

I/O Vivat

Jaargang 22, nummer 3,
April 2007
ISSN: 1389-0468
Oplage: 1800

I/O Vivat is het periodiek orgaan van Inter-Actief (de studievereniging voor Informatica, BedrijfsInformatie Technologie en Telematica) en ENIAC, de alumnivereniging voor Informaticastudenten van de Universiteit Twente.

Redactie

Alexander Spannenburg (voorzitter),
Jan-Willem Dijkshoorn, Michel Jansen,
Bram van de Laar, Tom Palsma,
Mark Timmer

Gastschrijvers

Sjoerd van der Spoel, Ronnie Scholten,
Sensami, Gerrit van der Hoeven,
Ander de Keijzer, Eelco Eerenberg,
Berend van den Brink, Heidi Vogel,
Ronald Brockmann, Erwin Boutsma
(Technisch Weekblad)

Drukker

Van den Bosch & Fikkert



Adressen

E-mail: vivat@inter-actief.net
Telefoon: 053-4893756
Fax: 053-4894571
Postadres: Inter-Actief
Postbus 217
7500 AE Enschede
Internet: www.inter-actief.net

Dank aan alle inzenders van kopij.
De studievereniging wil de adverte-
rende bedrijven bedanken voor de
goede samenwerking.

Copyright © 2007 Studievereniging
Inter-Actief.

Foto voorkant: Overturned Quantum
Universe, CS4FN, Queen Mary Uni-
versity of London, <http://www.cs4fn.org/quantum/quantumuniverse.php>

Full colour

“Het nut van een redactioneel” was de titel van mijn vorige redactioneel en alhoewel we hier niet in de I/O Vivat-vergaderingen over hebben gedebatteerd, is er uit de vorige editie een sprankje nut gekomen. In tegenstelling tot normaal was editie 22.2 geheel in full colour. Veel mensen viel dit ook op en veel complimentjes kwamen daarmee onze kant op. Altijd leuk! Zo'n groot en onaangekondigd lay-outverschil behoeft naderhand natuurlijk uitleg en daarmee toont de redactioneel weer een van zijn doelen. Het ging allemaal als volgt.

Voordat de I/O Vivat in grote hoeveelheden wordt gedrukt, maakt de drukker altijd eerst een proefdruk zodat we eventuele (schoonheids)foutjes nog kunnen verbeteren. Nadat we de vorige editie digitaal hadden aangeleverd bij de drukker duurde het opmerkelijk lang voordat we nieuws kregen over de proefdruk. Er waren namelijk wat problemen met een nieuwe machine die de drukker in gebruik had genomen, namelijk één die in plaats van vier pagina's nu acht pagina's op één vel papier drukt. Bij de lay-out was echter voor de indeling van de kleurenpagina's rekening gehouden met de oude machine. Verdere details zal ik u besparen, maar na vele pogingen koos de drukker voor de makkelijke weg en besloot om de gehele I/O Vivat in full colour te drukken.

Nu de meerwaarde van het full colour drukken van de I/O Vivat zo tastbaar is geworden, besloten we samen met het Inter-Actief-be-stuur om te onderzoeken of een full-colourversie financieel haalbaar was. Na contact met verschillende drukkers en een goede offerte van onze vertrouwde drukker (Van den Bosch & Fikkert) is er besloten om vanaf nu de I/O Vivat in full colour te laten drukken. Erg goed nieuws! Want het ziet er niet alleen veel mooier uit, het lay-outwerk is nu ook aanzienlijk makkelijker geworden. Voorheen moesten de artikelen met mooie kleurafbeeldingen op hetzelfde vel waar een full-colouradvertentie op stond gepositioneerd worden. Bovendien moesten veel afbeeldingen omgezet worden naar de bekende blauwe I-A-kleur. Nu zijn we daar niet meer van afhankelijk en kunnen we de artikelen rustig indelen aan de hand van de nieuwe tab-indeling.

Nieuw is ook de papiersoort waar de I/O Vivat vanaf deze editie op gedrukt zal worden. Het gramsgewicht van de cover is nu hetzelfde als van het binnenwerk, maar deze is wel gestegen ten opzichte van de vorige edities. Omdat de drukker hier veel meer verstand van heeft dan wij, hebben we onze contactpersoon bij de drukker gevraagd hierover te schrijven. In zijn bijdrage wordt ook het drukprocedé uit de doeken gedaan.

Nog een nieuwtje is de nieuwssectie. In deze sectie willen we nieuwsberichten plaatsen die interessant zijn voor onze lezers. Aangezien de I/O Vivat een laagfrequent verschijnend periodiek is zullen we alleen interessante of opmerkelijke berichten plaatsen waarvan de nieuws-waarde wat langer houdbaar is. In de volgende editie willen we deze sectie twee pagina's laten vullen. Goede suggesties voor deze nieuwe sectie zijn natuurlijk altijd welkom.

Ten slotte nog een heuglijk feit, want vanaf deze editie komt Michel Jansen het team versterken, iets waar we als redactie natuurlijk erg blij mee zijn.

Rest me u weer veel leesplezier te wensen! ■

Alexander Spannenburg, voorzitter redactie I/O Vivat

RECTIFICATIE

In de vorige editie is per abuis een verkeerde versie van het artikel 'Web 2.0' opge-nomen. Onze excuses hiervoor. De redactie.

Inhoud



Aan het werk in de VS (pag. 9)



Inaudible Melodies (pag. 22)

“Drukkerij Van den Bosch & Fikkert te Almelo: een uiterst moderne offsetdrukkerij die al sinds 1932 (75 jaar) actief is als drukkerij, van het visitekaartje tot hoog kwalitatief full-colourdrukwerk in uiteenlopende oplages”

4

“Ambient Intelligence, of kortweg Aml, is een visie op de nabije toekomst: de computer is alom vertegenwoordigd in de omgeving, zonder dat je echt in de gaten hebt dat ze er zijn.”

7

“Het idee dat er een relatie moet bestaan tussen muziek en kleur is al minstens zo oud als de oude Grieken. Filosofen als Aristoteles en Pythagoras speculeerden al over een verband tussen de ‘musical scale’ en het kleurenspectrum van de regenboog.”

22

Inter-Actief:

- 3 Voorwoord
- 4 Van de drukker
- 7 Sensami

Afdeling INF:

- 8 Wat zoekt Van der Hoeven...
- 9 Trio
- 14 The Interactive Virtual Storyteller
- 22 Inaudible Melodies
- 25 Improving recommender systems

ENIAC:

- 30 Van de penningmeester
- 31 Bestuur/actievelingen gezocht!
- 32 Wie, Wat, Waar Heidi Vogel
- 34 Wie, Wat, Waar Ronald Brockmann
- 35 ENIAC sponsort!

Overigen:

- 36 Nieuws
- 39 Op bezoek bij... TNO



Terwijl ik dit schrijf, vallen voor het eerst zonnestralen de Inter-Actief-kamer binnen. Voor het eerst sinds midden jaren zestig kan genoten worden van de vroege maandagochtendzon, onder het genot van een kopje koffie, op de roze bank. Bijna veertig jaar lang heeft Hal D, oftewel de Matrix, prominent in beeld gestaan, gezien vanuit wat de laatste vijf jaar de Inter-Actief-kamer is geweest.

Na (jaren)lang gesteggel is eindelijk het verouderde, verlaten, volgens velen oerlelijke gebouw tussen de Waaier en de Horst dan toch echt gesloopt. En met het verdwijnen van de zwart/grijs/witte doorn in het oog kan er nu eindelijk genoten worden van een panoramisch uitzicht op Horst, Matrix-resten en de mooie natuur op de campus. Heel verfrissend om op een rustige maandagochtend eindeloos naar buiten te kunnen staren, met ook daadwerkelijk uitzicht. Hoewel de eerste plannen voor de sloop van het laatst overgebleven deel van het oude Hallencomplex, waar de Waaier ook bij hoort, al voor 2000 uitgevoerd hadden moeten worden, is het nu dus pas eindelijk zover. En pas als je uitzicht hebt, merk je opeens hoe belangrijk het is. Gelijk als je de kamer in komt lopen, of het nu 's ochtends, als je net de deur open doet, of later op de dag na een praatje op de gang, wordt je overvallen door een breedbeeld

blikveld over een uitgestrekte vlakte met op de achtergrond de grote toren van de Horst. Je kent de kamer haast niet meer terug, het is als oogkleppen die opeens afvallen. De kamer, die toch bepaald niet klein is, lijkt opeens stukken groter. Zo merk je hoe belangrijk het is om uitzicht te hebben. Te hebben, maar (helaas?) niet om te houden. Het uitzicht voor de toekomst van de huisvesting van Inter-Actief

is niet deze betonnen vlakte met de vreemd uitstekende metalen resten van het geraamte van de oude kolos die Matrix heette, maar een geheel nieuwe ruimte, met een compleet ander uitzicht. De toekomst lijkt ons een nieuwe kamer in de voormalige INF/TO-bibliotheek in de Zilverling, tegenover de hoofdingang van de Waaier. Alweer een plan dat al jaren speelt, en de voorgangers van mijn voorgangers waarschijnlijk al jaren slapeloze nachten oplevert. Hoewel er nu concrete plannen getekend zijn voor een indeling, met een grote algemene ruimte te gebruiken voor alle studenten (maar impliciet vooral voor de leden van Inter-Actief, Abacus en Scintilla) met aan de randen daarvan kantoren voor de drie studieverenigingen, is dit niet de eerste keer dat de plannen zover zijn. Zeker één keer eerder waren de meubels al uitgezocht en konden de bouwvakkers aan de slag. Maar helaas, tot nu toe is er nog geen resultaat. In de tussentijd neem ik mij maar voor om, met mijn kopje koffie, op de bank nog maar even te genieten van het uitzicht, zolang het nog kan. ■



Sjoerd van der Spoel

Voorzitter I.C.T.S.V. Inter-Actief

Van de drukker

De nieuwe I/O Vivat

Introductie

Mijn naam is Ronnie Scholten, ik ben werkzaam als technisch adviseur en vertegenwoordiger bij **Drukkerij Van den Bosch & Fikkert** te Almelo: een uiterst moderne offsetdrukkerij die al sinds 1932 (75 jaar) actief is als drukkerij van het visitekaartje tot hoog kwalitatief full-colourdrukwerk in uiteenlopende oplages, en tevens diverse deelspecialismen in huis heeft, waaronder blisterkartonnage alsmede vervaardiging van luxe verpakkingen. Er is mij gevraagd een stukje te schrijven over de technische kant van de vervaardiging van de I/O Vivat.

Offsetprocédé

De technologie van het offsetdrukken bestaat al meer dan honderd jaar. De term "offset" verwijst naar het feit dat de drukcilinder met de te drukken informatie niet in aanraking komt met het papier.

Stap 1: Drukformvervaardiging

De eerste stap van het offsetdrukprocédé bestaat uit het aanmaken van een drukvorm op basis van het originele ontwerp. Vroeger gebeurde dit middels filmmateriaal (vraag maar aan de redactie hoe dat eruit ziet. Red.), tegenwoordig rechtstreeks vanuit de computer. De drukvorm (offsetplaat) wordt tegenwoordig direct vanuit de computer belicht en ontwikkeld, waardoor er geen film en fotografische vloeistoffen (filmontwikkelaar en fixeer) meer gebruikt hoeven

te worden, wel zo vriendelijk voor ons milieu en natuurlijk kostenbesparend. De offsetplaat is bedekt met een lichtgevoelige emulsie die inkt opneemt. Deze wordt blootgesteld aan licht waardoor de plaat wordt ontwikkeld en de emulsie enkel op de belichte plaatsen achterblijft.

Stap 2: Bevochtigen

Bij offsetdrukken wordt gebruikgemaakt van het water-versus-olie-principe. De plaat wordt eerst met water bevochtigd. Hierdoor blijven de zones zonder afbeelding vochtig, waardoor deze beter bestand zijn tegen de inkt op basis van olie die in dit proces worden gebruikt. De plaat kan nu worden geïnk.

Stap 3: Het inkten of offsetdrukken

Nadat de offsetplaat is ontwikkeld, wordt deze vastgehecht aan een cilinder op de drukpers. Terwijl deze ronddraait, brengt de plaat de geïnkte afbeelding over op een rubberdoek

waardoor een afbeelding in spiegelbeeld ontstaat. Het doek brengt deze afbeelding op zijn beurt over op het papier, waardoor deze weer wordt gespiegeld en juist wordt weergegeven. Dit is de uiteindelijke afdruk.

Stap 4: Meerdere passages om alle kleuren te krijgen



Ronnie Scholten

contactpersoon Drukkerij Van den Bosch & Fikkert



Om meerdere kleuren te drukken wordt in elk kleurstation (druktoren) de inktkleur veranderd, terwijl de afdruk door de pers loopt. Moderne persen hebben niet alleen vier kleurstations, maar brengen ook een vernislaag aan om vingerafdrukken op het papier te vermijden.

Drukkerij Van den Bosch & Fikkert werkt onder andere met twee van dit soort vijf-kleuren-machines, waardoor het mogelijk is full colour en vernis (of eventueel een steunkleur) in slechts één machinedoorgang te behalen. Tevens beschikken beide

colour: het bestuur heeft echter nu ook gekozen voor een selfcover wat eigenlijk niets meer of minder betekent dan dat het papier van de omslag hetzelfde gramsgewicht heeft als het binnenwerk. Wel is er gekozen om het gramsgewicht te verhogen naar 135 g^r/m² om voldoende cachet en body te behouden. ■

Mocht u naar aanleiding van dit artikel nog vragen hebben neemt u dan contact op met DRUKKERIJ VAN DEN BOSCH & FIKKERT
telefoon: 0546-813017
www.druk-bosfik.nl

“Fijne boek- en tekenpapieren zijn soms ook met lommen veredeld.”

machines over ieder een uniek principe: de ene is een kartongeschikte machine. Waar het bij de meeste drukkerijen niet mogelijk is papier te verwerken met een gramsgewicht hoger dan 350 g^r/m² kan deze machine wel 800 of 900 g^r/m² karton of bord aan. De andere machine is zelfs uniek in de wereld: deze is uitgerust met een rotatieve stansunit en is tevens voorzien van restmateriaalafzuiging waardoor een schoon full-colourproduct compleet gestanst in de uitleg komt (denk hierbij aan enorm gecompliceerde maar ook simpele stansvormen, onder andere voor verpakkingen).

Papiersoort I/O Vivat

Het MC-papier (Machinecoated) dat gebruikt wordt voor de vervaardiging van de I/O Vivat valt onder de categorie houtvrije papiersoorten, bestaande uit cellulose. Fijne boek- en tekenpapieren zijn soms ook met lommen veredeld. Houtvrij papier heeft ten opzichte van houthoudende soorten verschillende voordelen: de vezels zijn langer, elastischer en het papier is vaster en beter voor veredeling geschikt. Bovendien vergeelt het minder snel. De papierindustrie heeft tevens hulpstoffen ontwikkeld die het papier beschermen tegen het ouder worden. Deze editie van de I/O Vivat is ook weer compleet full



ADVERTENTIE

Océ

Na het succes van het eerste landelijk informaticasymposium, georganiseerd vanuit Enschede, wordt er op dinsdag 13 juni 2007 opnieuw een dergelijk, groot symposium georganiseerd, ditmaal door Utrechtse studenten informatiekunde en informatica. Het thema is 'Sensami', een samentrekking van 'Sense Ambient Intelligence'



Sensami

Experience the world of Ambient Intelligence!

Ambient Intelligence, of kortweg AmI, is een visie op de nabije toekomst: de computer is alom vertegenwoordigd in de omgeving, zonder dat je echt in de gaten hebt dat ze er zijn. Aangezien AmI een opkomend concept is en naar alle waarschijnlijkheid zeer belangrijk gaat worden in de ICT-wereld was dit een goed onderwerp voor het vervolg-symposium van de Stichting Nationaal Informatica Congres, beter bekend als SNiC. De titel van het symposium: Sensami.

Dinsdag 13 juni wordt een geheel verzorgde dag (inclusief lunch en diner), als je tenminste naar Apeldoorn komt! 's Ochtends staan er op verschillende locaties in Nederland bussen klaar om jullie naar congressentrum Orpheus te brengen. Eenmaal daar word je verwelkomd met een hapje en een drankje. Om 11:30u zal het symposium worden geopend met een lezing om er even goed in te komen. Daarna beginnen drie verschillende *tracks*, dus voor ieder wat wils! De tracks zijn ieder goed voor vier lezingen, verspreid over de hele dag.

Bottlenecks and pitfalls

Elke nieuwe techniek ondervindt in eerste instantie veel problemen in de ontwikkeling. In deze track worden zulke problemen besproken, alsmede mogelijke oplossingen. Ook wordt er

gekeken naar eventuele toekomstige problemen.

Business and society

Wat zijn de gevolgen van AmI in de maatschappij? Zal men deze technologie zomaar accepteren? De gevolgen met betrekking tot privacy zijn niet gering: hoe wordt de privacy gegarandeerd en wat voor beveiligingen worden er gebruikt? Daarnaast is het interessant om te kijken wat voor producten er nu op de markt zijn en wat je binnenkort in je eigen kamer kan hebben!

Science and technology

De techniek en het onderzoek achter AmI is voor zowel informatica- als informatiekunde-studenten erg interessant. In deze track zal het alleen maar over zulke onderwerpen gaan: de keiharde technologie is het basis-onderwerp.

Uiteraard mag je je eigen programma samenstellen uit de tracks: verschillende lezingen uit verschillende tracks kunnen de dag extra leuk en interessant maken.

Naast de lezingen zijn er ook workshops: deze zijn zo opgezet dat er een leuke en actieve omgeving ontstaat waarin je zelf aan de slag kan met het product of met het onderwerp, erg leuk en leerzaam dus! Wees er wel snel bij, want er kunnen maar een

beperkt aantal bezoekers deelnemen aan de workshops.

Naast deze theoretische en praktische diepgang is er ook ruim tijd om je netwerk-skills weer eens even te testen op de informatiemarkt. Ruim vijftien bedrijven zijn daar aanwezig om je alles te vertellen wat zij nu precies doen met AmI en natuurlijk om je voor te lichten over een vette carrière bij dat ene bedrijf!

En dat alles speciaal voor Inter-Actief-leden voor maar 7 euro 50!

Voor meer informatie en de laatste nieuwtjes over het symposium ga je naar www.sensami.nl of vraag ernaar bij het bestuur van je eigen studievereniging. Inschrijven kan vanaf 22 maart, via de website uiteraard! ■

Wat zoekt Van der Hoeven?

Leraren?

Aan sommige Nederlandse universiteiten verwachten de studenten weinig zorg van hun instelling. Op een moeilijke vraag krijg je er niet altijd een antwoord, maar dat is je eigen probleem. Krijg je wel een antwoord, maar het blijkt fout, dan is dat ook je eigen probleem. Zijn dat slechte universiteiten? Nee, maar ze zijn wel heel anders dan de Universiteit Twente. Studenten van de UT rekenen op meer zorg van hun instelling. Maar ze zijn nogal ontevreden, zo hoor ik in bestuurlijke kringen. Het krijgen van de juiste antwoorden op je vragen wordt steeds moeilijker.

Ik zie een aantal oorzaken. Bijvoorbeeld dat er vaak helemaal geen antwoord op de gestelde vraag is. En de instantie die het antwoord zou moeten hebben, denkt dat iemand anders het zou moeten hebben. Of twee instanties hebben een tegenstrijdig antwoord, dat ze op verschillende plaatsen hebben gepubliceerd. Zeg ik nu dat allerlei mensen hun werk niet goed doen? Nee. Ik denk dat wij het onszelf veel te moeilijk hebben gemaakt, met onze tientallen overgangsregelingen, organisatieveranderingen en uiteenlopende ondersteunende systemen.

Ik pleit in deze context voor kwaliteit door soberheid. Doe dingen eenvoudig, en doe ze vervolgens met aandacht en meer dan goed. Is dat "de student centraal", waar het College van Bestuur op inzet? Hm. Het

past een universiteit de student centraal te stellen, en als er één universiteit in Nederland is die zich daarom bekommert, is het de UT. Dat moeten we zo houden. Maar dat helpt mij niet zo vreselijk bij de analyse en de oplossing van het probleem. En laten we vooral ook duidelijk zeggen dat we het niet mooier kunnen maken dan het is.

Dit pleidooi voor soberheid is mijn reactie op de complexiteit die ik waarneem, en die we onvoldoende weten te beheersen. Maar het is erger. We zijn zo klein dat we wel sober móeten zijn. En die kleinheid is een uitdaging die in alle discussies terugkomt. Hier opnieuw.

Scholieren hebben problemen met bèta en techniek. Waarom zou je in die richting gaan studeren? De wereld wordt al overspoeld door techniek, daar hoeft je echt niet iets aan bij te dragen. Als je dat toch wilt, dan moet je er hard voor werken, begrip komt niet vanzelf. Wie kiest daarvoor? En maak je die keuze wel, dan voeg je je bij een club waar weinigen bij willen horen. Dat nooit! Of is dat verbeterd sinds Nederland is gegrepen door de Snippes?

Om te beginnen moeten alle bachelors maar eens inspirerende leraren worden! We moeten toch ergens beginnen. ■

Op 1 maart j.l. werkte Gerrit precies 25 jaar bij de faculteit/afdeling Informatica: de redactie wil hem op deze plek nogmaals van harte feliciteren met dit benglijke feit en toewensen dat we nog maar lang van zijn columns mogen genieten.



Gerrit van der Hoeven
opleidingsdirecteur



Nadat Ander de Keijzer is meegeweest met de studiereis Kryptos is hij tijdelijk aan het werk gegaan bij Stanford University, alwaar hij aan het Trio-project heeft gewerkt. Hier doet hij verslag.



Trio

Ik woon nu vier maanden in Mountain View, een stad in de San Francisco Bay Area in California op ongeveer tien kilometer van Stanford University. Ik werk daar in de InfoLabgroep van Jennifer Widom aan het Trio-project, een nieuw soort database voor data, uncertainty en lineage. Elke morgen rijd ik van Mountain View via Palo Alto naar Stanford en voordat ik Gates (het gebouw waar informatica zit) binnenga, haal ik eerst een goeie kop koffie bij Bytes. Goede koffie is moeilijk te vinden in de VS, dus ik heb geluk dat er een echt goede koffiezaak recht tegenover Gates zit. Afgezien van de koffie zit je qua eten wel goed in de VS. Het beste restaurant in de buurt, The Cheesecake Factory, zit elke dag vol. De gemiddelde wachttijd is een uur als je niet gereserveerd hebt, maar de cheesecake is dan ook geweldig.

Er werken op dit moment acht mensen aan Trio. Iedereen heeft zo zijn eigen onderdeel om aan te werken, maar er wordt ook erg veel samengewerkt. Overdag wordt er veel overlegd. Er is een duidelijke open-deur-mentaliteit. Bovendien worden er per dag gemiddeld tien tot twintig e-mailtjes rondgestuurd over de voortgang, geconstateerde problemen en nieuwe ideeën. Als iedereen om ongeveer 18:00u naar huis gaat, is er een korte pauze waarin er niets

lijkt te gebeuren maar daarna gaat het werk gewoon door, soms tot diep in de nacht.

Op vrijdag is er een InfoLunch, een lunch met alle leden van de InfoLabgroep. Tijdens de lunch worden reisverslagen gegeven, bijzondere gebeurtenissen besproken, maar er wordt vooral veel gelachen. Na afloop geeft een van de InfoLab-leden een informele presentatie over het huidige onderzoek. Het is bij deze presentatie absoluut verboden om slides te gebruiken. Alles moet dus ter plekke op een whiteboard worden geschreven, dit om het informele karakter te benadrukken. Bovendien ontstaat er op deze manier een interactieve discussie.

Een database voor Data, Uncertainty en Lineage

Iedereen weet wel wat een database is. Een database is feitelijk niets anders dan een grote bak met informatie en met behulp van het Database Management System (DBMS) kun je op eenvoudige en generieke wijze informatie in deze bak zetten, en belangrijker nog, er ook weer uithalen. Databases worden dan ook overal gebruikt om informatie op te slaan. Denk aan de klantenadministratie bij een bedrijf, de studentenadministratie bij een universiteit, patiëntgegevens in een ziekenhuis, maar ook een postzegelverzameling

bij iemand thuis. Alle informatie die over de klanten, studenten, patienten of postzegels bekend is, kan worden opgeslagen in een database. Natuurlijk moet er, zeker bij de eerste drie voorbeelden, op worden gelet dat als informatie wijzigt, bijvoorbeeld omdat iemand verhuist, de informatie in de database wordt aangepast. De database moet dus een beschrijving zijn van de huidige staat van (een deel van) de wereld.

In de afgelopen jaren is het aantal databronnen echter drastisch gestegen. Verschillende databronnen hebben informatie over dezelfde entiteiten. Deze informatie kan gelijk zijn, overlap vertonen, maar ook geheel verschillend zijn. Als iemand nu alle beschikbare informatie van een entiteit wil hebben, moeten deze databronnen dus worden geïntegreerd. Helaas is het integreren van informatie erg ingewikkeld en tijdrovend en ook nog zeer onnauwkeurig. Vaak wordt bij conflicten tussen de verschillende databronnen een keuze gemaakt uit de verschillende mogelijkheden. Als deze keuze later verkeerd blijkt te zijn, is de correcte data inmiddels mogelijk niet meer beschikbaar. Behalve een toename in databronnen is ook de soort data soms anders dan vroeger. Sensordata bijvoorbeeld kan ook worden opgeslagen in een database. Sensoren zijn echter niet altijd even betrouwbaar of nauwkeurig en

dus is ook de opgeslagen data dat niet, terwijl in de vorige voorbeelden de data altijd betrouwbaar en nauwkeurig was, al is dit natuurlijk wel afhankelijk van de persoon die de informatie heeft ingevoerd.

Aan Stanford wordt op dit moment gewerkt aan een nieuw soort database, Trio genaamd, waarbij niet alleen data kan worden opgeslagen maar ook onzekerheid (uncertainty) en afkomst (lineage) van de data. Met deze twee nieuwe aspecten kunnen we de eerder genoemde problemen voorkomen. Als we informatie integreren kunnen we, in plaats van te kiezen voor een alternatief, simpelweg alle alternatieven opslaan en bij elk alternatief eventueel een kans aangeven. Hetzelfde geldt voor de sensorinformatie waarbij we nu ook de nauwkeurigheid van de sensoren kunnen opslaan. Om het nieuwe datamodel en de query-taal toe te lichten, gebruik ik een eenvoudig 'crime-solver'-voorbeeld.

Crime Solver

We gaan uit van misdaden waarbij er een getuige is en een bestuurder van een auto. Voor de getuigen slaan we de naam van de betreffende getuige op en de kleur en het merk auto dat deze getuige gezien heeft. Van bestuurders slaan we de naam op en de kleur en het merk auto dat deze persoon heeft. Het schema van de database ziet er dus als volgt uit:

```
Saw(witness, color, car)
Drives(person, color, car)
```

Ook in Trio gebruiken we CREATE TABLE om een tabel aan te maken. Omdat we in Trio ook willen kunnen werken met gewone relationele data, voegen we het woord 'trio' toe om aan te geven dat het gaat om een trio-tabel. De opdracht om de SAW-tabel te maken is dan

```
CREATE TRIO TABLE SAW (witness
    varchar, color varchar, car
    varchar)
```

In plaats van een normale database, is Trio een Uncertainty and Lineage DataBase (ULDB). Dit houdt in dat

in plaats van tuples, 'x-tuples' worden opgeslagen. Een x-tuple is een set van alternatieve tuples. Stel dat zich een getuige, Amy, meldt bij de politie. Amy is alleen niet zeker over welke auto ze gezien heeft. Het was of een blauwe Honda, of een rode Toyota, maar het kan ook zijn dat ze helemaal niets gezien heeft. In Trio geven we dit als volgt weer:

SAW

witness	color	car	
Amy	Blue	Honda	?
Amy	Red	Toyota	?

De SAW-tabel hierboven heeft één x-tuple met twee alternatieven. Een vraagteken na een alternatief geeft aan dat het x-tuple mogelijk niet bestaat. Dit ene x-tuple definieert dus drie mogelijke werelden. Een eerste wereld (Amy, Blue, Honda), een tweede wereld (Amy, Red, Toyota) en een derde wereld waarin het x-tuple niet bestaat. We breiden de SAW-tabel

"In plaats van een normale database, is Trio een Uncertainty and Lineage DataBase (ULDB)."

uit met een aantal nieuwe getuigen. Het resultaat is hieronder weergegeven. De getallen achter de alternatieven zijn om de alternatieven later uniek en eenvoudig te kunnen onderscheiden.

SAW

witness	color	car	
Amy	Blue	Honda	? 11
Amy	Red	Toyota	? 12
Betty	Green	Mazda	13
Betty	Green	Toyota	14
Betty	Green	NULL	15
Cathy	Red	Acura	? 16
Diane	Red	Toyota	17
Diane	Blue	Toyota	18

De DRIVES-tabel ziet er ongeveer hetzelfde uit als de tabel SAW. De DRIVES-tabel heeft drie x-tuples. Het eerste x-tuple heeft twee alternatieven, het tweede x-tuple heeft een alternatief en het derde x-tuple heeft twee alternatieven. Bovendien heeft het tweede x-tuple een vraagteken en bestaat dus

misschien niet. Het totaal aantal mogelijke werelden dat door deze tabel wordt beschreven is acht. De volledige tabel is hieronder weergegeven.

DRIVES

person	color	car	
Frank	Red	Toyota	21
Frank	Blue	Toyota	22
Billy	Blue	Honda	? 23
Jimmy	Green	Mazda	24
Johnny	Green	Mazda	25

TriQL

Om de data te bevragen wordt in een relationele database SQL gebruikt. In Trio gebruiken we een op SQL gebaseerde taal Trio Query Language (TriQL). Net als in SQL, is ook in TriQL de basis een 'SELECT FROM WHERE'-statement. Daarnaast hebben we in TriQL een paar constructies waarmee we uncertainty en lineage kunnen manipuleren.

De 'SELECT FROM WHERE'-constructie in TriQL werkt hetzelfde als in SQL, met dit verschil dat de conditie in de where werkt op alternatieven. De query:

```
SELECT *
FROM DRIVES
```

levert dus de gehele tabel 'DRIVES' als resultaat. Aangezien de conditie werkt op alternatieven, kan het dus voorkomen dat slechts een beperkt aantal alternatieven van een x-tuple als resultaat wordt opgeleverd. Als voorbeeld geven we een query die alle personen oplevert die een rode Toyota hebben:

```
SELECT *
FROM DRIVES
WHERE color = 'Red' and car =
'Toyota'
```

Het resultaat van deze query is:

DRIVES

person	color	car	
Frank	Red	Toyota	?



De oplettende lezer zal direct geconstateerd hebben dat er ten opzichte van de tabel DRIVES een vraagteken is bijgekomen. Dit komt omdat het x-tuple niet geheel is geselecteerd, waardoor sommige werelden niet in de resultaten zijn opgenomen. Hierdoor bestaat het resultaat niet meer uit het volledige universum en is het antwoord dus onzeker geworden. Zodra niet alle alternatieven van een x-tuple in het resultaat zitten, wordt het x-tuple onzeker en krijgt een vraagteken.

Confidences

Tot nu toe hebben we alleen gekeken naar alternatieven. Alternatieven in een x-tuple sluiten elkaar uit en tussen x-tuples zijn alternatieven onafhankelijk van elkaar. In sommige gevallen is het niet alleen zo dat alternatieven bekend zijn, maar dat sommige alternatieven waarschijnlijker zijn dan andere. Door tijdens het maken van een tabel de optie 'WITH CONFIDENCES' toe te voegen, kunnen we aan alternatieven een kans toekennen. Als de totale kans

door de query die gebruikt is om tot de afgeleide relatie te komen.

Lineage

De afkomst van data is niet alleen handig, bijvoorbeeld om te kijken hoe data is ontstaan, maar het is zelfs noodzakelijk. Deze afkomst van data kan bijvoorbeeld afhankelijkheden tussen data laten zien, die nodig zijn bij het berekenen van kansen. Voor elk query-resultaat wordt per alternatief opgeslagen welke bronalternatieven bijgedragen hebben tot het resultaat. We gebruiken $\lambda(x_x, a_x) = (x_y, a_y)$ om aan te geven dat alternatief a_x uit x-tuple x_x als lineage-alternatief a_y uit x-tuple x_y heeft. Als voorbeeld geven we hier de lineage voor het alternatief (Frank, Amy, Red, Toyota) uit de SUSPECT-tabel.

$$\lambda(1, 31) = (1, 12)(1, 21)$$

Hiermee geven we aan dat alternatieven 12 en 21 in de lineage van alternatief 31 zitten. Na queries op een non-base table zal er dus lineage naar non-base tables zijn, waarvan dan weer lineage naar base tables gaat. Op deze manier ontstaat een lineage-graaf. Zodra er een cycle in de lineage-graaf ontstaat, is er mogelijk sprake van data-afhankelijkheid. Deze afhankelijkheid hoeft niet alleen van invloed te zijn op de bijbehorende kansen, maar ook op eventueel aanwezige vraagtekens.

"De afkomst van data is niet alleen handig (...), maar het is zelfs noodzakelijk."

Veel interessanter worden de queries als er gebruik wordt gemaakt van meerdere relaties. We kunnen met de crime solver ook verdachten identificeren met behulp van de query:

```
SELECT d.person as suspect,
       s.witness, s.car, s.color
FROM SAW s, DRIVES d
WHERE s.car = d.car AND s.color = d.color
```

SUSPECT

suspect	witness	color	car	
Frank	Amy	Red	Toyota	? 31
Frank	Diane	Red	Toyota	? 32
Frank	Diane	Blue	Toyota	? 33
Billy	Amy	Blue	Honda	? 34
Jimmy	Betty	Green	Mazda	? 35
Johnny	Betty	Green	Mazda	? 36

per x-tuple kleiner is dan 1, heeft het x-tuple een vraagteken.

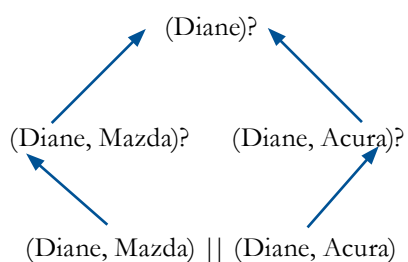
Zodra we een 'join' tussen twee relaties nemen, worden de kansen van de alternatieven vermenigvuldigd. Het resultaat is de kans voor het alternatief in de resulterende relatie. Aangezien kansen onafhankelijk zijn, kunnen we dit zonder problemen doen. Helaas werkt dit alleen voor *base tables*. Base tables zijn tabellen die niet het resultaat zijn van een query, of anders gezegd, tabellen waarvan de data door de gebruiker is ingevoerd. Zodra afgeleide relaties worden gebruikt, kan het namelijk gebeuren dat kansen afhankelijk geworden zijn

Bovenstaand resultaat, dat we de SUSPECT-tabel noemen, wordt verkregen door elke combinatie van alternatieven tussen SAW en DRIVES die aan de conditie voldoet, in het resultaat op te nemen. Als alternatieven afkomstig zijn van deze twee x-tuples in SAW en DRIVES, zijn die alternatieven ook onderdeel van hetzelfde x-tuple in het resultaat. Als niet alle alternatieven van beide x-tuples in SAW en DRIVES bijdragen aan het resulterende x-tuple, wordt een vraagteken toegevoegd. Als een x-tuple al een vraagteken had, blijft dit uiteraard ook in het resultaat bestaan.



De Keizer voor het Gates-gebouw

Onderstaande graaf laat zo'n afhankelijkheid zien.



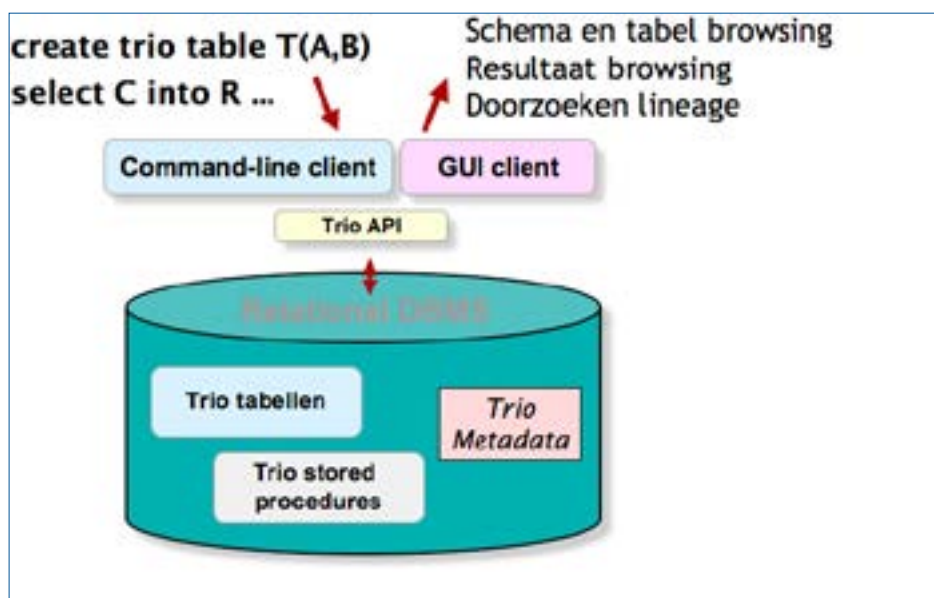
In deze graaf wordt in eerste instantie (onderste regel) uitgegaan van twee alternatieven zonder vraagteken. Door een query ontstaan twee verschillende x-tuples met toegevoegde vraagtekens. Door een volgende query worden de twee x-tuples weer samengevoegd, maar omdat beide bron-x-tuples hier een vraagteken hebben, heeft ook het resultaat een vraagteken. Het vraagteken hier is echter overbodig. Dit blijkt dan ook bij het doorlopen van de graaf.

Overige functionaliteit

Trio is veel uitgebreider dan tot nu toe beschreven. Zo is er een mogelijkheid om alternatieven als mini-relatie te bevragen. Dit wordt gedaan met *horizontale subqueries*, een soort subquery die werkt over alternatieven per x-tuple. Er zijn ook manieren om lineage te gebruiken in queries door het *lineage predicaat*. Met dit predicaat kunnen we bijvoorbeeld achterhalen welke data tot de relatie heeft geleid, zelfs als deze data niet meer beschikbaar is. Daarnaast zijn er op onderzoeksgebied nog heel wat interessante onderwerpen binnen Trio, zoals *mass confidence computation*, waarbij grote hoeveelheden kansen zo snel mogelijk, via lineage, worden berekend. Een compacte, maar relatief complete beschrijving is te lezen in [1].

Prototype

Tijdens het onderzoek wordt een prototype ontwikkeld om alle verschillende aspecten van de database met uncertainty en lineage te testen. Een schematische weergave van het prototype is gegeven in Figuur 1. Op dit moment wordt een bestaande



Figuur 1: Schematische weergave van Trio

RDBMS als basis gebruikt. Er is een online versie, inclusief uitgebreide taalhandleiding van Trio beschikbaar om te proberen. Er worden in deze online versie twee voorbeeld databases aangeboden, een uitgebreide versie van de crime solver en een IMDb/Netflix-database. Het is ook mogelijk om een database aan te maken en te vullen met eigen data. De eenvoudigste manier om het prototype te bereiken is door te zoeken naar 'Stanford Trio'. ■

Ander de Keijzer
a.dekeijzer@utwente.nl

Referenties

[1] Michi Mutsuzaki, Martin Theobald, Ander de Keijzer, Jennifer Widom, Parag Agrawal, Omar Benjelloun, Anish Das Sarma, Raghobham Murthy, Tomoe Sugihara, "Trio-One: Layering Uncertainty and Lineage on a Conventional DBMS (Demo)", *CIDR 2007:269-274*, 2007



ADVERTENTIE

Quinity

The Virtual Storyteller [TFHN03] is een systeem dat automatisch verhalen genereert en vertelt. De verhalen zijn iedere keer anders. Ze komen voort uit een door de computer gesimuleerde wereld waarin de karakters met elk hun eigen gevoelens en belangen gesimuleerd worden. Het verhaal wordt verteld door een Embodied Conversational Agent (ECA). Dit is een door de computer gegenereerd virtueel personage.

Er is al meer onderzoek gedaan naar interactieve verhalenvertelsystemen ([Rei96], [MS03], [RVC03]). Echter, de systemen die al bestaan maken gebruik van vormen van interactie die niet natuurlijk en onopvallend zijn voor de gebruiker. In dit project probeerden we een zo natuurlijk mogelijke vorm van interactie te vinden voor het Virtual Storyteller-project.

The Interactive Virtual Storyteller

Een verhaal over het vertellen van verhalen



Het systeem is in feite een multi-agent systeem waarin elke agent acties kan uitvoeren die een logisch gevolg zijn van het doel dat deze agent wil bereiken en de middelen die aanwezig zijn. Deze acties beïnvloeden de staat van de wereld waarin het verhaal zich afspeelt. Deze staat van de wereld en de veranderingen daarin worden in formele vorm beschreven en door een apart stukje programma omgezet tot natuurlijke taal. Daarnaast wordt er ook gewerkt aan visualisatie van het verhaal dat hieruit voortkomt.

Het genereren van het verhaal vindt op dit moment plaats zonder input van een gebruiker. Ons project richt zich op het onderzoeken van mogelijke vormen van (passieve) interactie en intelligentie voor dit systeem om zo tot het uiteindelijke doel te komen, een adaptief en perceptief systeem: The Interactive Virtual Storyteller.

Het meten van aandacht

In de eerste iteratie van het onderzoek stond het interactief maken van het systeem centraal. Het systeem zou op een voor de gebruiker natuurlijke manier interactief moeten zijn. Hier-

mee wordt bedoeld dat het systeem zonder actieve tussenkomst van de gebruiker moet detecteren wanneer de gebruiker een verandering in het verhaal wenst, en hierop reageert. Het systeem moet dus perceptief en adaptief zijn. Het moet de gebruiker observeren, en kunnen detecteren of en wanneer de gebruiker zijn interesse verliest. Op deze gebeurtenissen moet dan gereageerd worden. Het verhaal kan een plotselinge wending krijgen (bijvoorbeeld door ineens een element van spanning of horror te introduceren) of het verhaal kan sneller verteld worden. Als we deze gedachtegang volgen ontstaan er verschillende vragen, zoals: “Kan men invloed uitoefenen op het aandachtsniveau van de luisteraar door middel van korte en lange versies van het verhaal?”, waarbij direct weer volgende vragen worden opgeroepen, zoals: “Wat is eigenlijk aandacht en hoe meet je dat?” en “Vertoont de lichaamshouding of het gezicht van een persoon bepaalde kenmerken afhankelijk van zijn aandacht?”

De verwachting was dat men wel invloed kan uitoefenen op het aandachtsniveau van de luisteraar. Op de vragen op welke manier dit kan, en hoe men aandacht meet, probeerden we in dit onderzoek een antwoord te vinden. Dit eerste deel had dus een meer verkennende insteek om te ont-

dekken wat er nu eigenlijk nodig is voor het meten van aandacht.

Opzet

Om te onderzoeken of men invloed kan uitoefenen op het aandachtsniveau van de gebruiker is er een experiment uitgevoerd waarin de lengte van het verhaal wordt gevarieerd. Een bestaand verhaal wordt opgedeeld in stukken, en van ieder stuk wordt een lange en een korte variant gemaakt. Alle stukjes samen maken weer het complete verhaal, ongeacht of er per stukje gekozen wordt voor kort of lang. De bedoeling van het experiment is dat de gebruiker een vorm van natuurlijke interactie ervaart en er eigenlijk niks voor hoeft te doen. Het verhaal werd tijdens het experiment voorgelezen door de Microsoft Agent die gebruik maakt van de Lernaut & Hauspie TTS3000 speech engine [L&H03]. Simultaan met de gesproken tekst verschijnt de tekst woord voor woord in een tekstballonnetje in beeld. Om de proefpersonen ook iets anders te bieden dan enkel een tovenaars met een tekstballonnetje waren er ook fullscreen foto's te zien die de omgeving en de setting van het verhaal illustreerden. Deze foto's bleken niet overheersend te zijn met betrekking tot het aandachtsniveau maar precies genoeg om de luisteraars zich beter in te laten leven.



Het experiment werd uitgevoerd in een Wizard of Oz-opstelling, waarbij een lid van het onderzoeksteam als wizard fungeerde. Deze wizard bepaalde wanneer een lange en wanneer een korte variant gekozen werd. In feite zorgde deze persoon dus voor de 'intelligentie' in het systeem. Een dergelijke Wizard of Oz-opstelling is een manier om functionaliteit te testen die niet daadwerkelijk geïmplementeerd is.

Achteraf werd de proefpersonen gevraagd hun mening te geven over

schouwd als een signaal van desinteresse maar als een trekje of gewoonte van de persoon. Andere personen frunniken, friemelen of wrijven enkel op bepaalde momenten aan hun gezicht, haar et cetera. In zo'n geval wordt het wel als een teken van desinteresse opgevat. Met name het wrijven in de ogen duidt op vermoeiing of aandachtsverlies.

Sommige signalen zijn duidelijker dan andere en sommige signalen duiden sterker op desinteresse dan anderen. Eén van de meest duidelijke

kleine bewegingen of verzitten. Een voorbeeld daarvan is het bekijken van de mobiele telefoon. Dit duidt op een verlies van aandacht en daarom werd er op zo'n moment ook voor de korte versie gekozen.

De zaken waar we uiteindelijk op geannoteerd hebben zijn:

* Nulposities: welke basisposities neemt de persoon aan? Over het algemeen zijn dit houdingen die langer dan dertig seconden aangehouden worden.

* Wegkijken: dit kenmerk leek ons een van de meest opvallende kenmerken van desinteresse.

* Opvallende dingen: hieronder vallen alle andere zaken zoals: frunniken, friemelen, wrijven, telefoon bekijken, haar (los)schudden, een andere blik opzetten, gaan verzitten of bewegen van arm, hand, been of hoofd.

Deze kenmerken komen voort uit literatuuronderzoek (in belangrijke mate het paper van Frank Leenaars [LEE06]), uit een interview met Wim Wolbrink (een professionele verhalenverteller [WER07]) en uit een pre-evaluatie waarin we deze kenmerken hebben gecontroleerd.

Experiment 1

Voor het uitvoeren van het experiment hebben we geprobeerd een onbekend verhaal uit te kiezen om zoveel mogelijk beïnvloeding uit te sluiten. "De Stille Ronde van Bergen Op Zoom" [Coh21] was achteraf ge-

"Elke proefpersoon of gebruiker is anders en heeft natuurlijk zijn eigen trekjes en standaard bewegingen."

verschillende aspecten van het experiment; "Hoe leuk of spannend vonden ze het verhaal?", "Wat vonden ze van het systeem als geheel?" et cetera.

Tevens werden er tijdens het experiment opnames gemaakt die naderhand geannoteerd werden. Tijdens het annoteren werd er gelet op verschillende lichaamskenmerken die het aandachtsniveau zouden kunnen kenmerken, de zogenaamde *cues*.

Cues

Elke proefpersoon of gebruiker is anders en heeft natuurlijk zijn eigen trekjes en standaard bewegingen. Een groot deel van de proefpersonen frunnikt of friemelt continu aan zijn of haar gezicht, arm, hand et cetera. In zo'n geval wordt het niet be-

lijk signalen is het wegkijken van het scherm, bijvoorbeeld naar buiten of naar een ander deel van de kamer. Dit soort gedrag kwam weinig voor en als een gebruiker al wegkeek van het scherm was het in de richting van de evaluator. Het wegkijken kwam dus vooral voort uit oogbeweging en niet zozeer door het volledig draaien van het hoofd.

Andere duidelijke signalen zijn het sluiten van de ogen, zuchten en in de ogen wrijven. Er waren personen die duidelijk wat vermoeid waren en aan het eind van het verhaal de aandacht ietwat verloren. Op zo'n moment kiest de evaluator consequent voor de korte versies.

Een enkele keer kwam het voor dat de gebruiker andere dingen deed dan



Figuur 1: Het scherm voor de evaluator.



Figuur 2: Het scherm voor de proefpersoon.

zien bij alle acht proefpersonen onbekend. Dit verhaal hebben we zoals gezegd opgedeeld in meerdere stukken die in een XML-bestand gezet zijn. Door middel van zelfgeschreven software werd dit XML-bestand ingelezen waarmee eenvoudig een keuze te maken was voor een korte of lange versie van het verhaal.

Dit programma bestaat uit twee schermen. Eén scherm (figuur 1) is voor de evaluator en het andere scherm (figuur 2) is voor de luisteraar. Deze schermen kunnen op verschillende monitoren gezet worden zodat de luisteraar niet mee kan kijken met de evaluator en dus niet ziet wat de evaluator uitvoert.

Een foto van de opstelling zoals gebruikt tijdens het experiment is te zien in figuur 3. De gebruikte webcam is, zoals (niet) te zien is, verdekt opgesteld zodat de proefpersonen zich er niet van bewust waren dat ze opgenomen werden om zo beïnvloeding te voorkomen.

Resultaten

Om conclusies te kunnen trekken uit de resultaten van de evaluatie hebben we gekeken naar correlaties tussen de geantwoorde kenmerken en de antwoorden op de enquête van de respectievelijke proefpersonen. De correlaties tussen de frequentie van weggijken en het cijfer dat de perso-

nen hebben gegeven aan respectievelijk 'spannend', 'leuk', en 'interessant' zijn $-0,62$, $-0,81$ en $-0,72$. Dit houdt in dat hoe vaker men weggijkt, hoe minder spannend, leuk en interessant men het vindt. Daarnaast was er een correlatie zichtbaar tussen het cijfer dat de personen hebben gegeven aan 'spannend' en de gemiddelde tijdsduur dat ze stilzitten, namelijk $0,47$. Dit betekent dat hoe langer men stilzit, hoe spannender men het vindt. Dit verband is ook goed terug te zien in het scatterdiagram (figuur 4).

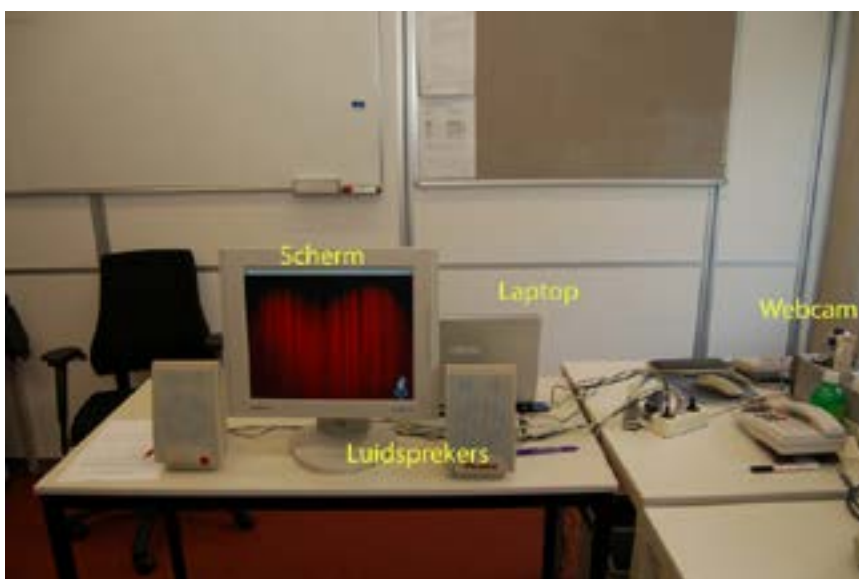
het aantal keren weggijken een zeer goede indicator van aandacht is. Het meten van de nulpositie kan ook een goede indicator zijn, maar dat is met dit aantal proefpersonen nog niet significant aangetoond. We denken dat, aan de hand van deze gegevens, het uiteindelijke systeem goed zal kunnen detecteren of de luisteraar nog geïnteresseerd is en of de verteller moet uitwijken qua detail of juist moet inperken om zo het tempo erin te houden.

Wel of niet expressief?

In de tweede iteratie van het onderzoek is onderzocht hoe de relatie is tussen het gebruik van lichaamstaal door een verteller en de aandacht van de luisteraar. Met lichaamstaal wordt voornamelijk gezichtsexpressie en handgebaren bedoeld. Dit leidde tot de onderzoeksvraag: "Kan een virtuele verteller, die veel gebruik maakt van lichaamstaal, de aandacht van de luisteraar beter bij het verhaal houden dan een verteller, die weinig gebruik maakt van lichaamstaal?" De verwachting was dat het gebruik van veel lichaamstaal meer gewenst is dan het gebruik van weinig lichaamstaal. Het gebruik van overdreven gezichtsuitdrukkingen

Uit deze resultaten kunnen we de conclusie trekken dat het meten van

zal ook meer gewaardeerd worden, omdat de expressies en emoties mak-



Figuur 3: De testopstelling van het eerste experiment





“Brian was playing the fiddle”



kelijker en sneller herkend worden. Dit uit zich in het beter vasthouden van de aandacht. Echter, te veel en te overdreven gebruik van lichaamstaal zal waarschijnlijk irritant zijn, net zoals totaal geen expressie dat is,

Omdat beide filmpjes in principe hetzelfde verhaal vertellen zouden we waarschijnlijk te maken krijgen met volgorde-effecten (als een proefpersoon twee keer hetzelfde verhaal te horen krijgt zal hij de tweede keer

kregen alle proefpersonen een vragenlijst om in te vullen. De eerste helft bestond uit het beoordelen van kenmerken (op een schaal van 1 tot 5, de Likert-schaal) van het filmpje en de verteller, het tweede deel bestond uit open vragen. Dezelfde vragenlijst kregen ze na het tweede filmpje. Na deze vragenlijst moesten de proefpersonen nog een vragenlijst invullen waarin gevraagd werd beide filmpjes te vergelijken.

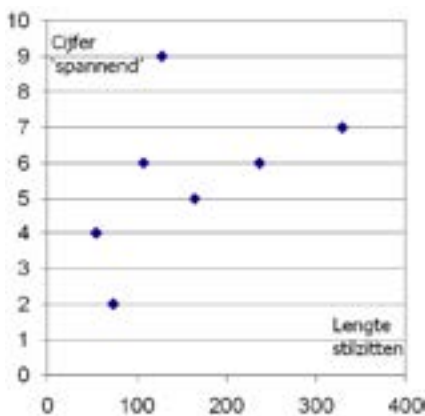
De getallen die de proefpersonen hebben toegekend na afloop van beide filmpjes zijn gepaarde waarnemingen. Wanneer goed gebruikt kunnen gepaarde waarnemingen voor redelijk sterke statistiek zorgen, omdat beide waarnemingen voor en na een bepaald moment direct met elkaar vergeleken mogen worden. Wij hebben gekozen voor de Wilcoxon Signed-rank Paired Test om steeds twee non-parametrische afhankelijke gepaarde variabelen met elkaar te vergelijken.

Wilcoxon-test

De Wilcoxon-test test in hoeverre de verdeling van twee gepaarde variabelen in twee gepaarde observaties hetzelfde is. Deze test neemt de grootte van de verschillen tussen de gepaarde variabelen mee. Het resultaat omvat een zogenaamde Ranks-tabel, die voor elk paar het aantal valide cases en ‘ties’ laat zien met daarbij de positieve en negatieve afwijkingen, de bijbehorende gemiddeldes en sommen van de rangen die toegekend zijn. Daarnaast wordt ook een tabel getoond waarin de Z-waarde en de waarschijnlijkheid p getoond wordt (zie voor een grafiek van deze tabel in figuur 4).

“Deze test neemt de grootte van de verschillen tussen de gepaarde variabelen mee.”

wat ervoor zorgt dat de aandacht zal verslappen.



Figuur 4: Scatterdiagram spannend/nulpositie.

Om te onderzoeken of lichaamstaal en expressie van de verteller een rol spelen bij de aandacht van de luisteraar, is er een experiment opgezet. In dit experiment kijken proefpersonen naar twee filmpjes, waarvan één filmpje veel lichaamstaal en expressie vertoont, en de ander weinig. Voor de filmpjes hebben we de hulp ingeroepen van Wim Wolbrink. Om te meten of lichaamstaal en expressie van de verteller een rol spelen bij de aandacht van de luisteraar, moesten de proefpersonen een aantal vragenlijsten invullen. De proefpersonen werden, net als tijdens het eerste experiment, onbewust opgenomen waarna deze filmpjes werden geannoteerd.

zijn aandacht er waarschijnlijk minder bij hebben). Daarom hebben we gebruikgemaakt van een ‘randomised crossover experiment’. Dit wil zeggen dat er twee even grote groepen zijn die willekeurig eerst ‘veel’ en dan ‘weinig’ te zien krijgen en andersom. Vanwege deze twee groepen zijn er veel meer proefpersonen nodig dan in het eerste experiment om de resultaten significant te laten zijn. Uiteindelijk hebben twintig mensen meegewerkt aan het tweede experiment.

Experiment 2

Voor het tweede experiment hebben we de filmpjes opgenomen met een goede videocamera en is wederom gekozen voor een zo onbekend mogelijk verhaal, namelijk het Ierse “The Story of the Man Who Had No Story” [Yol86]. De proefpersonen kregen allereerst een introductiefilmpje te zien om te wennen aan de omgeving en de stem van de verteller. Tevens werden hierin vijf (lastige) woorden uitgelegd. Na de eerste versie van het echte filmpje



“Brian! the old man said”



Een proefpersoon die niet oplet



Een proefpersoon die wel oplet

De Z-waarde wordt als volgt berekend: de absolute waarde van de verschillen tussen de waarnemingen wordt berekend (het teken van het verschil in het achterhoofd houdende), de gepaarde variabelen worden gesorteerd op grootte van het verschil en er wordt een rang toegekend waaraan het teken weer wordt toegevoegd. W is de som van deze resultaten. Bij een aantal (n) groter dan 10 mag de W genormaliseerd worden naar de normale verdeling (gecorrigeerd met een constante) middels:

$$Z = \frac{W - \mu W \pm 0,5}{\sigma W}$$

naar Z . Wanneer de Z -waarde boven de kritieke waarde voor een directionele test bij een α van 0,05 komt mag men spreken van een significant verschil. Voor $\alpha=0,05$ bij een directionele test is de kritieke Z -waarde 1,960 (de blauwe lijn in figuur 4). We gaan uit van een directionele test omdat we in onze hypothese stellen dat de variant 'veel' aantrekkelijker is en dus hoger zou moeten scoren bij de relevante vragen dan de variant 'weinig'.

Resultaten experiment

De uitkomsten van de Wilcoxon-test laten interessante resultaten zien. Zo is er bij veel kenmerken waarop de proefpersonen de verteller beoordeeld hebben een significant verschil te zien. Een resultaat is significant als er meer dan 95% zekerheid is over het feit dat het verschil tussen de filmpjes niet op toeval berust. Bij de termen *Leuk*, *Veel emotie*, *Menselijk overkomen*, *Tolerant*, *Vriendelijk*, *Zelfverzekerd* en *Sociaal* is dit het geval en wel in het voordeel van het filmpje *Veel*. In het geval van *Introvert* is het filmpje *Weinig* hoger beoordeeld. De

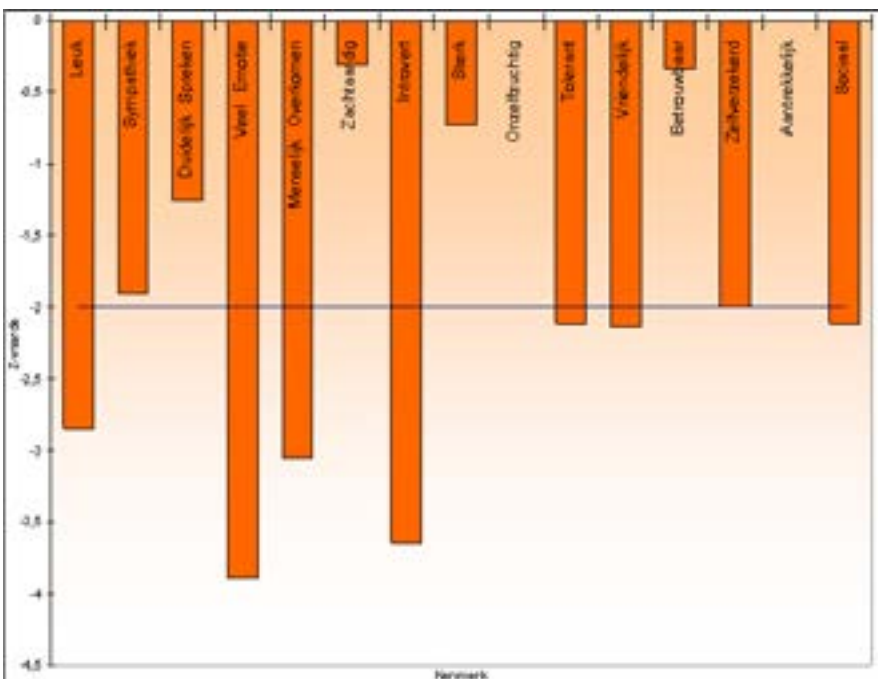
kenmerken *Duidelijk spreken*, *Zachtzaam*, *Sterk*, *Onzelfzuchtig*, *Betrouwbaar* en *Aantrekkelijk* laten allemaal geen significante resultaten zien wat overeenkomt met onze opzet. Deze kenmerken hebben we bijgevoegd om de proefpersonen niet te duidelijk te laten merken wat de bedoeling van onze enquête was om zo beïnvloeding te voorkomen.

Naast de kenmerken van de verteller hebben de proefpersonen ook het verhaal beoordeeld. De kenmerken *Spannend* en *Interessant* lieten een significant verschil zien in het voordeel van het filmpje met veel lichaamstaal. Hieruit blijkt dus dat een verhaal met veel lichaamstaal spannender en interessanter wordt ervaren dan een verhaal dat gebracht wordt met weinig lichaamstaal.

Het tweede deel van de enquête bestond uit vragen waarin beide filmpjes vergeleken moesten worden. Bij deze vragen zijn ook antwoorden gegeven die zeer duidelijke resultaten laten zien. Er is een χ^2 -test uitgevoerd om aan te tonen dat deze resultaten genoeg verschillen van de nulhypothese om zinvolle conclusies te kunnen trekken. De nulhypothese is dat er geen verschil tussen beide filmpjes is en dat dus beide filmpjes even vaak gekozen zouden moeten worden. Zoals de tabel laat zien kunnen we de nulhypothese voor elk kenmerk behalve *Natuurlijkst* verwwerpen.

Conclusie

Uit de evaluatie blijkt dat proefpersonen de voorkeur geven aan de variant met veel gezichtsexpressies en handgebaren: het verhaal wordt spannender en interessanter gevonden en de verteller wordt als leuker, menselijker, toleranter, vriendelijker, zelfverzekerder en sociaal ervaren, en men vindt dat de verteller meer emotie toont. Als men voor de keuze gesteld wordt tussen de twee varianten, vindt men de variant met veel gezichtsexpressies en handgebaren



Figuur 4: De resultaten van de Wilcoxon-test

Frequenties		Leukst	Spannendst	Natuurlijkst	Prettigst	Aandacht	Gaan Zitten	Emotie
waargenomen	veel	18	19	12	16	18	17	19
	weinig	2	1	8	4	2	3	1
	totaal	20	20	20	20	20	20	20
verwacht	veel	10	10	10	10	10	10	10
	weinig	10	10	10	10	10	10	10
	totaal	20	20	20	20	20	20	20
p volgens χ^2		0,000347	5,7E-05	0,371093	0,00729	0,000347	0,001745	5,7E-05

Tabel 1: Resultaten

leuker, spannender en prettiger. Men vindt dat deze variant de aandacht beter vasthoudt, meer emotie toont en men zou er eerder voor gaan zitten. Al deze uitkomsten maken dat

Referenties

[Coh21] http://beleven.org/verhaal/de_stille_ronde_van_bergen_op_zoom, laatst bekeken 25-10-2006, originele bron: "Nederlandsche sa-

Entertainment, S. Göbel, N. Braun, U. Spierling, J. Dechau and H. Diener (eds), *Fraunhofer IRB Verlag*, ISBN 3-8167-6276-X, 2003.

[Wer07] <http://www.werder.nl/>, "Verteltheater Werder", laatst bekeken op 10-01-2007.

[Yol86] "Favorite Folktales From Around The World" Jane Yollen. *Random House Inc. pp 20-23*, ISBN: 0394751884. 1986

"Met dit resultaat is onze hypothese dus bevestigd."

we kunnen stellen dat gebruik van veel lichaamstaal de aandacht van de luisteraar beter bij het verhaal houdt. Met dit resultaat is onze hypothese dus bevestigd. ■

Bram van de Laar

Aan dit project werkten mee:

Bram van de Laar, Jurrie Overgoor, Joan ter Weele en Niek Hoeijmakers.

Dankbetuigingen

Graag willen we Wim Wolbrink bedanken voor zijn tijd en kennis. Zonder hem zou dit project niet mogelijk zijn geweest. Ook willen we de docenten Mariët Theune en Dirk Heylen bedanken voor hun ondersteuning en nuttige kritiek. Uiteraard willen we ook alle proefpersonen bedanken voor hun tijd en moeite.

Meer informatie?

Ben je nieuwsgierig geworden en wil je meer informatie omtrent dit project? Bekijk dan het volledige verslag op: <http://wwwhome.cs.utwente.nl/~overgoor/HMIpj/verslag.pdf> of mail ons op hmipj@yorinf.net

gen en legenden" door Josef Cohen met 32 illustratiën in kleurendruk en zwart door Pol Dom. Eerste deel, Zutphen, W. J. Thieme & Cie, 1921.

[L&H03] <http://www.microsoft.com/msagent/downloads/user.asp#tts>, laatst geüpdated op 15 mei 2003, laatst bekeken op 06-11-2006

[Lee06] Leenaars, F., "Modeling Attention in Intelligent Tutoring Systems", *5th Twente Student Conference on IT*, Enschede 26th June, 2006.

[MS03] Mateas, M. and Stem, A., "Façade: An Experiment in Building a Fully-Realized Interactive Drama", *Game Developers Conference, Game Design track*, March 2003.

[Rei96] Reilly, W.S.M. "Believable social and emotional agents", *School of Computer Science, Pittsburgh, PA: Carnegie Mellon University, Ph. D. thesis*, 1996.

[RVC03] Ryokai, K., Vaucelle, C. and Cassell, J., "Virtual peers as partners in storytelling and literacy learning", in *Journal of Computer Assisted Learning*, vol19, issue 2, pp 195-208, 2003

[TFHN03] M. Theune, S. Faas, D.K.J. Heylen and A. Nijholt, "The virtual storyteller: Story creation by intelligent agents", in *E 03: Technologies for Interactive Digital Storytelling and*

ADVERTENTIE

Belasting - dienst/CICT

Voor het vak Art & Mediatechnology moesten studenten in groepen van twee onderzoek doen op het raakvlak van kunst en techniek en hiervoor een eigen kunstwerk produceren. Tussen “digital storytelling” en het wiskundig genereren en reproduceren van kunst was er meer dan genoeg ruimte voor artistieke vrijheid. Wim Bos en Michel Jansen bouwden met hun “Inaudible melodies” een heus

kleurenorgel: een combinatie van beeld en muziek.

Inaudible Melodies

Visuele kenmerken van muziek toegepast op kleurenorgels

Inleiding

Al sinds mensenheugenis spelen kunstenaars met combinaties van beeld en geluid. Door aan beeld geluid toe te voegen of andersom, wordt geprobeerd een nog intensere ervaring te creëren. Door de jaren heen zijn er verschillende manieren geweest om deze zogenaamde “visuele muziek” te maken. Door één van deze, het kleurenorgel, hebben wij ons laten inspireren. Nadat we een aantal voorbeelden van klassieke kleurenorgels hadden gezien vonden we dat dit ook anders kon in een digitaal tijdperk. Een duik in de geschiedenis van visuele muziek en de theorieën die daaraan ten grondslag liggen, om uiteindelijk een eigen digitaal kleurenorgel te kunnen bouwen, is wat volgde.

Een stukje geschiedenis

Het idee dat er een relatie moet bestaan tussen muziek en kleur is al minstens zo oud als de oude Grieken. Filosofen als Aristoteles en Pythagoras speculeerden al over een verband tussen de ‘*musical scale*’ en het kleurenspectrum van de regenboog. Onder hen die voor het eerst experimenteerden met deze beginselen waren grote kunstenaars uit de Renaissance als Leonardo da Vinci en Athanasius Kircher.

Louis Bertrand Castel was in de 18^e eeuw met zijn “Clavecin Oculaire” de eerste die muziek en geluid daadwerkelijk samenbracht in één apparaat. Dit kleurenorgel, dat het eerste van zijn soort was, was in essentie een klavier met vijf octaven dat, wanneer een toets werd ingedrukt, een gekleurd stukje papier of zijde tegen een zwarte achtergrond liet zien. Iedere toets was gekoppeld aan een specifieke kleur en per octaaf werden deze kleuren een tint lichter.

Een andere belangrijke bijdrage aan de geschiedenis van visuele muziek werd geleverd door de componist Alexander Scriabin, met zijn werk “Prométhée ou le poème du feu” (Prometeus of het gedicht van het vuur, opus 60). Een muziekstuk geschreven voor een compleet orkest, dat werd begeleid door een nog niet bestaand instrument met de naam “Clavier à Luce”. Dit instrument moest de kleuren spelen die begeleidend waren bij de muziek.

Oskar Fischinger was tot slot één van de pioniers in het samenbrengen van film en muziek. Zijn bekendste werk, “Motion Painting No. 1”, stamt uit 1947. Deze film, geschilderd op acrylglas op wijze van Bachs “Brandenburg Concerto no. 3”, wordt nog altijd gezien als één van de belangrijkste werken uit het genre. Fischinger maakte ook nog zijn eigen kleu-

renorgel, dat hij de ‘Luminograph’ noemde.

Tegenwoordig is de kunst van het combineren van beeld en geluid alomtegenwoordig. Van sfeerbevooroderende muziek voor een film tot lichtshows bij concerten. Zelfs de gemiddelde mediaspeler komt niet meer zonder een scala aan muziekvisualisaties. Dat betekent echter niet dat er niets meer te leren valt van de ‘oude kunsten’ van de kleurenorgels.

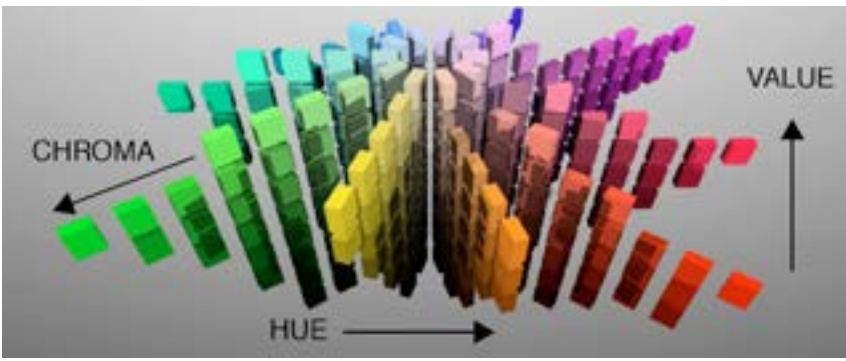
Een kleurenorgel bouwen

Een kleurenorgel is feitelijk niet veel meer dan een koppeling tussen visuele en muzikale elementen. Bij het maken van een kleurenorgel ligt het nogal voor de hand om muziek als leidraad te nemen en hier de visualisatie bij te genereren, aangezien de muziek is wat voorhanden is als er op wordt gespeeld. Per muzikaal aspect wordt dus bekeken welke visuele kenmerken eraan worden gekoppeld.

Tonen en kleuren

Tonen zijn de bouwstenen van muziek zoals kleuren dat voor visuele composities zijn. Zowel toon- als kleurverschillen zijn directe representaties van verschillen in de frequenties van respectievelijk geluid en licht. Bij mensen met de zeldzame aandoening synesthesie zijn tonen en kleuren zelfs gekoppeld in hun





Figuur 1: Combineren van tonen en kleuren

hersenen. Het bouwen van een kleurenorgel begint dan ook bij een succesvolle koppeling tussen deze twee

bij elkaar, zodat voor dissonantie het omgekeerde geldt.

“Anders gezegd: als de muziek harmonieus is, moeten de kleuren dat ook zijn.”

elementen. Omdat een orgel of piano een toetseninstrument is en dus uitsluitend zuivere tonen kan produceren, kunnen we aan iedere toon een kleur koppelen. Als we kleuren niet zien als een combinatie van rood, groen en blauw maar als een drietuple ‘hue’ (kleur), ‘chroma’ (verzadiging) en ‘value’ (lichtheid) zoals in figuur 1, is de “hue” wat een kleur definieert in termen van zijn frequentie, dus kiezen we deze parameter als basis voor onze koppeling.

Wanneer twee of meer tonen worden gecombineerd in een akkoord, ontstaat er een interactie tussen deze tonen die in de muziektheorie harmonie wordt genoemd. Een akkoord is consonant als het ‘stabiel’ klinkt en geen onrustig gevoel oproept. Een dissonant akkoord is het tegengestelde hiervan en doet wel onrustig aan.

Het is dus niet voldoende om kleuren aan tonen te koppelen. We willen ook dat de interactie tussen tonen tot uiting komt in de resulterende visuele compositie. Anders gezegd: als de muziek harmonieus is, moeten de kleuren dat ook zijn. Om dit te bereiken, legden we een bekende representatie van de belangrijkste harmonieuze toonverdeling, de “circle of fifths”, over een kleurenwiel, zoals in figuur 2. Als twee tonen samen consonant zijn, liggen de kleuren dicht

Deze verdeling herhaalt zich voor elke octaaf op het orgel. Om het onderscheid tussen de verschillende octaven waar dezelfde toon toe kan behoren te verduidelijken, koppelen we de hoogte van de toon ten opzichte van het gehele keyboard aan de “value”-parameter oftewel de lichtheid van de kleur. Een hoge toon is daarbij lichter dan een lage.

Tonen en positie

Op een piano heeft iedere toon zijn eigen toets en daarmee zijn eigen horizontale positie op het keyboard. De lage tonen beginnen links en naar rechts staat elke toets voor een hogere toon. Aangezien een beeldscherm twee assen heeft, ligt het erg voor de hand deze zelfde verdeling van links naar rechts op de X-as te representeren. Van links naar rechts, van laag naar hoog. De tweede as blijft daardoor vrij voor een andere kwantitatieve grootte.

Dynamiek en beweging

Tot slot maken de variaties tussen stukken die hard en zacht worden gespeeld muziek veelal interessanter. De toetsen van een keyboard kunnen zacht (piano) of hard (forte) worden aangeslagen, waarbij ze een harder danwel zachter geluid

produceren. Om deze dynamiek te visualiseren, maken we gebruik van twee visuele kenmerken: dichtheid en snelheid.

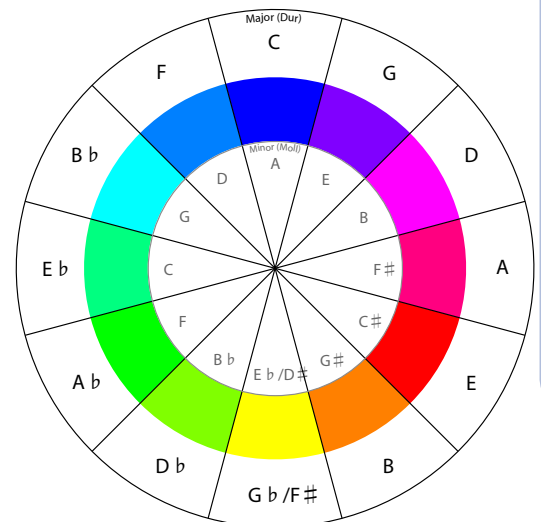
We hadden reeds bepaald dat elke toon een kleur en een positie heeft. In plaats van dit te representeren als stilstaande objecten, kiezen we ervoor een aangeslagen toon te visualiseren als een “fontein” die gekleurde deeltjes omhoog spuit vanaf een vaste basis op het scherm. Dit is te zien in figuur 3. Het aantal deeltjes dat omhoog wordt gespoten koppelen we aan de aanslag; harder betekent meer, zachter minder deeltjes.

Vervolgens koppelen we ook de snelheid waarmee de deeltjes hun oorsprong verlaten aan de aanslag, alsof ze met minder kracht omhoog worden gespoten. Dit heeft ook als gevolg dat zachte tonen minder hoog komen dan wanneer een toets hard wordt aangeslagen en tot bovenin het scherm reikt.

Technische implementatie

De hiervoor beschreven koppeling tussen muziek en beeld is geïmplementeerd in een programma met de naam “Inaudible Melodies Colour Organ”, dat beschikbaar is onder de GNU Public License. Figuur 4 geeft een schematische representatie van de componenten van dit systeem.

Het systeem doet in essentie niets anders dan het omzetten van muziek



Figuur 2: Kleurenwiel



Figuur 3: Screenshot

naar beeld; het is dan ook opgedeeld in twee delen, een muziekdeel en een visueel deel, die gekoppeld worden door een component die de omzetting verzorgt. Het linkerdeel handelt de muziek af, het rechterdeel het beeld. De informatiestroom loopt dus van links naar rechts, door de in groen afgebeelde ColorOrganController, waar de translatie plaatsvindt.

Als input gebruikt het programma een MIDI-keyboard, aangezien dit het toelaat rechtstreeks en in real-time uit te lezen welke toets hoe hard en hoe lang is ingedrukt, zonder dat daarbij complexe audio-analyse nodig is.

Voor het realiseren van de hiervoor genoemde 'fontein' van geluid is een zogenaamd 'particle system' gebruikt. Een groot aantal OpenGL *polygon quads* die altijd naar de camera wijzen en voorzien zijn van een gekleurde texture, vormen een levendige, dynamische stroom.

Toepassingen en toekomst

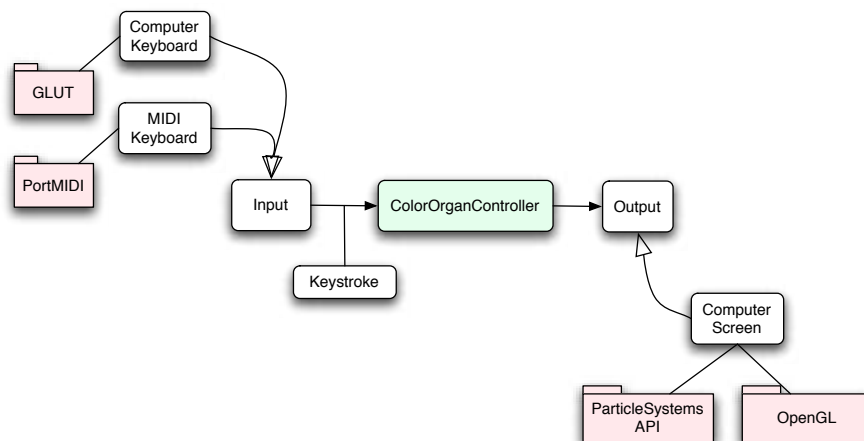
Zoals alle kleurenorgels die eraan voorafgingen, is ook het hier beschreven kleurenorgel bedoeld als kunst. Kunst die ernaar streeft om een prettige, volledige ervaring voor de toeschouwer te creëren. Dat wil niet direct zeggen dat deze kunstvorm niet nuttig kan zijn. Tijdens het testen van het kleurenorgel ontdekten we al dat de toevoeging van

beeld maakt dat een enkele fout gespeelde noot nog makkelijker wordt opgemerkt. In dezelfde lijn wordt visualisatie van muziek bijvoorbeeld al gebruikt bij muziekles aan slechthorende of zelfs dove kinderen. Het bouwen van dit kleurenorgel heeft ons in ieder geval een dieper begrip van de grondbeginselen van muziek en visuele theorie gegeven. ■

Michel Jansen

Nieuwsgierig?

Meer informatie, waaronder het oorspronkelijke artikel, de broncode van de applicatie en een aantal filmpjes, vind je op <http://micheljansen.org/inaudible-melodies/>



Figuur 4: Architectuur



Doing business over the Internet is becoming more and more popular. In the first half of 2005 4,26% of the Dutch retail turnover was accomplished by home sales, whereas this was only 1,71% in 1999 [THU]. This increase is due to an increase in shopping over the Internet. Besides that, more Internet shops are being started each day and more and more products are being added to the catalogues of Internet shops. As a result it is getting har-

der for customers (or users of the system) to find what they are looking for and to browse the catalogues to find interesting products.

Improving recommender systems

This problem is addressed by recommender systems [SKR01], which also have the effect of increasing sales and improving the shopping experience for users. A recommender system is a piece of software which uses an algorithm to mine collected data from a company to predict how much a user likes a product. If this value is above a certain threshold, then the product will be recommended. An example of data, that can be used are for example top sellers on the site, analysis of the buying behavior of users and demographics.

However, once a company uses a recommender system to make suggestions to users it clearly wants to make the most useful suggestions, thus increasing profits.

In practice a number of problems arise when making recommendations with a recommender system:

- How to recommend newly added items?
- How reliable are ratings from users?
- How to correctly recommend products that belong to a sequence?

In this article we will provide you with a description of these problems, give you an overview of existing solutions to these problems and present our own architecture for a recommender system. This architecture smartly combines existing solutions into one solution, thereby making it

easy for further development or researchers to build upon it.

Problems explained New item problem

Most recommender systems nowadays are built on algorithms like content-based and collaborative filtering [BaS97] which make recommendations based on two main criteria:

1. Recommend items similar to items the user liked in the past (content-based).
2. Recommend items similar users (e.g. from the same level of education) liked or bought (collaborative).

For making this type of recommendations, the system needs to know which users liked which items: for each item there needs to be at least one user rating. The system also needs to know how a new product relates to existing products. The problem arising from this system is evident: how to recommend new products?

Explicit rating problem

The collaborative filtering just described uses user ratings of items to predict how much another user likes this item. So the system checks for a given user what users are similar and takes the item ratings from these similar users to decide whether to recommend the corresponding item or not. The problem with this approach is that not all users who are similar

to a given user actually rate the items [Cla01]. Also, how reliable are the given ratings? For example: one might find the music on a CD outstanding, but if the cover contains shocking artwork, the user might give the CD a low rating. A similar user with other norms about artwork does not get a recommendation for the CD. Another problem is that most users feel interrupted in their browsing if they have to rate an item and therefore do not rate it. So, the ratings that users give themselves (therefore: explicit ratings), aren't always reliable upon which to base recommendations.

Sequence recognition problem

The filtering techniques used to recommend users a set of products are not able to see patterns in buying behaviour or patterns in rating behaviour [Pra05]. For example: suppose there are three books in an e-commerce system on the subject of computer science. These books, {A1,A2,A3}, are all bought by a certain group of users. So when a new user plans to buy book A2, the system will recommend book A1 and A3. This seems to be no problem, but the system does not know why, in which order or when the books have been bought. Maybe book A2 is a more up-to-date book than A1, logically users who bought book A1 therefore bought book A2 at a later moment. In the described case it

does not make sense for the new user who is planning to buy book A2, to recommend him or her A1. This problem is noted as the ‘sequence recognition problem’ in literature.

Current Solutions

Plenty of research has been conducted on the three problems mentioned above, all scattered in different journals and papers. Here we present you an overview of the current state of art.

New item problem

In the studied literature on the ‘new item problem’, all the articles surveyed use the same kind of solution, namely clustering [BaS97] [Bye04] [CYH04].

Clustering is grouping objects together in –as the name suggests– clusters. This is done by a clustering algorithm. For a clustering algorithm to work, we must have a set of objects which are in some way comparable. More formally said: all the objects must have at least one shared aspect (or attribute) allowing a comparison between the objects. For example: all the trees in the world are objects and their size is an aspect on which we can compare these objects. Note that for a clustering algorithm to work it is not needed to have the same types of objects; a shared comparable aspect is enough. Therefore it is the responsibility of the designer to define useful sets of objects to enable clustering.

The clustering algorithm itself works basically by constructing classes out of the given objects and their aspects [TSK06]. Considering the tree example: a clustering algorithm might come up with three classes: 0 – 20 metres, 20 – 40 metres, and 40 metres and up. After the classes are established, the objects are also divided over the classes by the clustering algorithm.

Clustering is by far the most frequently used method to recommend

new items. The most used algorithm is k-means [HaW79]. By running a clustering algorithm, newly added items are assigned to a cluster. When making recommendations to the user, the system picks objects from the most similar cluster compared to the product the user is currently browsing and therefore also recommends newly added items [BaS97], [Bye04] and [CYH04].

In the case of Internet business, the objects vary widely (home equipment, audio, video) and so do the aspects. So for example, how to handle a newly added CD? Suppose that the aspects are: genre and artist. These aspects cannot be retrieved from the object itself, so metadata is used. In practice this means that there must be a human effort made to enter enough metadata about the object for a useful clustering to be made. It is obvious that this process cannot be automated, because it is not easy to retrieve the genre from the raw audio signal. Besides, most metadata is also needed to allow users to search for products, so no extra effort has to be made.

Some interesting research has been done on retrieving data like tempi and even melody from an audio signal [CKW05]. This automatic generation of aspects is not covered in this article.

Explicit rating problem

Explicit ratings are ratings users give a product when asked for it and implicit ratings are ratings derived by the system from the behaviour of a user [GoS00]. As we saw in the problem description, these explicit ratings are often absent or not reliable. To solve this problem, we take a look at implicit ratings.

Implicit ratings are collected by the system and are based on the browsing behaviour of a user: implicit indicators. For example: if a user bookmarks a page about a product, he or she likes the product to a degree that it is worthwhile to bookmark it. A good classification of interest indicators is given by Clayspool [Cla01]:

- Explicit interest indicators: a user gives his or her rating at an item.
- Marking interest indicators: a user bookmarks an item, e-mails it, prints it, et cetera.
- Manipulation interest indicators: a user does something with the information, e.g. copying and pasting it, but also scrolling and mouse behaviour.
- Navigation interest indicators: a user spends time on a page about an item, follows a link (or not), et cetera.
- External interest indicators: a user starts to sweat, increases in heartbeats per minute, etcetera.
- Repetition interest indicators: a user who visits certain information repetitively is probably interested.
- Negative interest indicators: if no other interest indicators are de-

rived about a certain item, then it is likely to be of no interest.

The interest indicators must be gathered by the whole system and serves as input for the recommender system. It can be used in two ways: to rule out unreliable explicit ratings while they are made and to serve as input for the recommender system.

Implicit interest indicators are hard to obtain. The external interest indicators can only be measured in a lab situation with volunteers. For a real-time recommender system these indicators are of no interest (while eye movement patterns are one of the best ways of telling what a user likes [SaA01]).



Also, marking interest indicators and manipulation interest indicators can only be recorded by the users' browser. There are browsers that record this information and send it to a server, but they are mainly used for research and by volunteers [Cur]. Of course, such a browser will never become popular. Also privacy legislation makes these indicators hard to acquire [CBL01] (note that most spyware applications record these indicators).

The remaining implicit interest in-

parent nodes (for example: I want to go to the supermarket and I want to relax for a while). Also, when a plan is added to the network which can explain the observed action, another node is added representing how the fact fits into the plan (for example: to go to the supermarket; the walking part is the transportation part of the plan here). It is obvious that when there are many plans explaining a fact and these plans are big (i.e. there are many parts like the transportation part stated above), a fairly large

T-REC (our architecture)

In this section we will present a proposal for a system that solves the mentioned problems. This system is capable of handling new items, verify user ratings and is able to recognize plans in purchases. The system is named after the biggest and strongest among the dinosaurs, the Tyrannosaurus Rex or T-Rex. This big and strong recommender system is called T-REC.

The most abstract view of T-REC is drawn in figure 1. Although this figure is fairly simple, it is important because it nicely summarizes the situation and also gives a good overview on the systems domain. The model might even be interpreted as a general view of recommender systems.

From the model it can be concluded that there are only two actors who interact with the system, ordinary users and administrators. Administrators add product information into the system. Users can buy these products and rate them. Again, this is just a simplified view of the real world in such way that we can argue about T-REC without, for example, taking order processing in account.

From this broad view, we decompose the process from figure 1 in such a way that we obtain figure 2.

This model contains three data sources, one for product information, one for user ratings and one for purchases. This is strictly divided because each problem is about separate data.

The new item problem is solved by using clustering on product data. In this way, the function that allows the

"The sequence in which a user buys and rates items is defined by some kind of 'plan'."

dicators –navigation, repetition, and negative interest indicators– can be captured by the website. Once captured, they can be used to make recommendations [GoS00].

Sequence recognition

When the system does not know anything about the sequence in which products are bought, it can obviously not use this information when considering recommendations. The sequence in which a user buys and rates items is defined by some kind of 'plan'. Plan recognition is a widely studied theme in computer science, especially for recommender systems [Pra05].

To achieve automatic reasoning about plans, it is necessary to formalize a plan in some way, using predicates, functions, and terms. Representation of knowledge about a plan and actions are done in directed acyclic graphs with probabilities on the arcs. These graphs are referred to as a Bayesian network. Arcs point in the direction that reflects the causality. Basically, Charniak and Goldman [ChG93] give an algorithm to construct a Bayesian network from the observed facts. When a fact is observed (for example: I went walking) it is added to the network as a node. Then from some sort of knowledge base, all the possible explaining plans are added as

network arises. It is also stated above that probabilities are added to arcs; this is a complex process. But the main idea is that a plan with most of its parts explained by observed facts has a higher chance of being true. For interested readers; this process is also used in other Bayesian networks and is called Bayesian updating.

For using these plans in a recommender system, we take a look at the last items a user has bought and look in all the generated plans from other users where the same items are bought in the same sequence. From these plans we can derive which products are bought because of the need of buying them, after buying the first set of products [Pra05].

Combined solutions

In this article we will not cover existing ideas about how to combine some of the abovementioned problems, in the corresponding paper [Eer06] there is a section about these combined solutions.



Figure 1: General design

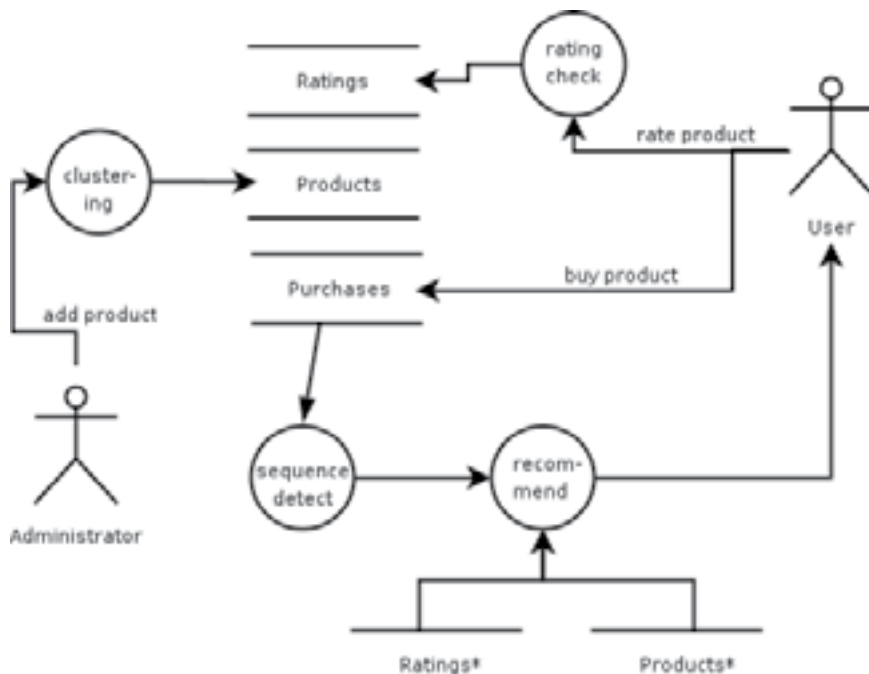


Figure 2: Decomposed design

system to also recommend new items, works on the product data source.

The problem with user ratings being possibly incorrect is a step before storing user ratings into the database. Because this extra step involves only data about user ratings and about the surfing behaviour, this can also be modelled separately.

In order to solve the problem of not being able to make useful recommendations because of the lack of historical data, we designed an algorithm that works on historical data. This algorithm makes recommendations only based on the historical data it has. Again, this part only needs one data source, namely historical data about purchases.

These are all design choices and can be implemented in another way in a real-life environment. The reason for choosing these separate data sources for each subpart of the system is that in this way, there is no need for some kind of normalization of the data. If we choose one big algorithm then all the data structures must be transformed to a uniform data structure or some encapsulation must be used. T-REC has different system components using different data sources (again: these might be on one database server).

In the corresponding paper [Eer06] we cover the implementation of each subsystem (the circles in figure 2). These implementations are based on the solutions given above and are beyond the scope of this paper. We will however present an outline of the recommend subsystem, because that is where the magic takes place: combining existing solutions into a workable system.

Recommend subsystem

This part of the system does the actu-

al recommendation. It uses the information about the product database, especially the item group vector. It combines this information with the ratings from users on the products. It queries the sequence part about specific recommendations based on historical data.

The latter case (products coming from the sequence part) is simple. T-REC prefers recommendations which can be underpinned by historical data above other recommendations. So, if the web shop using T-REC wants to display five recommendations and there are two recommendations from the sequence part,

other sources are used to make the three remaining recommendations.

The only thing left to specify is how to recommend the remaining products after the sequence part has been called. When a user browses a certain product, the item group vector from that product is retrieved. The most similar cluster is chosen based on the vector. If this cluster has more items than the number of recommendations to be made, then the top rated item is recommended.

The definition of 'top rated' is a bit fuzzy: is a good rated product by non-similar users better to recommend than a medium rated product by similar users? We must admit that this is a delicate question. However, from experimental results, a strategy called Cosine to combine user similarity and item similarity, works best [Sar01], [Lik03]. We will not elaborate on the Cosine algorithm; this would be too much detail for this article. The Cosine algorithm basically gives an ordering of all products, based on clusters and user similarity. The best product to recommend is on top and the worst on the bottom

of this list. So, we can simply take the top items from the list.

The only bad thing that can happen, is that in the top of this list the last few items have a bad score (but better than the rest). In most recommender systems on the market, some recommendations are chosen at random to give the user a broader choice of products. Therefore, in T-REC we will pick random recommendations in the mentioned case. When there are three items to be recommended with a score below some threshold, then exactly one of these recommendations is replaced by a random recommendation (of course being a product not already recommended).



When we describe the above in pseudocode, we get the fairly clean and simple code presented in figure 3.

```

N = number of recommendations
x = threshold value for good recommendation
rec = set of products to be recommended
sequence_part() = function which returns recommendations done by the sequence part of the system
cosine_recommends(n) = function which returns the top n products coming from the Cosine recommendation algorithm.
rec <- sequence_part();
if (length(rec) < N) {
  list = cosine_recommends(N-length(rec));
  foreach(item in list)
    if (item.score < x)
      randomize (rest_of_list);
  rec <- rec + list;
}
else if (length(rec) > N) {
  truncate (rec, N);
}

return rec;

```

Figure 3: algorithm blueprint

CONCLUSION

In this article we discussed the main problems with today's recommender systems and tried to combine solutions into one system which solves common problems for recommender systems.

The presented system outline showed that it is possible to make a streamlined system combining solutions to solve all the problems stated in this paper. All solutions are inspired by well known research and are considered correct. The presented system is the only system we are aware of which combines these solutions.

Another conclusion is that designing such a system with software design principles in mind, makes reasoning about T-REC much easier, if not trivial. The system is easy to comprehend because of the clear decomposition. In many papers we saw that the design step is skipped and an algorithm is presented to the reader. The clear system design we presented

might inspire others to start thinking at a more abstract level about recommender systems.

We have added our own ideas to T-REC, emphasizing the power of the combination. First, we adopted the idea of case based reasoning to give our solution on the item rating problem. Second, we came up with the connection between historical recommendations and normal ones. Last, we combined the idea of item vectors with other algorithms, like Cosine.

We presented a system outline which is based on modern ideas about software design and is therefore flexible and might be used in further research to test different algorithms on the same subject. ■

Eelco Eerenberg

REFERENCES

- [BaS97] Balabanović, M. and Shoham, Y. Fab: content-based, collaborative recommendation. *Communications of the ACM, Volume 40, Issue 3 (Mar 1997), Pages 66 - 72*
- [Bye04] Byeong Man Kim et all. A New Collaborative Recommender System Addressing Three Problems. *Lecture Notes in Computer Science, Volume 3157, (Jan 2004), Pages 495 - 504*
- [CBL01] Claypool, M., Brown, D., Le, P., Waseda, M. Inferring User interest *Internet Computing, Volume 5, Issue 6 (Nov/Dec 2001), Pages 32 - 39*
- [ChG93] Charniak, E. and Goldman, R. P. A Bayesian model of plan recognition. *Artificial Intelligence, Volume 64, Issue (November 1993), Pages 53 - 79*
- [CKW05]Cano, P., Koppenberger, M., Wack., N. An Industrial-Strength Content-based Music Recommendation System *Proceedings of the 28th Annual International ACM SIGIR Conference, 2005, Pages 673 - 673*
- [Cla01] Claypool, M., et all. Implicit interest indicators. *Proceedings on the 2001 international conference on intelligent user interfaces, 2001, Pages 33 - 40*
- [Cur] Curious Browser <http://perform.wpi.edu> visited 2006-05-06

[CYH04] Chen, J., Yin, J. and Huang, J. Recommendation of new items based on indexing techniques *Proceedings on 2004 International Conference on Machine Learning and Cybernetics, 2004 Volume 2, Pages 1168 - 1172*

[Eer06] Eerenberg, E. Improving recommender systems *Proceedings on 5th Twente Student Conference on IT, 2006 Volume 1, Pages 309 -316*

[GoS00] Goecks, J. and Shavlick, J. Learning users' interests by unobtrusively observing their normal behavior. *Proceedings on International Conference on Intelligent User Interfaces, Pages 129 - 132*

[HaW79] Hartigan, J.A., Wong, M.A. Algorithm AS 136: A K-Means Clustering Algorithm *Applied Statistics, Vol. 28, No. 1. (1979), pp. 100-10*

[LiK03] Li, Q., Kim, B. M. An approach for combining content-based and collaborative filters. *Proceedings of the Sixth international Workshop on information Retrieval with Asian Languages volume 11, 2003, Pages 17-24*

[Pra05] Prasad, B. HYREC: A Hybrid Recommendation System for E-Commerce. *Lecture Notes in Computer Science, Volume 3620, September 2005, Pages 408 - 420*

[SaA01] Salvucci, D. D., Anderson J.R. Automated eye-movement protocol analysis *Human-Computer Interaction, Volume 16, Issue 1, 2001, Pages 39 - 48*

[Sar01] Sarwar, B. M., et all. Item-based Collaborative Filtering Recommendation Algorithms *Proceedings 10th international WWW conference 2001, pages. 285-295*

[SKR01] Schafer, J.B., Konstan, J.A., and Riedl, J. E-Commerce Recommendation Applications. *Data Mining and Knowledge Disco, Issue 5, 2001, Pages 115 - 153*

[THU] Thuiswinkelen in Nederlands perspectief (*Dutch report about home shopping behaviour*) http://www.thuiswinkel.org/uploadedFiles/Thuiswinkelen.nl/Nieuwe_website/Thuiswinkelen_in_Nederlands_perspectief.pdf visited: 2006-02-27

[TSK06] Tan, P.N., Steinbach, M., Kumar, V. Introduction to Data Mining *Chapter 4, pages 145-199*

Als penningmeester houd ik me nu alweer enkele jaren bezig met de financiën van onze vereniging. Zo heb ik het eigen vermogen zien groeien van € 33421,89 op 1 januari 2002 tot dik € 45000,- op het moment dat ik dit schrijf. Een fors bedrag en ook nog steeds een stevige jaarlijkse groei. Diegenen die de moeite nemen jaarlijks de financiële jaarverslagen te bekijken of op de ALV mijn toelichting aan te horen, hebben al gemerkt hoe ik

daar over denk. Maar dat is maar een zeer klein deel van de leden. Het leek mij dan ook een aardig idee iets te schrijven over de rijkdom van onze vereniging.

Van de penningmeester

Het eigen vermogen van ENIAC of wat maakt een vereniging rijk

Maar wat is de rijkdom van een vereniging als de onze: is die af te leiden van het tegoed op een spaarrekening? Of eerder uit het aantal activiteiten dat er georganiseerd wordt? Of uit het aantal periodieken dat de leden ontvangt?

Een rijk verenigingsleven, hoe zou dat er voor ENIAC uitzien? Een hoge opkomst op de ALV? En wat is hoog? Vijf procent van de leden misschien? Dat zou neerkomen op zo'n 40 leden. En wat zou het bestuur kunnen doen om die leden aanwezig te laten zijn? Een boeiend bijprogramma? Vermaak of inhoud? Partnerprogramma of kinderopvang? Of is een ALV per definitie niet interessant? Dat lijkt me niet, we hebben toch de laatste jaren ook ALV's gehad met een redelijke opkomst.

Veel verschillende zaken organiseren, zodat er voor elk wat wils tussen zit? Of juist heel weinig doen, zeg twee evenementen per jaar, zodat de deelnemers "geconcentreerd" worden en de kans dat je een bekende uit je studietijd tegenkomt veel groter wordt? Of activiteiten koppelen aan (een) jaargroep(en) en die jaargroep(en) gericht benaderen?

Of moet het bestuur de leden vragen wat zij willen? Tegen de tijd dat jullie dit lezen is de ledenenquête al

weer enige tijd achter de rug. Ik ben benieuwd hoeveel van onze leden de moeite hebben genomen die enquête in te vullen en ons willen laten weten hoe wij onze vereniging beter kunnen laten aansluiten bij hun behoefte.

Of moeten we enthousiast doorgaan met wat we doen en het beter laten weten aan de leden? Met een grote ansichtkaart met een plattegrond van je eigen buurt en een speciale brochure waarin het georganiseerde evenement uitgebreid wordt toegelicht, en op die manier een opkomst van vijf procent van alle INF/BIT/TEL-alumni te bereiken (of tien procent van onze leden)? Want dat is de UT op die manier wel gelukt met de alumnidag.

Natuurlijk gaat het niet alleen om kwantiteit van de opkomst, sterker nog, naar mijn idee gaat het juist om kwaliteit van het samenzijn. Maar als je kwaliteit weet te bieden is het natuurlijk wel leuk als daar veel leden van meegenieten. Wij blijven ons best doen. Als jullie suggesties hebben om zaken te verbeteren, dan staat het bestuur altijd open voor positieve feedback. De financiële rijkdom van de vereniging biedt ruimte voor goede ideeën, waarbij wij als bestuur natuurlijk wel verantwoord met de financiën willen blijven om-

gaan. Maar daarbij hoort juist ook dat we het geld besteden aan zaken waar de leden wat aan hebben en niet alsmaar laten aangroeien op de spaarrekening. Zoals gezegd: jullie suggesties zijn van harte welkom en jullie inbreng als bestuurslid, commissielid, eenmalig vrijwilliger en/of deelnemer meer dan dat. Die is hoogst noodzakelijk voor een rijk verenigingsleven binnen ENIAC. Dus graag tot ziens! ■

Berend van den Brink
Penningmeester ENIAC

P.S.: wil je de enquête nog invullen:
<http://saampjes.dnsalias.org/eniac>
un: eniac
pw: eniacEnquete

Net zoals Inter-Actief niet zonder bestuur kan, kan ENIAC dat ook niet. Omdat, eveneens net als bij Inter-Actief, ook bij ENIAC het bestuur weleens wisselt zijn er ook steeds weer nieuwe bestuursleden nodig, dus:



Bestuur/actievalingen gezocht!

Met spoed gezocht:

Enthousiaste

Ervaren en/of Onervaren

Bestuurders (m/v)

ENIAC is hard op zoek naar enkele enthousiaste mensen die het bestuur willen komen versterken. Er zijn diverse open plaatsen en in onderling overleg met het interim-bestuur kan bepaald worden hoe de taakverdeling binnen het nieuwe bestuur er uit gaat zien. Daarbij streven we er naar ieders kwaliteiten optimaal te benutten: ook die van

jou!!!!

Daarnaast staat het ENIAC-bestuur altijd open voor:

ACTIEVELINGEN (m/v)

Mensen die zich niet willen vastleggen voor een jaar, maar bijvoorbeeld wel een inhoudelijk, sportief of gezelligheids-evenement willen (mee)organiseren, een artikel willen schrijven of plaats willen nemen in een scriptieprijscommissie. Het bestuur van ENIAC wil initiatieven van leden graag stimuleren en faciliteren en heeft daarvoor budget beschikbaar. Meld je aan en bespreek de mogelijkheden met het bestuur!

Neem contact op via mail of telefoon: bestuur@eniac.utwente.nl of 06-29074616



Naam: Heidi Vogel
Studiejaar: 1991
E-mail: heidi.vogel@avinity.net



Wie, Wat, Waar

Heidi Vogel

Waar liggen je roots en hoe kwam je in Enschede?

Ik kom uit een klein plaatsje in Groningen. Ik wilde een technische studie, wellicht omdat techniek met de papelepel thuis ingegoten is. Ik heb lang getwijfeld tussen Informatica en Toegepaste Wiskunde (TW), maar toen we tijdens de Thea-dagen alleen maar roofdier-prooi-modellen kregen bij TW was de keuze snel gemaakt. Ik heb ook nog in Eindhoven gekeken, maar Twente was beter bereikbaar vanuit Groningen en zag er zeker zo goed uit.

Wat dacht je toen je als eerstejaars in Enschede begonnen was?

Vrijheid! Op eigen benen staan. Je voelt je dan toch heel wat, wat achteraf wel weer mee valt natuurlijk. Toch onverwacht dat er zo weinig vrouwen waren: van de andere kant zat ik in een doegroep met vier meiden, dus dat viel wel weer mee. Wellicht was ik toch een beetje naïef. Ik vond het heerlijk om mijn eigen gang te gaan, al moest ik toch behoorlijk blokken om mijn vakken te halen.

Hoe zag in globale lijnen je studiepakket, afstudeer-richting of opdracht eruit?

Mijn keuze was op multimedia-toepassingen gevallen. Voor mijn afstuderen heb ik multimedia-toepassin-

gen voor het onderwijs onderzocht. Ik heb lang getwijfeld tussen Databases en Informatiesystemen, uiteindelijk is het IS geworden; Methodische Systeemontwikkeling bij Sjaak Brinkemper. Mijn opdracht kwam bij Toegepaste Onderwijskunde vandaan bij de groep Instrumentatietechnieken; de naam refereert aan instrumentatietechnieken die een docent in kan zetten. Ik heb onderzoek gedaan naar nieuwe media als nieuw instrument.

Ik heb veel geleerd, een multimedia database-prototype gebouwd in Informix. Het systeem heb ik getoetst bij de Hogeschool Enschede, het was helaas nog te theoretisch.

Als je terugkijkt naar je studiekeus wat vond je het leukst?

Mijn studiebeleving ging met rasseschreden omhoog toen ik zelf vakken mocht kiezen. Uitstapjes naar andere faculteiten vond ik erg leuk. Ik heb ook de BICA-aantekening (Bedrijfskundige InformatiCA) gehaald. Informatica zelf vond ik wat beperkt, ik vind het belangrijk dat je er ook wat mee moet kunnen. Daarom heb ik bijvoorbeeld ook mijn afstudeeropdracht buiten Informatica gezocht.

Verder heb ik het studentenleven altijd erg leuk gevonden. Ik had ook een geweldig gezellig studentenhuus.

Ik had ook wel graag in het buitenland af willen studeren, maar dat ging niet makkelijk. De ondersteuning vanuit de faculteit was ook wel erg beperkt. Toen er eindelijk een opdracht in Ethiopië was mocht ik niet van Olf Broenink omdat dat te vrouwonvriendelijk was. Maar hij heeft later wel de bekostiging voor een ticket weten te regelen, dat dan weer wel.

Wat was je eerste baan en hoe kwam je hier terecht?

Na mijn afstuderen heb ik met een aantal mensen een bedrijf gestart in virtuele leeromgevingen. Dat was niet zo'n succes. Laat ik daar maar niet te veel over zeggen. Het klikte niet. Snel daarna ben ik naar KPN Research gegaan.

Wat doe je nu, hoe ben je daar ingerold en wat vind je het interessantst in je huidige baan?

Ik ben mede-oprichter van Avinity, een start-up. We richten ons op internet-tv; videostreaming via TCP/IP met en zonder draad. Ik houd me bezig met het overbruggen van de technische wereld van de embedded systemen en de content-wereld. Avinity



zit op het Mediapark in Hilversum: dé plek. Nederland blijkt een goede kweekvijver: goede breedbandfaciliteiten, veel streaming informatie, het medialandschap is in beroering en zoekt nieuwe toepassingen. Omroep.nl is hier bijvoorbeeld opgericht. Je ziet een move naar het internet: Talpa begint met internet-tv, steeds meer afleveringen komen ook via het internet beschikbaar.

Hoe verhoudt de op de UT opgedane kennis tot de be-

ben je flexibeler en is alles toch makkelijker te regelen.

Wat zijn je (globale) plannen voor de aankomende paar jaar?

Het uitrollen van een nieuwe dienst: een combinatie van internet televisie en het bijbehorende technische product; het kastje. Hiervoor is het nodig om de content te faciliteren en te organiseren. Internet wordt zo een nieuwe bron van content die je op je tv kan bekijken. ■

“Toepasbaarheid moest erbij. Op de UT kon je toen althans veel doen zonder ooit te programmeren.”

nodigde kennis, wat moest erbij?

Toepasbaarheid moest erbij. Op de UT kon je toen althans veel doen zonder ooit te programmeren. Ik moest bijvoorbeeld bij KPN Research ineens in C++ programmeren. Nou dan moet je maar even heel snel een cursus doen. Je leert wel je analytische vaardigheden ontwikkelen, maar hands-on ervaring en toepasbaarheid zijn beperkt. Verder heb ik ook nog andere trainingen en cursussen gevolgd als “marketing in the networked economy” bij IMD in Lausanne en een leergang strategie aan de universiteit van Groningen.

Welke wens heb je nog open staan

De grootste en de beste. Lachend: “Ja, dat klinkt misschien wel arrogant, maar dat wil ik graag meemaken”. Als we hier over 10 jaar nog steeds met dezelfde bezetting zitten dan is er iets niet goed gegaan en trek je de stekker eruit. Wat Ronald (Brockmann, partner en informatica-alumnus) heeft meegemaakt met “No Wires Needed” zou ik graag met Avinity meemaken.

Verder wil ik ook heel graag onze kids noemen, echte doegroepkindjes. Daar zijn we erg blij mee. Dat is toch erg prettig aan een eigen zaak, dan

Naam: Ronald Brockmann
 Studiejaar: 1991
 E-mail: ronald.brockmann@avinity.net

Wie, Wat, Waar

Ronald Brockmann

Waar liggen je roots en hoe kwam je in Enschede?

Ik ben geboren in Laren, Noord-Holland en getogen in Eemnes. In Hilversum heb ik de middelbare school doorlopen. Wellicht dat de interesse voor een technische studie er al van huis uit inzit, aangezien mijn pa bij Lucent werkt. Twente stond goed aangeschreven en sprak aan bij de open dag. Ik ben ook nog in Delft wezen kijken, maar Twente sprak toch meer aan: alles bij elkaar op een campus.

Wat dacht je toen je als eerstejaars in Enschede begon?

Heel veel mensen met hetzelfde profiel, tsjonge dat viel wel een beetje tegen. Allemaal technische bètamannetjes, een behoorlijk homogene groep. Gelukkig kwam ik in een doegroep met vier meiden terecht. Erg gezellig.

Hoe zag in globale lijnen je studiepakket, afstudering of -opdracht eruit?

Ik ben bij TIOS afgestudeerd, zo herinner ik me nog vakken als PRIMP/PRONT (Protocol Implementatie en Protocol Ontwerp). Tsja, het afstuderen heb ik bij No Wires Needed gedaan (*Redactie: Ronald was een van de medeoprichters van No Wires Needed, een van de eerste bedrijven die producten ont-*

wikkelde voor WiFi. Later is dit bedrijf overgenomen door Intersil). Dat was een erg gave tijd. Ook erg leuk dat je je afstuderen bij je eigen bedrijf doet. Ik heb onderzoek gedaan naar quality of service (QoS) voor WiFi, een erg interessant onderwerp. Bij mijn huidige bedrijf (Avinity) is het overigens nog steeds een hot topic, na al die jaren.

Als je terugkijkt naar je studiekeus wat vond je het leukst?

Tsja, doordat we al vrij snel begonnen met No Wires Needed heb ik niet veel collegezalen van binnen gezien. Dat eigen bedrijf was het leukste natuurlijk. Overdag hard werken in Hal D waar we ruimte hadden gehuurd. En de avond natuurlijk vullen in de stad :-). Er gebeurden ook altijd veel leuke dingen op de campus. Altijd wel iemand te spreken die iets weet over netwerken of ATM.

Verder zijn we zijdelings betrokken geweest bij de oprichting van XS4ALL: dat was erg inspirerend. In het bijzonder de gedachte dat een eigen bedrijf een vehikel is om leuke dingen te doen komt uit die tijd.

Wat was je eerste baan en hoe kwam je hier terecht?

Zoals al gezegd was No Wires Needed mijn eerste job. We zijn met een

groep studenten begonnen met een klein bedrijfje op het gebied van draadloze netwerken, het huidige WiFi. 802.11 was net een standaard. Met wat geluk en wijsheid zijn we een tijdje marktleider geweest in dit segment: een bijzondere tijd.

Wat doe je nu, hoe ben je daar ingerold en wat vind je het interessantst in je huidige baan?

Onlangs heb ik samen met Heidi Vogel, Maarten Hoeben en Remi Blokker Avinity opgericht. Onze business is internet-tv en streaming video over WiFi en gewoon netwerk. We willen de embedded systemen wereld en de content wereld overbruggen. We zitten nu op het Mediapark in Hilversum, een ideale plek. Nederland blijkt een goede kweekvijver te zijn. Er zijn goede breedbandfaciliteiten, het medialandschap is in beroering. Omroep.nl is in dit gebouw opgericht. Je ziet ook dat alle omroepen bezig zijn met tv op internet.

Hoe verhoudt de op de UT opgedane kennis tot de be-

nodigde kennis, wat moest erbij?

Toepasbaarheid moest erbij. Abstracte kennis is ook erg nuttig hoor. Zo heb ik veel gehad aan *group theory* bij cryptografie en aan statistiek voor netwerkbelasting. Maar je moet toch ook echt kunnen programmeren. Op de UT kun/kon je veel doen zonder te programmeren. In de praktijk heb je vooral de combinatie én analytisch én hands on ervaring nodig, dat maakt het onderscheid. Die combinatie kreeg je niet mee van de UT.

Het merendeel van de toepasbaarheid heb ik mezelf geleerd door het te doen. Ik heb nog wel een cursus "ondernemer worden" gevolgd bij Twente School of Management. Vrij basaal en algemene wijsheden, maar wel nuttig. Een aantal zaken pas ik nog steeds toe, zoals "zitten en tekenen", niet eeuwen wachten met een deal maken.

Welke wens heb je nog open staan?

Tsja waar doe je het allemaal voor. Dat vroeg ik me bij No Wires Needed ook wel eens af. Zit je met z'n allen in zo'n hok, dag in dag uit... Achteraf bezien is dat wel de leukste tijd. Maar er hoort natuurlijk ook groei bij en liefst de grootste worden. Als je na tien jaar nog steeds met z'n tien in dat hok zit, dan stop je er mee, dan is het niet gelukt.

Wat zijn je (globale) plannen voor de aankomende paar jaar?

Avinity tot een groot succes maken, met z'n allen in dat hok zitten ;-) ■

-advertentie-

ENIAC sponsort!

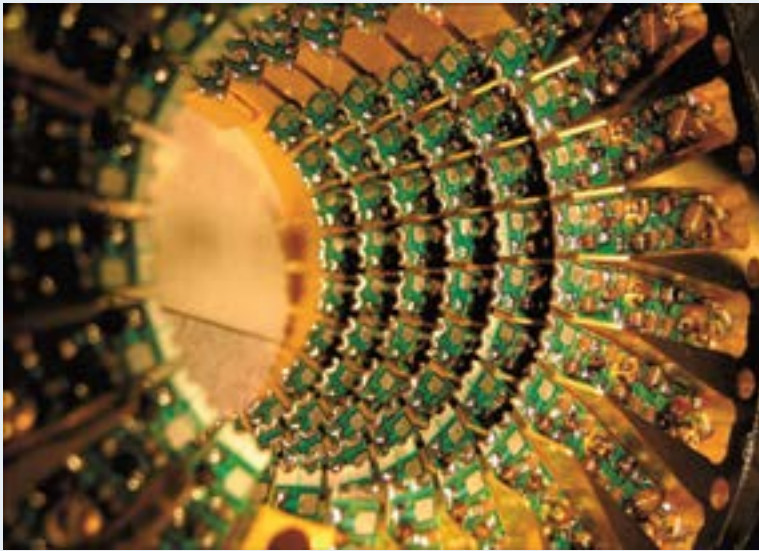
Heb jij een goed idee voor een bijeenkomst van alumni van INF/BIT/TEL? ENIAC wil dergelijke initiatieven graag ondersteunen. Als jij minimaal 10 alumni bij elkaar hebt, staat ENIAC achter je!

De voorwaarden zijn simpel:

- de activiteit komt op de ENIAC-agenda (met voldoende tijd voor aankondiging);
- de activiteit is vrij toegankelijk voor alle ENIAC-leden;
- jij zorgt voor een minimale deelname van 10 personen.

Het maakt ons niet uit of het om een inhoudelijke bijeenkomst gaat, een meer sportieve activiteit of gewoon een gezellig samenzijn. Naast jouw inspanning om het minimaal benodigde aantal personen bij elkaar te halen zal ENIAC de activiteit bekend maken op de agenda en in de elektronische nieuwsbrief. Voor de initiatiefnemer van het evenement is er bovendien een aardigheidje!

Neem contact op via mail of telefoon: bestuur@eniac.utwente.nl of 06-29074616.



Een van de 'filtering stages' van Orion, die de essentiële taak van het verwijderen van ruis uit het input-sigitaal voor hun rekening nemen

'Hoopvol scepticisme' over quantumcomputer

Erwin Boutsma

Het Canadese bedrijf D-Wavedemonstreerde vorige week 's werelds eerste quantumcomputer. Dat is veel eerder dan de wetenschap voor mogelijk hield.

Prof.dr.ir. Leo Kouwenhoven aarzelt als hem gevraagd wordt naar zijn mening over de demonstratie van de quantumcomputer van het Canadese bedrijf D-Wave. 'Het is lastig om er inhoudelijk iets over te zeggen. De demonstratie was overtuigend, maar je moet je bedenken dat het een privaat bedrijf is en dat ze geld nodig hebben van investeerders.'

D-Wave claimt als eerste een quantumcomputer te hebben gerealiseerd met zestien 'qubits'. De qubit vormt de basale rekeneenheid van de quantumcomputer en onderscheidt hem van een gewone pc, doordat een qubit 0 én 1 tegelijk kan zijn. Dit opent de weg naar berekeningen die buiten het bereik van gewone pc's liggen, zoals het kraken van encrypties en het ontwikkelen van specifieke medicijnen.

D-Wave demonstreerde de werking van hun quantumcomputer, 'Orion' gedoopt, de afgelopen week. Daarin kreeg het apparaat een Sudoku-puzzel voor zijn kiezen, die het probleemloos wist op te lossen. Ook met een klassiek vraagstuk over een tafelschikking (waarin persoon A niet naast persoon B wil zitten, persoon B alleen naast een vegetariër, etc.), had Orion geen enkele moeite. Beide vraagstukken zijn eenvoudige voorbeelden van een zogenaamde 'NP-compleet probleem', waarin quantumcomputers excelleren.

Maar Kouwenhoven, zelf als onderzoeker op het gebied van quantumcomputers verbonden aan het Kavli Institute of Nanoscience, waarschuwt: 'Pas bij zo'n duizend qubits ga je een snelheidsvoordeel krijgen ten opzichte van een gewone pc. Ik moet nog zien of dat ze lukt.' Hij doelt op de techniek die D-Wave gekozen heeft, die is gebaseerd op koeling tot dichtbij het absolute nulpunt (-273 graden Celsius) om de benodigde quantummechanische randvoorwaarden te creëren. Bij deze temperatuur veranderen de elektronen in de supergeleidende magneetjes van aluminium en niobium in zogenaamde 'Cooper pairs'; deze bezitten de juiste quantummechanische eigenschappen om als qubits te fungeren. De techniek staat onder wetenschappers echter bekend als lastig op te schalen.

Kouwenhoven is derhalve 'hoopvol maar sceptisch' ten aanzien van de mogelijkheid om de zestien qubits op te schalen naar 32 (eind 2007), 512 (zomer 2008) en 1024 (eind 2008), zoals D-Wave van plan is. 'Maar misschien hebben ze wel een techniek ontwikkeld waardoor het mogelijk is. Dat zou fantastisch zijn. Het probleem is dat niemand het op dit moment kan controleren. Er is nog geen wetenschappelijke publicatie over en het is hun medewerkers zelfs verboden erover te spreken.'

D-Wave wil al op korte termijn commercieel aan de weg timmeren door reken-tijd op Orion te verhuren.

Overgenomen uit Technisch Weekblad, Jaargang 38, #8, 24 februari 2007

Royal Navy stapt over op Windows

Het onder Windows-gebruikers beruchte 'Blue Screen Of Death' (BSOD), het blauwe scherm met de onbegrijpelijke foutmelding, krijgt binnenkort een geheel nieuwe dimensie. Het Britse bedrijf Insyte, dat commandosoftware ontwikkelt voor de schepen van de Royal Navy, stapt namelijk over op Microsoft Windows 2000 als basis voor haar software. Dit meldt de Britse website The Register.

Nieuwe Type 45 destroyers worden momenteel al afgeleverd met 'Windows for Warships' en verschillende oudere oppervlakteschepen en onderzeeboten worden er in de komende maanden ook mee uitgerust. Critici vinden het onverantwoord Britse oorlogsschepen te laten besturen door Amerikaanse software die dagelijks onder vuur ligt van virussen, malware en aanvallen van kwaadwillende hackers, en bovendien een slechte naam heeft op het gebied van stabiliteit. De kapitein van een Vanguard-klasse onderzeeër, met aan boord tientallen nucleaire raketten (icbm's), zit niet bepaald te wachten op een BSOD.

Toch beschouwt de Royal Navy de introductie van Windows als een grote vooruitgang. De oude systemen, nog altijd in gebruik, vereisen namelijk de permanente aandacht van vier technici om ze aan de praat te houden; de commandosoftware draait nog op 24-bits computers met een 1 Mhz processor en 25 kB geheugen. (cb)

Overgenomen uit Technisch Weekblad, Jaargang 38, #9, 3 maart 2007

Oproep: lever nieuwsberichten aan

Deze pagina is (zoals waarschijnlijk wel duidelijk is) een nieuwe rubriek in de I/O Vivat; de redactie is van zins elke editie één of meerdere pagina's te vullen met nieuws dat interessant is voor de lezers van dit magazine.

We roepen bij dezen alle lezers op om interessante ICT-artikelen te sturen naar iovivat@interactief.utwente.nl

Omdat veel ICT-nieuws, zeker het hardware-nieuws, vaak een nogal beperkte houdbaarheidsstermijn heeft, dat wil zeggen het nieuws is snel verouderd, zoeken we berichten die niet na een week volledig achterhaald zijn maar juist hun noviteitswaarde lang behouden; de artikelen moeten gaan over voor ICT'ers (denk INF, BIT, TEL) interessant of grappig nieuws zijn dat we mogen plaatsen onder vermelding van de bron; dit betekent mede dat een ingestuurde tip of ingestuurd bericht altijd een bron moet hebben!

ADVERTENTIE

SAI

ADVERTENTIE

TNO

Wie afgestudeerd is of bijna gaat afstuderen heeft een belangrijke keuze te maken: wil hij het bedrijfsleven in, of doorgaan in de academische wereld in de vorm van promoveren en eventueel een positie aan een universiteit? Voor sommigen is deze keuze wellicht makkelijk te maken, anderen zullen het er moeilijker mee hebben. Met name degene die zowel wetenschappelijk onderzoek als het bedrijfsleven interessant vindt, zal eens een kijkje

moeten nemen bij de Nederlandse Organisatie voor Toegepast-Natuurwetenschappelijk Onderzoek, beter bekend als TNO.

Op bezoek bij...

De organisatie

Toen ons land aan het eind van de Eerste Wereldoorlog te kampen had met een groot aantal problemen, werd er een commissie opgericht met als doel de beoefenaren van de natuurwetenschappen te betrekken bij het oplossen van deze problemen. Nadat deze commissie haar werk had voltooid, realiseerde men zich dat het zinvol zou zijn om door te gaan met het onderzoek naar hoe wetenschappelijke kennis het best toegepast kan worden in het bedrijfsleven en bij de overheid. Uiteindelijk resulteerde dit in 1932 in de 'TNO-wet', die onder andere de taken en organisatievorm van het toen opgerichte TNO vastlegde. Overigens is TNO, ondanks wat deze constructie wellicht doet verwachten, geen overheidsinstelling.

Ondertussen is TNO uitgegroeid tot een grote organisatie met zo'n 5000 medewerkers, maar de missie luidt nog steeds: wetenschappelijke kennis toepasbaar maken om het innovatief vermogen van bedrijfsleven en overheid te versterken. Om dit te bereiken neemt TNO voornamelijk deel aan onderzoeksprojecten met marktpartijen. Er is echter ook geld en tijd beschikbaar om zelf onderzoek te doen, met uiteraard als voorwaarde dat dit uiteindelijk op een innovatieve manier toegepast moet kunnen worden. Een behoorlijk aan-

tal deeltijdhoogleraren helpt hieraan mee om het wetenschappelijke niveau hoog te houden.

Het werk bij TNO is opgedeeld in vijf kerngebieden: Kwaliteit van Leven, Defensie en Veiligheid, Industrie en Techniek, Bouw en Ondergrond, en Informatie- en Communicatietechnologie.

TNO Informatie- en Communicatietechnologie

Het kerngebied dat voor (oud-)studenten Informatica, Telematica en BIT het meest interessant is, is 'Informatie- en Communicatietechnologie'. Dit onderdeel van TNO is gevestigd in Delft (de hoofdvestiging), Groningen en Enschede, met in totaal zo'n 385 medewerkers. Binnen het kerngebied zijn de ICT- en Telecomdisciplines van TNO gebundeld en worden bedrijven, overheden en (semi-)publieke organisaties geholpen met het succesvol innoveren met ICT. Er wordt voornamelijk gekeken hoe ICT en telecomtechnologie op een strategische manier nuttig gemaakt kunnen worden. De inhoud van uitgevoerde projecten is erg divers. Zo wordt er gewerkt aan slimme toepassingen van technologieën als sensoren, GPS en RFID, maar bijvoorbeeld ook aan nieuwe ontwikkelingen op het gebied van (mobiele) breedbandoplossingen.

TNO ICT



Dennis Krukkert

Wij spraken met een medewerker van TNO Informatie- en Communicatietechnologie: Dennis Krukkert. Dennis studeerde Informatica aan de Universiteit Twente, waar hij afstudeerde op 'automated business process matching within ebXML'. Dennis vertelt dat de vestiging in Enschede waar hij werkt zich voornamelijk bezighoudt met e-business. Hij werkt aan manieren waarop informatie het best elektronisch kan worden uitgewisseld. Zo deed TNO mee aan een onderzoek voor de veiligheidsregio Utrecht, waarbij een berichtenstandaard is gemaakt die in geval van een ramp gebruikt wordt om gegevens over slachtoffers uit te

wisselen tussen verscheidene betrokken instanties.

Een ander interessant project waar TNO haar kennis op heeft losgelaten betreft de ontwikkeling van een "slimme" dijk. Het stijgen van de zeespiegel leidt ertoe dat dijken steeds belangrijker worden in de bescherming tegen het water. Het simpelweg ophogen van de dijken leidt niet alleen tot horizonvervuiling, maar is ook nog eens erg kostbaar. Daar komt bij dat de hoogte van de dijken bij recente doorbraken niet zozeer het probleem was. TNO heeft daarom onderzoek gedaan naar het plaatsen van allerlei sensoren (druk, vochtigheid, geluid) in de dijken, met als doel het kunnen voorspellen wanneer een doorbraak op komst is. In dat geval kunnen op tijd verstevigingen aangebracht worden, om problemen te voorkomen. Om de projecten tot een goed einde te brengen wordt er met veel partijen samengewerkt. Er zijn contacten met alle universiteiten, met name met de drie technische. Bovendien heeft TNO contacten met vele hogescholen.

De medewerkers

Enthousiast geworden over TNO? Als technisch (oud-)student pas je prima binnen de organisatie. Haast alle medewerkers hebben een universitaire achtergrond, waarvan een behoorlijk aantal zelfs gepromoveerd is. Ben je nog niet zover dat je kunt beginnen met je (eerste) baan, dan biedt TNO volop mogelijkheden voor een stage- of afstudeeropdracht. Wie al even aan het werk is, maar er over denkt om alsnog te gaan promoveren, kan ook bij TNO terecht om dit te combineren.

TNO werkt altijd op basis van projecten, en zoekt dus met name mensen die goed kunnen samenwerken. Daarnaast moet je affiniteit voor innovatie hebben en zelfstandig met

ideeën kunnen komen. Wie op bepaalde vlakken nog wat bijschaving kan gebruiken, heeft bovendien de mogelijkheid om (interne) opleidingen te doen, zowel gericht op algemene vaardigheden als op technisch-inhoudelijke capaciteiten. Tijdens de eerste paar jaar hebben nieuwe medewerkers de kans om te ontdekken in welke richting hij of zij zich wil ontwikkelen. Zo is er de mogelijkheid om richting consultancy te gaan (advisering van organisaties), om innovator te worden (meer onderzoek) of om je met (project)management bezig te houden. Wie duidelijk aan-



ladder te blijven hangen, maar kan zich net zo goed op het inhoudelijke vlak opwerken.

Waarom specifiek bij TNO werken? Dennis stipt aan dat projecten veel inhoudelijke diepgang hebben, maar tevens erg divers van aard zijn. Dit wordt versterkt door de combinatie van onderzoeks- en marktprojecten wat werken bij TNO erg afwisselend en uitdagend maakt. Zoals de missie van TNO al aangeeft is alles bovendien gericht op innovatie. Er is in de werkweek dan ook altijd ruimte voor het opdoen van nieuwe kennis. Mocht je interesse hebben in een carrière bij TNO, neem dan contact op met Dennis via dennis.krukkert@tno.nl, of via *Inter-Actief*. ■

Mark Timmer

"Te strakke regels werken toch niet bij academici"

geeft wat hij graag wil, kan zich over het algemeen ook in die richting ontwikkelen.

De bedrijfscultuur

Qua werkomstandigheden is TNO behoorlijk soepel. Je kan je tijd zelf indelen, zolang je maar aan je uren komt. "Te strakke regels werken toch niet bij academici", aldus Dennis. Op kantoor zal de gemiddelde persoon die je tegenkomt ongeveer 30 jaar oud zijn, en een pak wordt meestal alleen gedragen bij een bezoek aan de klant. Wat Dennis ook goed bevalt, is dat de innovators (researchers) even hoog gewaardeerd worden als managers. Wie niet richting het management wil hoeft dus niet onderaan de

