nyenrode, onder leiding van prof. dr. Haico Ebbers, peilt op gezette tijden met behulp van een enquête de verwachtingen voor de toekomst bij deze groep. De vragen gaan onder meer over het functione-

ren van de overheid, het financiële systeem, sociale zekerheid, het ondernemingsklimaat en persoonlijke welvaart. Ebbers deed onderzoek bij een panel van academici, mensen uit het bedrijfsleven en ambtenaren uit de vier grootste steden van China.

De eerste meting van deze China Structural Development Monitor, een gezamenlijk initiatief van Nyenrode en pensioenuitvoerder/vermogensbeheerder Mn Services, toont dat de middenklasse behoorlijk tevreden is over de huidige situatie waarin men

nu verkeert en dat zij optimistisch is over de toekomst. De Chinese middenklasse vindt dat de overheid het over het algemeen goed doet. Volgens hen heeft de overheid bijvoorbeeld zeer goed gereageerd op de economische crisis. Toch zien zij natuurlijk ook verbeterpunten voor het overheidsbeleid, zoals het aanpakken van de inkomensongelijkheid en de problemen op milieugebied. Maar, vertelt Ebbers, de middenklasse heeft wel het gevoel dat iedereen kan meedelen in de nieuwe welvaart.

Er blijven met name op economisch gebied nog wel wensen over voor de middenklasse. In hun ogen zijn er bijvoorbeeld nog te veel barrières voor ondernemingen. Zo is het een hele toer om de benodigde vergunningen te verkrijgen en is het niet makkelijk om een onderneming te kunnen starten en om aan aan kapitaal te komen. Veel ondernemers hebben het gevoel dat zij worden achtergesteld bij staatsbedrijven, zegt Ebbers. Zij kunnen nog niet in volledige vrijheid handelen. Ook verwacht

bijna iedereen in het panel dat de wisselkoers van de Chinese munt omhoog gaat. Dat is slecht voor de export en daardoor een punt van zorg. Daarnaast worden staatspensioenen afgebouwd en moet men daarvoor nu zelf zorgdragen, wat onzekerheid met zich meebrengt.

De enquête zal ieder kwartaal door dezelfde groep mensen worden ingevuld. Hierdoor kunnen veranderingen in de opvattingen van de middenklasse op de voet worden gevolgd.