

Revolutionaire tool of lichtzinnig speeltje?

De vraag of je met een eye-opener te maken hebt of met een blingbling stukje speelgoed, dringt zich snel op als je kennis maakt met Trendalyzer, de grafiekenmanipulator van het Zweedse bedrijf Gapminder. Mooi om te zien is het zeker, en bijna nog mooier de manier waarop uitvinder Hans Rosling, verbonden aan het Zweedse Karolinska Instituut, de werking demonstreert. In een presentatie



voor TED, een Amerikaans forum voor megaknappe koppen, ondernemers en zieners (www.ted.com/index.php/talks/view/id/92) rent hij over het podium alsof hij het zelf is die op het beeldscherm de grafische bollen, lijnen en vlakken tussen de X- en de Y-as op en neer moet duwen. Dat is natuurlijk niet zo, want dat doet de software waaraan hij zeven jaar heeft gewerkt. Het principe van die software is eenvoudig, zoals meestal het geval is met revolutionaire vindingen. In een gangbare grafiek heb je twee assen en je kunt er dan ook de invloed van twee



variabelen op elkaar mee illustreren. Met wat creativiteit kan er nog een derde bij, bijvoorbeeld door de pun-

ten waar de x- en de y-waarde elkaar ontmoeten een bepaalde omvang te geven. Dan worden het bollen of lijnen waarvan de grootte ook weer een waarde aangeeft. Maar daarna zijn de mogelijkheden wel ongeveer op.

Met Trendalyzer begint het dan echter pas. Je voegt variabelen toe of je splitst variabelen en de daarbij behorende waarden, en in plaats van totale chaos of een onafzienbare reeks grafieken, krijg je een fraaie animatie te zien waarin de grafiek zich van de ene naar de volgende weergave transformeert. Trendalyzer lijkt een goede naam. Je analyseert er trends mee. Of liever: je geeft er je analyse van de trends gaaf mee weer. En in dat laatste onderscheid schuilt misschien het antwoord op de vraag: tool of speeltje?



Rosling zal er niet lang over na hoeven denken. Hij is een serieuze professor, die zich behalve met Trendalyzer bezig houdt met ontwikkelingsvraagstukken. Zijn non-profit bedrijf Gapminder heeft volgens een Zweedse nieuwssite (www.thelocal.se) als missie om 'informatie te verzamelen en verspreiden die helpt om mondiale duurzame ontwikkeling en de realisering van de millenniumdoelen te bevorderen'. Dat is iets anders dan speeltjes ontwikkelen. Vertrouwenwekkend. Maar datzelfde kan niet gezegd worden van het feit dat de Trendalyzer-software een paar maanden geleden werd opgekocht,

juist door alleseter Google. Wat moet die met duurzame ontwikkeling en de millenniumdoelen, of met wetenschappelijke analyse *for that matter*? Vertrouwen wekt trouwens evenmin Gapminders eigen claim dat 'Trendalyzer saaie data omtovert in aangename interactieve animaties'.

Saaï...saaï? Sinds wanneer zijn



data saaï? Data stromen door de levensader van de wetenschap en die dienen geanalyseerd te worden met alle mitsen en maren, zorg en voorbehoud die daarbij horen. Aan wie dat saaï vindt kun je die data beter niet toevertrouwen. Je moet dan ook je hart vasthouden voor de eerste analyse die tot stand is gekomen *met behulp van* Trendalyzer, die juist een schitterend hulpmiddel biedt om mitsen en maren weg te poetsen bij het presenteren.

Nee. Trendalyzer is geen analysegereedschap. Trendalyzer is een presentatiegereedschap. Je presenteert er een analyse mee die je eerder hebt gemaakt, en die je hopelijk goed hebt gemaakt. Of dat laatste het geval is, dat weten de ademloos geboeide kijkers niet. Maar dat wisten ze ook al niet toen ze het nog met tweedimensionale printgrafieken moesten doen. Met grafieken kan je prachtig liegen, zoals Disraeli in de negentiende eeuw al opmerkte. Trendalyzer voegt ook in dat opzicht een dimensie toe. Maar gelukkig geldt net als in de tijd van de Engelse politicus nog altijd: liegen mag niet. (MdG)

Column
Edward Vanhoutte

Humanities Computing

Men vraagt mij wel eens wat ik doe voor de kost. Omdat een onderzoekscentrum runnen bij mijn gemiddelde gesprekspartner meteen het beeld oproept van gelabojaste wetenschappers met pipet, erlenmeyer en bunsenbrander, en de Koninklijke Academie in Gent al te vaak wordt verward met die hogere kunstopleiding in dezelfde stad, zeg ik dan maar voor het gemak dat ik les geef aan de universiteit. Daar kan men zich wel wat bij voorstellen. Grote aula's met ex cathedra docerende proffen, bijvoorbeeld, en straffe verhalen over mondelinge examens. Ik vertel er dus maar direct bij dat ik jaarlijks gemiddeld tussen de vijf en de vijftien studenten heb en dat die geen examen hoeven af te leggen maar op basis van permanente evaluatie en een groepsproject worden beoordeeld. En dan komt de onvermijdelijke vraag, namelijk wat ik dan wel doceer. Humanities Computing. Gegarandeerd valt het gesprek dan even stil.

Het vervolg verloopt gewoonlijk volgens een van de vier volgende scenario's. Scenario één: mijn gesprekspartner knikt begrijpend, vermijdt vervolgens dieper op het onderwerp in te gaan of wandelt doodleuk weg. Scenario twee: mijn gesprekspartner deelt mij mee dat een vriend, kennis of familielid ook in de computer- of informaticabusiness zit, en dat daar wel wat mee te verdienen valt. Scenario drie: mijn gesprekspartner begint prompt over problemen met tekstverwerkers, printers of draadloze netwerken, waarop ik het eerste scenario toepas. Scenario vier: mijn gesprekspartner toont een zeldzame belangstelling en vraagt wat Humanities Computing dan wel is. En dan word ik gegarandeerd even stil.

Een goede Nederlandse vertaling bestaat er immers niet. 'Alfa-informatica' zegt de Vlaming niets. 'Informatica voor de humane wetenschappen' of de letteren, of 'digitale geesteswetenschappen' dekken de lading niet helemaal, en 'informatiekunde' is dan weer net niet wat Humanities Computing is. Een goede definitie voor het veld bestaat evenmin. Verder dan een hermetisch 'de applicatie van computationele technieken en methoden op de humane wetenschappen, waarbij theorie, experiment en toepassingen worden geïntegreerd' kom ik nauwelijks.

Humanities Computing is een breed veld van activiteiten, waarvan mijn specialiteit slechts een onderdeel is. Het grote probleem van Humanities Computing is dat er geen methodologisch besef is waarbij het veld zichzelf kan legitimeren. Nochtans staat of valt de interne communicatie binnen en de externe communicatie over het veld met een dergelijk besef. Zonder een duidelijke beschrijving van de methodologische netwerken die Humanities Computing uitmaken, zal dit activiteitenveld nooit in staat zijn om aan een groot publiek duidelijk te maken wat het is en doet.

Willard McCarty schrijft in zijn boek Humanities Computing (Palgrave, 2005) dat de graad van bekwaamheid om in simpele termen te communiceren recht evenredig is met de aandacht en de ontvankelijkheid voor onze wetenschappelijke activiteiten. Zolang dat diepere methodologische besef niet expliciet is geformuleerd, murmel ik mijn gesprekspartner iets toe over teksten en computers en probeer ik het daaropvolgende weekend toch weer een of ander draadloze netwerk of printer aan de praat te krijgen.

Edward Vanhoutte is coördinator van het Centrum voor Teksteditie en Bronnenstudie van de Koninklijke Academie voor Nederlandse Taal- en Letterkunde en doceert Humanities Computing aan de Universiteit Antwerpen.

www.surf.nl/download/Factsheet_DARE_Entreeopmaat.PDF

P. Dekker, A. van der Horst, H. Kox e.a., *Markplaats Europa, Vijftig jaar publieke opinie en marktintegratie in de Europese Unie. Europese verkenning 5*. Den Haag, SCP/CBS, 2007. ISBN 978-90-377-030-54

De verkenning is een gezamenlijke publicatie van het CPB en het SCP en

komt jaarlijks tot stand in opdracht van het ministerie van Buitenlandse Zaken. In deze vijfde verkenning wordt nader ingegaan op de publieke opinie in de lidstaten en de ontwikkeling van de interne markt in de afgelopen vijftig jaar. www.scp.nl/publicaties/boeken/9789037703054.shtml

Gelezen

Sociale samenhang in beeld, het SSB nu en straks. Voorburg, Centraal Bureau voor de Statistiek, 2007. ISBN 978-90-357-1942-2

Dit boek bevat een verzameling opstellen, gebaseerd op de lezingen en referaten die op 1 december 2005 gehouden zijn tijdens een symposium over het Sociaal Statistisch Bestand (SSB) in Naturalis te Leiden. Het SSB maakt gebruik van een breed scala aan registraties in plaats van vele enquêtes. Hierdoor ontstaan veelal integrale bestanden over verschillende maar verwante onderwerpen, die op persoonsniveau kunnen worden gekoppeld. Op die manier kan het CBS voor het eerst samenhangende uitkomsten op alle terreinen van maatschappelijk belang opstellen en daarmee de sociale samenhang in beeld brengen. www.cbs.nl

Mishne, G.A., met bijdragen van M. de Rijke. *Applied text analysis for blogs. Amsterdam, Instituut voor Informatica UvA, 2007. ISBN 9789057761638* De voornaamste vraag die de auteurs proberen te beantwoorden is: hoe kunnen de karakteristieke eigenschappen van blogs worden gebruikt om effectief kennis te vergaren uit de

blogspace? Het betreft de feitelijke informatie rondom een persoon zoals interesses, culturele voorkeuren en leefomgeving, maar ook de niet-feitelijke informatie zoals emoties, meningen, en stemmingen. Dit is de informatie die grotendeels uniek is voor de blogspace.

De auteurs laten zien dat deze twee soorten van kennis inderdaad effectief vergaard kunnen worden uit blogs. Hierbij wordt gebruik gemaakt van een reeks methoden, veelal gebaseerd op statistische taalmodellen en tekstclassificatie. Deze vorm van informatievergaring uit een groot aantal blogs (gebruikmakend van tekstanalyse) legt informatie bloot die niet beschikbaar is in individuele blogs. <http://dare.uva.nl/record/219974>

Wrzesinska, M.: *Handling complexity and change in grid computing. Amsterdam, Proefschrift Vrije Universiteit, 2007*

Het resultaat van deze dissertatie is een programmeeromgeving die het eenvoudiger maakt om gridapplicaties te schrijven. Problemen die zich bij grids vaak voordoen zoals rondom communicatie, de taakverdeling tussen processoren of foutbestendigheid

worden automatisch opgelost door de programmeeromgeving. De toepassingen die worden ingevoerd binnen deze programmeeromgeving kunnen efficiënt draaien in gridomgevingen en zijn bestand tegen situaties die kenmerkend zijn voor dat soort omgevingen, zoals wegvallende processoren of veranderende belasting op processoren en netwerkverbindingen. <http://hdl.handle.net/1871/10756>

Entree op maat, meerwaarde van repositories. SURFfoundation, 2007.

Alle Nederlandse universiteiten, de KNAW, NWO en de KB hebben in de jaren 2003-2006 een gezamenlijk netwerk van repositories, digitale schatkamers, opgezet. Hiermee maken zij zo veel mogelijk resultaten van wetenschappelijk onderzoek voor iedereen beschikbaar via internet. Dit gezamenlijke DARE-programma, gecoördineerd door Stichting SURF, heeft Nederland bij de vernieuwing van de wetenschappelijke informatievoorziening een toonaangevende positie bezorgd. Het factsheet 'Entree op maat' geeft een kort overzicht van de resultaten van het DARE programma en de meerwaarde van repositories.