

Pagina 4 • DAREnet en NARCIS gaan samen • De afdeling Onderzoekinformatie van de KNAW werkt dit jaar aan de integratie van de twee portals op het gebied van wetenschappelijk onderzoek.

Pagina 4 • Gruuthuse digitaal • De Koninklijke Bibliotheek legde de hand op het Gruuthuse handschrift. Wat ging daaraan vooraf en hoe komt een en ander beschikbaar?



Pagina 5 • Muzikale databankier Louis Grijp • Musicoloog en bijzonder hoogleraar Louis Grijp, verbonden aan het Amsterdamse Meertens Instituut, raakte in de jaren zeventig in de ban van de oude muziek. Hij ging die zelf spelen, en legde van het Nederlandse lied een databank aan die uniek is in de wereld.

Pagina 6 • Het Nederlands Interdisciplinair Demografisch Instituut • Directeur Frans Willekes werpt in de rubriek Focus licht op de activiteiten van het NIDI.

EN VERDER

Agenda.....	2
Nieuws.....	3
Achtergrond.....	4
Ontwikkelingen.....	7
Column.....	8
Gelezen.....	8

OCW betaalt studie data-opslag

Het ministerie van Onderwijs, Cultuur en Wetenschap (OCW) heeft middelen beschikbaar gesteld voor een verkennend onderzoek naar de opslag en beschikbaarheid van data in Nederland. Dat werd op 21 mei bekend op een vergadering in de Koninklijke Bibliotheek in Den Haag, bij de oprichting van de Nationale Coalitie voor Digitale Duurzaamheid (NCDD).

Met de OCW-steun is een goede start verzekerd van de coalitie, waarover al een klein jaar wordt gepraat. Voorzitter van de NCDD is drs. René Smit, die ook bestuursvoorzitter is van de Vrije Universiteit en Hogeschool Windesheim. Deelnemers zijn de nationale onderzoekorganisaties NWO en KNAW, SURF, het CBS, Beeld en Geluid, het Nationaal Archief, DANS en de Koninklijke Bibliotheek en als toevoegde technische universiteit Delft. De coalitie wil er voor zorgen dat Nederland over vijf jaar de infrastructuur heeft om de duurzame toegankelijkheid te garanderen van digitale gegevens van wetenschap, bedrijfsleven en cultuur. Daarnaast is het doel om informatie uit te wisselen en te verspreiden, internationale samenwerking te bevorderen en het profiel van Nederland als expertiseland accent te geven op het gebied van digitale duurzaamheid en data infrastructuur. Ten opzichte van de vroegste plannen, toen de coalitie nog vooral de onderzoekwereld als werkterrein zag, is de horizon nu



Jansen: 'In gesprek met VNO-NCW'

verbreed. Hans Jansen, directeur Research & Development van de KB en mede initiatiefnemer van de coalitie, licht toe dat in het bedrijfsleven 'ontzettend veel data' worden geproduceerd. 'Wij zijn dan ook in

gesprek met VNO-NCW over deelneming in de coalitie. Dat ziet er positief uit maar in dit stadium kunnen we er nog niets definitiefs over zeggen. Daar moet natuurlijk ook met de besturen en de achterbannen over worden gesproken.

Het ministerie van Onderwijs, Cultuur en Wetenschap stelt vier ton beschikbaar voor de verkenning, die na de zomer moet beginnen en 'ongeveer een jaar' zal duren. Daarmee gaat een belangrijke wens in vervulling van Jansen, die vorig jaar in *e-data@research* nog een studie bepleitte naar alleen data welke onder auspiciën van de overheid zijn verzameld. Net als de coalitie zelf krijgt de verkenning nu echter een veel breder gezichtsveld. 'Wie bewaart wat bij overheid, bedrijfsleven, cultuur en wetenschap, dat is de opdracht. En wie neemt welke verantwoordelijkheid, want een belangrijke

vraag is niet alleen hoe data worden opgeslagen maar vooral hoe ze voor de toekomst beschikbaar worden gehouden en wie daarvoor zorgt. Het is goed dat het bedrijfsleven erbij wordt betrokken, want daar worden niet alleen enorm veel gegevens, bijvoorbeeld op financieel gebied, verzameld maar je kunt je afvragen of dat voor de lange termijn wel goed gebeurt'. Met de verkenning in de hand zal de coalitie een strategie ontwikkelen voor de inrichting van een nationale infrastructuur. De nodige activiteiten worden daarin benoemd met de partijen die daarvoor verantwoordelijk zijn.

In de zomer wordt er een office manager aangesteld om als spil te fungeren in het beoogde slanke apparaat van de NCDD. Een website is er al in beta-versie maar de url was bij het sluiten van deze *e-data* nog niet bekend. (MdG)

Hugo de Groot bereikbaar

Op tien april jl. werd aan minister Donner het eerste exemplaar uitgereikt van *Hugo de Groot, Een leven in strijd om de vrede, 1583-1645*. De nieuwe biografie van De Groot is geschreven door Henk Nellen, medewerker van het Huygens Instituut. Het is, zoals de minister in zijn dankwoord memoreerde, de eerste grote biografie van De Groot sinds 1925. Daarmee is kennis van het leven van de pionier van het volkenrecht en strijder voor tolerantie nu voor ieder bereikbaar.

De biografie is mede gebaseerd op de enorme briefwisseling (17 delen) van De Groot, die door het Huygens Instituut in samenwerking met het Instituut voor Nederlandse Geschiedenis is uitgegeven. Donner: 'Het is passend dat we deze intellectuele zwaargewicht eren met een niet minder zware biografie, die ook de vele brieven van zijn hand toegankelijk maakt'. Directeur Henk Wals van het Huygens Instituut kondigde aan dat de briefwisseling eind 2007 via internet beschikbaar zal zijn. Met anderen werkt het Huygens Instituut bovendien aan een omvangrijkere database van correspondentie van 17de-eeuwse geleerden. Daarmee



G. van Nijmegen, portret van Hugo de Groot (ontleend aan M.Jz. van Mierevelt)

worden de sociale netwerken, de verspreiding van kennis én het dagelijks leven van de betrokkenen voor wetenschappers en belangstellenden

van nu een stuk beter toegankelijk. Toekomstig biografen zullen het zo wellicht iets makkelijker hebben dan Henk Nellen. (PB)

Data infrastructuren en het Europees Kaderprogramma

Het nieuwe Europese Kader Programma (KP7) is nog maar net gestart maar er worden al volop projectvoorstellen ingediend. De eerste belangrijke deadlines zijn zelfs al weer achter de rug. Zo ook voor het onderdeel *Infrastructures* van het deelprogramma *Capacities*, waarvan de termijn voor indiening afliep op 2 mei jl.

Bij deze eerste *Infrastructures* oproep ging het om voorstellen voor verbetering van bestaande infrastructures – uitbreiding van repositories en digitale toegang tot alle mogelijke onderzoeksinfrastructures, waaronder ook databanken – maar ook om voorstellen voor de voorbereiding van grote, nieuwe infrastructures zoals die zijn opgenomen in de *European Roadmap for Research Infrastructures*.

Op het gebied van de alfa- en gammawetenschappen zijn verschillende projectvoorstellen ingediend. Bijvoorbeeld voor het CLIO-INFRA

project voor verbeterde toegang tot en behoud van historische sociaal-economische data, dat wordt gecoördineerd door het Internationaal Instituut voor Sociale Geschiedenis. En het DRIVER II project, een vervolg op het eerste *Digital Repository Infrastructure Vision for European Research* project, waarin SURF een belangrijke rol speelt, bijgestaan door een aantal andere Nederlandse organisaties.

Ook vrijwel alle ESFRI Roadmap projecten in de sociale wetenschappen en humaniora hebben een voorstel ingediend. Het gaat daarbij steeds om voorstellen voor het voorbereiden van infrastructures – het maken van gedetailleerde projectplannen voor de constructiefase en vooral ook het verkrijgen van financieringstoezeggingen hiervoor. De constructie zelf zal dan, afhankelijk van het project, over ongeveer twee jaar starten.

• Vervolg op pagina 7 •

AGENDA

8 juni

Den Haag, NWO-gebouw
Studiedag Historisch GIS
 DANS en de Vereniging voor Geschiedenis en Informatica (VGI) organiseren deze dag over de toepassing van Geografische Informatie Systemen (GIS) in historisch onderzoek.
www.dans.knaw.nl/nl/dans_symposia/2007_04/

11 juni

Utrecht, Academiegebouw Universiteit Utrecht
Subsidies voor Digitaal Erfgoed 2007
 Het plenaire programma met lezingen over ervaringen uit de praktijk wordt gevolgd door *speed dating* en themadiscussies tussen professionals uit de erfgoedinstellingen en medewerkers van diverse subsidieregelingen en fondsen. Het symposium richt zich op professionals uit cultureel erfgoedinstellingen die fondsen willen werven voor ICT-projecten.
www.den.nl/docs/20070501122230/

juni – september

Ann Arbor, Verenigde Staten
ICPSR Summer Program in Quantitative Methods of Social Research
 Aanbod van een groot aantal cursussen van hetzij vier weken of enkele dagen, op het gebied van methoden en technieken van sociaal-wetenschappelijk onderzoek. Organisator is het Interuniversity Consortium for Political and Social Research van de universiteit van Michigan. Voor deelnemers uit Nederland is via DANS een tegemoetkoming in de kosten mogelijk. Informatie kan aangevraagd worden via: info@dans.knaw.nl
www.icpsr.umich.edu/sumprog

10 juli

Nijmegen, Golden Tulip Val Monte
Themadag Gaming and the art of digital learning
 Steeds meer jongeren voeren een digitale lifestyle. Het spelen van entertainment games is daar een gewaardeerd onderdeel van. Binnen het onderwijs groeit de belangstelling voor spelend leren. Soms is het niet meer dan een vergulde pil om saai onderwijs aantrekkelijk te maken. In andere gevallen blijken 'serious games' in te zetten voor kwalitatief beter onderwijs. Succesvolle voorbeelden uit de onderwijspraktijk weten didactische principes goed te paren aan gaming kenmerken. Hoe ze dat doen komt onder meer aan de orde op deze door SURFfoundation en ISAGA georganiseerde themadag.
www.surf.nl/smartsite.dws?id=13041&showinfo=true

19 juli – 5 augustus

Ljubljana, Slovenia
ECPR Summer School in Methods and Techniques
 The European Consortium for Political Research organises its second Summer School in Methods and Techniques in cooperation with the University of Ljubljana. It is targeted at postgraduate students, teachers and researchers in political science and related disciplines.
www.essex.ac.uk/ecpr/events/summer-schools/ljubljana/index.aspx

27 augustus – 31 augustus

Tilburg, Tilburg Innovation Centre for Electronic Resources (Tilburg University)
Digital Libraries a la carte
 This modular course prepares librarians for the future. Some modules are also relevant to publishers, researchers, lecturers and IT specialists.
www.tilburguniversity.nl/services/lis/ticer/07carte/index.html

22 – 24 november

Vilnius, Vilnius University, Lithuania
Textual Scholarship and the Canon
 The fourth International Conference of the European Society for Textual Scholarship
www.textualscholarship.org/ests/conference07/index.html

Informatiemiddag van werkgroep Willem Witsen

Op 26 maart vond in *Arti et Amicitiae* op het Rokin in Amsterdam voor zo'n zeventig belangstellenden een voorlichtingsmiddag plaats over de activiteiten van de Willem Witsen Werkgroep. Een passende locatie want de schilder, etser en fotograaf was ooit bestuurslid van deze kunstenaarsvereniging. Witsen (1860-1923) was bevriend met Tachtigers als Willem Kloos, Jacobus van Looy, Frans Erens, Albert Verwey en Frederik van Eeden – de laatste twee tevens zijn zwagers – en met kunstenaars als Isaac Israëls, Jan Veth, Breitner en Toorop.

De Werkgroep Willem Witsen streeft ernaar om tekst- en beeldcollecties digitaal toegankelijk te maken als bron van onderzoek naar leven en werk van Witsen en het culturele leven van zijn tijd. De Werkgroep telt deelnemers uit zes instellingen: de Stichting Willem Witsen, het Huygens Instituut, het Rijksbureau voor Kunsthistorische Documentatie, de UB Amsterdam, de Digitale Bibliotheek voor de Nederlandse Letteren, en de Koninklijke Bibliotheek.

Marita Mathijssen, hoogleraar Moderne Nederlandse Letterkunde aan de Universiteit van Amsterdam, ver-

telde inspirerend over de boeiende persoonlijkheid van Witsen en over zijn correspondentie als rijke bron van onderzoek. Door de inlevende manier waarop zij passages uit de brieven voorlas, leek de kunstenaar bijna in levende lijve aanwezig in de sfeervolle zaal met onder andere een werk van zijn tijdgenoot Breitner aan de muur – een zaal die er niet veel anders zal hebben uitgezien in de tijd dat Witsen hier zijn kunstenaarsvrienden ontmoette.

Leo Jansen, lid van de Stichting Willem Witsen en hoofdredacteur van de brievenuitgave, vertelde over de problemen waarmee de 'transcribenten' werden geconfronteerd die met eindeloos geduld handschriften hebben ontcijferd, niet alleen van Witsen zelf, maar ook van een kleine honderd anderen die aan Witsen hun brieven richtten. En dan waren er nog de agenda's en kasboekjes die Witsen bijhield in – aldus Jansen – 'een combinatie van telegramstijl en stenografie.' Tenslotte vertelde Marianne Peereboom, projectleider bij de KB, over de verschillende lopende en geplande projecten van de Werkgroep onder meer op het gebied van digitalisering ten behoeve



Zelfportret van Willem Witsen uit het prentenkabinet Leiden

van het grote publiek en van wetenschappelijk onderzoek. (zie ook p. 3)

(Marianne Peereboom)
 Informatie: elsbeth.kwant@kb.nl

Gebruikersdag CBS: data delen, kennis delen

Begin 2007 organiseerde het Centrum voor Beleidstatistiek (CvB), dat binnen het Centraal Bureau voor de Statistiek sinds eind 2005 verantwoordelijk is voor alle microdata diensten, een gebruikersmiddag voor onderzoekers die in 2006 met CBS microdata hebben gewerkt. Met elkaar in contact te brengen van onderzoekers en ideeën opdoen om de dienstverlening te verbeteren waren het doel.

Een aantal onderzoekers presenteerde resultaten en ervaringen met de microdata. De onderwerpen waren zeer divers; er was onder andere aandacht voor stromen op de arbeidsmarkt, evaluatie van een subsidieregeling en loonstructuur en mobiliteit. Voor deze brede aanpak was opzettelijk gekozen om de diversiteit van de beschikbare data en de zeer verschillende toepassingen daarvan te illustreren.

In een paneldiscussie na afloop van de presentaties was er aandacht voor de mogelijkheden om opgedane kennis over microdata met andere onderzoekers te delen en de rol die het CBS daarbij kan spelen. Zo werd er gepleit voor een voorziening om onderzoekers ervaringen met elkaar te laten uitwisselen, naast de documentatie die nu al per bestand beschikbaar is. Een van de deelnemers aan de discussie en een zeer ervaren gebruiker van CBS data was prof.dr. Lex Borghans, die als onderzoeker is verbonden aan het Researchcentrum voor Onderwijs en Arbeidsmarkt

(ROA) van de Universiteit Maastricht. Ook hij bleek van mening dat onderzoekers hun kennis over gegevensbestanden moeten uitwisselen om zo van elkaar te kunnen leren. Deze praktijk is al veel gebruikelijker onder Amerikaanse onderzoekers, die via internet veel kennis en programma's uit wisselen. Het CBS ziet hier ook het belang van in en is intussen gestart met een project op het gebied van kennisdelen rond microdatabestanden.

@ www.cbs.nl/microdata

Seminar toont nut van EAD voor archieven

Wat heb je aan Encoded Archiving Description (EAD)? Een seminar over die vraag trok op 19 april zo'n 120 belangstellenden naar Utrecht, vooral vanuit archieven maar ook van bibliotheken en andere erfgoedinstellingen.

EAD is een manier om archieven te beschrijven, gebaseerd op de internationale standaard ISAD(G). Daarbij wordt gebruik gemaakt van XML, een markeertaal die wordt toegepast op internet. In XML worden alle elementen afzonderlijk gedefinieerd en gemarkeerd. Het gebruik van EAD maakt alle onderdelen van een inventaris altijd afzonderlijk herkenbaar, in welke taal dan ook.

Wat heeft een archiefinstelling hier nu aan? Veel, zo bleek op deze dag. Het Gemeentearchief Amsterdam gebruikt bijvoorbeeld EAD als schakel tussen het beheersysteem en

de website. Het biedt een veilige en duidelikescheidingtussendebackoffice en het publieksgedeelte en maakt bovendien mogelijk dat een zelfde inventaris in verschillende formaten getoond wordt, bijvoorbeeld als HTML, als pdf en als Excelbestand. De Zuid-Hollandse archieven zijn een samenwerkingsverband aangegaan waarbij de inventarissen van de verschillende archiefinstellingen allemaal worden geconverteerd naar EAD en op een gemeenschappelijke website kunnen worden gepubliceerd. Voor de conversie gebruiken ze scripts waarmee bestaande wordbestanden automatisch worden omgezet. Het Nationaal Archief is bezig al haar toegangen om te zetten naar EAD. De eerste 1400 staan

in Nederland en ook buitenlandse toppers gaan Nederland onderzoeken! aldus Borghans. 'Er is hier geen marktwerking en het CBS zal niet rijk worden van de stappen die het de laatste jaren heeft gezet. De onderzoekers zijn ervoor verantwoordelijk dat het CBS de credits krijgt die het verdient. Alleen zo wordt ook het maatschappelijk belang van dit werk inzichtelijk.' Suggesties voor kennisdelen: CvB@cbs.nl. (LS)

op internet. De opgedane kennis en de ontwikkelde hulpmiddelen stelt het archief vrijelijk beschikbaar voor alle geïnteresseerden via een *EAD toolbox*. Het bedrijf Pictura Imaginis demonstreerde in Utrecht ook een specifieke EAD-editor die als open source software gratis beschikbaar komt. De editor kan via het web aangeroepen worden en men hoeft bovendien geen XML te kennen om er mee te kunnen werken.

Niet alle bezoekers waren overtuigd van het directe nut van EAD, maar het seminar maakte wel duidelijk dat het gebruik steeds eenvoudiger en aantrekkelijker wordt en dat we nog veel van EAD gaan horen. (Gijsbert Kruihof)

@ www.taskforce-archieven.nl/dossiers/www.loc.gov/ead/

Kern Sociaal Statistisch Bestand beschikbaar

De kern van het Sociaal Statistisch Bestand (SSB) is sinds kort beschikbaar bij het Centrum voor Beleidsstatistiek van het CBS. Daarmee is voor veel partijen in wetenschap & beleid een belangrijke bron toegankelijk geworden.

De kern van het SSB bestaat uit onderling gekoppelde bestanden met demografische en sociaal-economische gegevens. CBS-medewerker Bart Bakker, verantwoordelijk voor de ontwikkeling van het SSB, licht toe: 'Hoewel de naam dat wel doet vermoeden, is de SSB-kern niet één fysiek bestand. Uit de gekoppelde registers wordt per eenheid een bestand gemaakt: een banenbestand van werknemers in Nederland, een bestand met zelfstandigen, één met werknemers die in het buitenland werken, zes met gegevens over uitkeringen en een personenbestand waarin demografische informatie is opgenomen'. Die informatie komt voornamelijk uit de *Gemeentelijke Basisadministratie persoonsgegevens* (GBA). Bakker: 'Van alle personen die vanaf 1995 tot de geregistreerde bevolking van Nederland hebben behoord of behoren zijn die gegevens beschikbaar. Bovendien zijn alle veranderingen daarin, zoals verhuizingen en echtscheidingen, ook opgenomen'. Daarnaast bevat de SSB-kern veel informatie over werknemers. 'Daarbij wordt gebruik gemaakt van de gegevens uit de Verzekerenadministratie van het Uitvoeringsinstituut Werknemersverzekeringen (UWV), de loonbelastinggegevens van de Belastingdienst en het bedrijfsenquête 'Enquête Werkgelegenheid en Lonen', aldus Bakker. 'Ook over zelfstandigen en mensen met

een uitkering worden verschillende registers gebruikt, onder andere van de Belastingdienst en het UWV'.

Door de registers onderling te koppelen is een groot aantal gegevens over personen, banen en uitkeringen integraal beschikbaar voor het samenstellen van statistische overzichten. De informatie is kwalitatief goed, voor meerdere jaren beschikbaar en longitudinaal gecorrigeerd, waardoor mensen gevolgd kunnen worden in de tijd. Bakker: 'Zo kan

worden nagegaan wie er uitstromen uit de WAO en wie daarvan een baan vindt'. Bovendien behoren ook duuranalyses tot de mogelijkheden. Daarmee kunnen vragen worden beantwoord over hoe lang iemand een werkloosheidsuitkering heeft en of dat verschilt tussen verschillende bevolkingsgroepen. Ook kan meer en betrouwbaarder informatie worden samengesteld naar kleine regionale eenheden en over kleine groepen in de samenleving. 'Voorbeelden hier-

van zijn in- en uitstroom van werkzame personen in de gezondheids- en welzijnszorg en het percentage van de bevolking van de gemeente Pekela met een werkloosheidsuitkering', aldus Bakker. Op dit moment is voor de verslagjaren 1999-2004 de SSB-kern bij het Centrum voor Beleidsstatistiek beschikbaar. 2005 komt naar verwachting in juli 2007 beschikbaar. Daarna zijn jaarlijks aanvullingen te verwachten. (LS)

www.cbs.nl/microdata

EU-project Planets voor digitale duurzaamheid

Europese archieven, bibliotheken, erfgoed- en wetenschappelijke instellingen worden geconfronteerd met een toenemende stroom aan digitale gegevens. Ondanks de vele inspanningen die al werden geleverd, zijn we nog steeds onvoldoende in staat om deze informatie veilig te bewaren en toegankelijk te houden. Op het ogenblik worden de kosten van verlies aan digitale informatie binnen de EU geschat op circa drie miljard euro per jaar. Om dit probleem het hoofd te kunnen bieden, investeert de Europese Commissie fors in projecten die een bijdrage leveren aan digitale duurzaamheidsoplossingen.

Het Planets project, een initiatief van de Europese Unie (EU) met een looptijd van vier jaar, is bedoeld om procedures, methoden en technieken, instrumenten en programma's

te ontwikkelen die de langetermijnbewaring en permanente toegankelijkheid van digitale informatie verzekeren. Naast de ontwikkeling van een testbed, communicatie, training en public relations, en een apart project waarin het overkoepelende kader wordt ontwikkeld, geven drie subprojecten hieraan invulling. Het project *Preservation Planning* maakt het voor organisaties mogelijk om concrete vragen over duurzaamheid in een duidelijk contextueel kader te zetten en te beantwoorden. *Preservation action* is het subproject dat invulling geeft aan de technische kant van dit antwoord: zijn er *tools* om de *preservation action* mee vorm te geven? Zo ja, hoe ga je na welke oplossing de voorkeur geniet, en waarom? En als er nog geen gepaste *tool* voorhanden is, aan welke eigenschappen zal een nieuw te bouwen *tool* dan moeten voldoen? Het project *Characterisation* tot slot beschrijft objectformaten zodat toegankelijkheid en leesbaarheid

van objecten gegarandeerd kunnen worden.

In Planets staat daarbij de praktische benadering centraal. Via implementatie van al bestaande standaarden en vertaling ervan naar bruikbare software toepassingen hoopt het Planets consortium in de nabije toekomst aan bibliotheken, archieven, erfgoed- en wetenschappelijke instellingen bruikbare oplossingen en programma's aan te bieden.

Binnen het Planets project werken zestien nationale bibliotheken en archieven, universiteiten en software bedrijven samen. In Nederland nemen de Koninklijke Bibliotheek en het Nationaal Archief deel. Beide instellingen hebben een ruime expertise op het vlak van digitale duurzaamheid opgebouwd. In Planets richten ze zich met name op de verdere ontwikkeling van methoden en technieken om digitaal materiaal leesbaar en toegankelijk te houden. (Hans Hofman, Frank Houtman)

www.planets-project.eu

Collecties Witsen open

Een aantal belangrijke tekst- en beeldcollecties met werk van de schilder, etser en fotograaf Willem Witsen wordt gedigitaliseerd met financiële hulp van het papieren-erfgoedprogramma Metamorfoze.

Dat vertelde Marianne Peereboom, projectleider bij de Koninklijke Bibliotheek (KB), eind maart op een voorlichtingsbijeenkomst van de werkgroep Willem Witsen in Amsterdam (zie p.2). Het gaat onder meer om grote collecties correspondentie van en aan Witsen uit de KB (circa 1300 brieven) en de UB Amsterdam (circa 400), een collectie uniek contemporair fotomateriaal uit het archief van het Rijksbureau Kunsthistorische Documentatie, glasnegatieven van door Witsen gemaakte foto's uit het Prentenkabinet Leiden, en verscheidene kleinere brievencollecties. Eind 2007 komen ze beschikbaar in het Geheugen van Nederland. Daarnaast maakt de Werkgroep een nieuw projectplan

om ook een wetenschappelijke toegang op het materiaal te realiseren, met uitgebreidere zoek-, navigatie- en presentatiemogelijkheden, en gekoppeld aan een werkomgeving die bijvoorbeeld het gezamenlijk editeren en annoteren van teksten en andere onderzoeksactiviteiten mogelijk moet maken. Deze toegang wordt net als het Geheugen gebaseerd op open standaarden zoals XML, OAI-PMH en SRU, aldus Peereboom. Daardoor kan hij in de toekomst gemakkelijk worden uitgebreid met nieuwe collecties, bijvoorbeeld over de Tachtigers of de Haagse School, om zo uit te groeien tot toegang en forum voor breed en veelzijdig interdisciplinair onderzoek naar deze culturele bloeiperiode.

Persoonsgegevens VOC

Sinds medio maart zijn de persoonlijke gegevens van mensen die in dienst zijn geweest van de Verenigde Oost-Indische Compagnie (VOC) online te raadplegen. De databank is nog niet compleet. In 2012 moeten alle 655.000 personen te vinden zijn die tussen 1700 en 1795 in dienst van de VOC zijn geweest. Op de speciale website van het Nationaal Archief zijn nu al de gegevens van vijf van de zes kamers van de VOC vrijwel compleet te raadplegen: Delft, Hoorn, Rotterdam, Enkhuizen en Zeeland. Met financiële steun van NWO worden de gegevens ingevoerd van de grootste kamer van de VOC, Amsterdam. De persoonsgegevens zijn afkomstig uit de *scheepssoldijboeken*, de personeelsadministratie van de VOC. <http://vocopvarenden.nationaalarchief.nl>

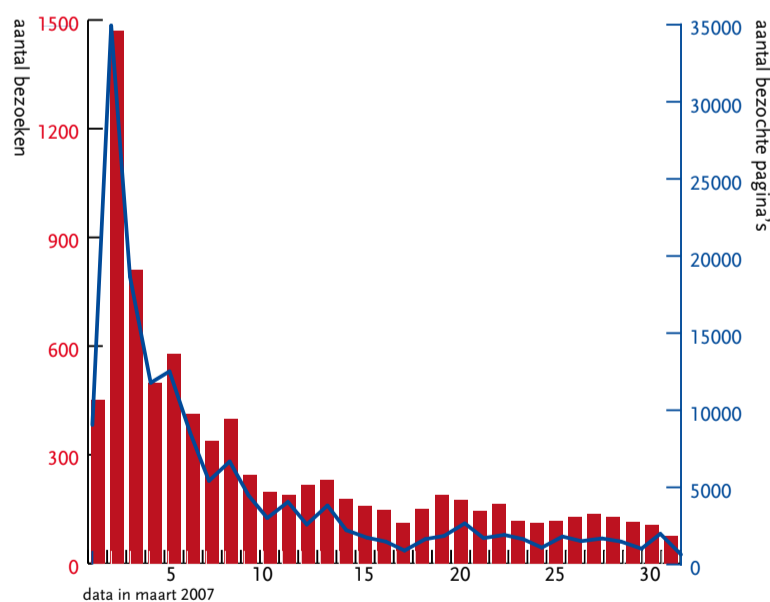
Leerstoel UU en CBS

Met zijn oratie 'Between chaos and creation' accepteerde prof.dr. Ger Sniijkers op dinsdag 27 maart het ambt van bijzonder hoogleraar aan de Universiteit Utrecht. De bijzondere leerstoel Methodenleer van bedrijfsonderzoek is tot stand gekomen door samenwerking tussen de universiteit en het Centraal Bureau voor de Statistiek. Prof. dr. Wubbels, vice-decaan van de faculteit Sociale Wetenschappen, benadrukte in zijn toespraak het belang van die samenwerking. In de komende vier jaar zal Sniijkers zich richten op verdere professionalisering van dataverzameling bij bedrijven. Dat moet leiden tot minder lastendruk, maar mag niet ten koste gaan van de economische data die het CBS krijgt (Catalogus bedrijfsgegevens CBS: www.cbs.nl/microdata). Sniijkers gebruikt een zeer brede multidisciplinaire aanpak waarin naast de bedrijfseconomische studies ook gedragswetenschappen een belangrijke rol spelen.

Emulatieproject

Volgende maand publiceren de Koninklijke Bibliotheek en het Nationaal Archief de resultaten van een project, waarin gewerkt wordt aan emulatie als concrete oplossing voor digitale duurzaamheid. Ze worden gratis en als *open source* beschikbaar gesteld. In 2005 begonnen beide instellingen aan het project. De Koninklijke Bibliotheek en het Nationaal Archief werken aan verschillende oplossingen om de dreiging af te wenden dat documenten in de toekomst niet meer te openen zijn of programma's niet meer te gebruiken. Met emulatie, een van die oplossingen, kan een computerplatform of -programma in software worden nagebootst op een ander platform of programma. Hierdoor kunnen documenten worden geopend of programma's worden gedraaid op een computer die hier eigenlijk niet voor bedoeld is. Bovendien is deze oplossing niet direct meer afhankelijk van hardware, maar van een software programma: de emulator. (Jeffrey van der Hoeven) www.kb.nl/hrd/dd/dd_projecten/projecten_emulatie.html

Gruuthuse trekt extra bezoek



Op 15 februari maakte de Koninklijke Bibliotheek bekend dat ze 'het laatste topstuk van de middeleeuwse cultuur in particuliere handen' had verworven: het Gruuthuse-handschrift. Vanaf 1 maart zou het stuk integraal te zien zijn op de KB-website, zo werd aangekondigd (zie ook pagina 4). Liefhebbers konden zich dus warmlopen en dat deden ze ook. In de middag van 1 maart werd het handschrift op de site gezet en diezelfde dag nog meldden zich 450 belangstellenden. Een dag later vertoonde de extra aanloop een

piek met bijna 1500 bezoekers. In de daarop volgende maand daalde het bezoek gestaag maar bereikte niettemin een daggemiddelde van 274. Opvallend is volgens KB-webmaster Jan van Oorschot dat de Gruuthuse-bezoekers gemiddeld op veel meer pagina's komen dan de andere webgasten: ruim 18 tegen 3,5. Relatief is het aantal bezochte Gruuthuse-pagina's dan ook nog indrukwekkender. Over de hele maand bedroeg dat een vijfde van het totaal aantal bezochte KB-pagina's. (MdG)

Musicoloog, speelman en databankier Louis Grijp: 'Ik was er vroeg bij en ik ging fanatiek door'

MARTIJN DE GROOT

Louis Peter Grijp ontwikkelde in zijn studententijd een liefde voor oude Nederlandse muziek en ging die ook spelen met het ensemble Camerata Trajectina. Hij hield dat tot op de dag van vandaag vol en zag tegelijk kans zijn liefhebberij om te zetten in wetenschappelijke productie.



DIEDERIK VAN DER LAAN

Louis Grijp bij de inmiddels weggedigitaliseerde kaartenbakken van het oude Volksliedarchief

'Hier neem dit lied bijvoorbeeld: *Toen ik op Neerlands bergen stond*. Datzelfde liedje is vierhonderd jaar geleden ook gezongen en het is al die tijd blijven bestaan. Vierhonderd jaar! Dat is ongelooflijk, dat zoiets al die eeuwen overeind is gebleven en dan overwegend gewoon door mondelinge overlevering. Vooral de teksten trouwens, die zijn het taaiste. Ongelooflijk!'

Als Louis Peter Grijp deze getuigenis van zijn gedrevenheid aflevert loopt zijn interview met *e-data&research* alweer ten einde. Maar een mooiere samenvatting van water schuil gaat achter zijn wetenschappelijke en uitvoerende werk had hij moeilijk kunnen geven. Grijp is van huis uit musicoloog, werkt sinds de jaren negentig bij het Meertens Instituut als onderzoeker en tegenwoordig als hoofd van het Documentatie- en Onderzoekscentrum voor het Nederlandse lied.

'Ongelooflijk dat zoiets eeuwen overeind is gebleven door mondelinge overlevering'

Hij bekleedt sinds 2001 een bijzondere leerstoel in de muziekwetenschap aan de Universiteit van Utrecht. Daarnaast is hij artistiek leider van het muziekgezelschap *Camerata Trajectina* dat naam heeft gemaakt met Nederlandse muziek van de middeleeuwen tot aan de Gouden Eeuw.

Liederen uit die laatste periode hebben altijd zijn bijzondere belangstelling gehad, maar inmiddels heeft

Grijp het bereik van zijn deskundigheid en betrokkenheid uitgebreid tot de hele muziekgeschiedenis van ons land. Zijn hoofdredacteurschap van de in 2001 verschenen 'Muziekgeschiedenis der Nederlanden' getuigt ervan.

'Opvallend is dat het juist de tekst is die vaak redelijk ongeschonden stand houdt door de eeuwen heen, en dan gaat het bijna altijd om ballades, tragische verhalende liederen. Ik heb pas uitgezocht hoeveel liederen uit de middeleeuwen tot in onze tijd nog te vinden zijn gebleven. Het zijn er een goede twintig, en allemaal ballades die grotendeels door mondelinge overlevering in stand zijn gebleven'. Dat Grijp zoiets kan uitzoeken is één van de verworvenheden van de Nederlandse Liederbank van het Meertens Instituut, die inmiddels zo'n 125 duizend beschrijvingen bevat van liederen met tekst en melodie en die deze maand online gaat. De musicerende onderzoeker stond zelf in de jaren tachtig aan de wieg van het project, bleef er intensief bij betrokken en kan inmiddels vaststellen dat de Liederbank qua opzet en omvang uniek is in de wereld: 'Ik was er vroeg bij en ik ben fanatiek doorgedaan'.

Hoe de Liederbank werkt wil Grijp graag laten zien. 'Hier heb ik bijvoorbeeld de *Geeraerd van Velzen* van P.C. Hooft, een toneelstuk over de moord op Floris de Vijfde'. Was dat niet 'der keerlen god', door de edelen omgebracht?

'Ja, en daar speelden natuurlijk nog allerlei andere motieven door-

heen die je in dit stuk ook tegenkomt. Als je nu naar de verzen kijkt dan zie je eigenlijk al aan die lange regels dat het om alexandrijnen gaat, achter elkaar door, bladzij na bladzij. Maar opeens kom je dan een stuk tegen met veel kortere versregels, en ook nog verschillend van lengte. Hier: (leest voor) 'O *Heyl'ghe Drietal Welkom/Die 't wereldlijk bezwaren/Komt ontvaeren*'. Vier, drie, twee beklemtoonde lettergrepen en dat gaat zo verder. Dat moet een lied zijn, dat kan je vrij snel zien. Maar hoe kom je nu aan de melodie van dat lied als je die niet weet? Nou dat kan dus met deze Liederbank'.

Grijp neemt plaats achter zijn Mac ('Zeer gebruikersvriendelijk platform, ik zou er moeilijk buiten kunnen!') en is al in het systeem verdiept nog voor dat hij zijn gesprekspartner heeft kunnen uitnodigen om mee te kijken. Hij vult in: 4, 3, 2, 4, 3, 2, 4, 4, 3, 3, 2. 'De beklemtoonde lettergrepen'. En dan: A, b, b, A, b, b, C, C, d, d, d. 'Dat is het rijmschema. De hoofdletters duiden op manlijk rijm, dus met de klemtoon op de laatste lettergreep, en de kleine letters op vrouwelijk rijm'. Nadat op deze manier alle relevante metadata zijn ingevoerd drukt Grijp op de knop 'Zoek' en daar komen ze: acht vermeldingen van liederen met exact hetzelfde *strofeschema*, met daarbij de titel, de eerste regels, de auteur als die bekend is, en nog een aantal voor de ongeofende toeschouwer onbegrijpelijke data. 'En kijk hier', wijst de musicoloog, 'op deze regel staat de oudste bekende versie van

het lied'. Eerder had hij al uitgelegd dat de zogenaamde contrafactuur – een nieuwe tekst op een bestaande melodie – in de Nederlandse geschiedenis een ook door grote schrijvers veel beoefend genre is. 'En die oudste versie is bij de meeste vermeldingen hetzelfde: *De lustelijke mei is nu in de tijd*. Nou dan heb je gevonden waar naar je zocht'.

Een prachtig instrument, zoveel is duidelijk. Maar ook een instrument dat vele arbeidsuren heeft opgeslokt voordat het tot wasdom kon komen. Grijp ontkent het niet. 'Het is eigenlijk in mijn studietijd begonnen, toen ik voor Camerata Trajectina regelmatig op zoek was naar muziek van oude liederen. Geuzenliederen zocht ik toen bijvoorbeeld. Daarvoor moest je hier op het Meertens Instituut zijn want hier stond een muur vol met kaartenbakken met oude muziek. Een unieke verzameling, maar wel allemaal op kartonnen kaartjes. Ik was vaste klant. Als je wist wat je zocht dan ging het vaak goed, want op die kaartjes stonden wijsaanduidingen – 'op de wijs van' – en zo kon je de bronmelodie vinden'.

'Hier op het Meertens stond een muur vol met kaartenbakken met oude muziek. Ik was vaste klant'

Maar vaak stond het er ook niet bij of was er gewoon geen wijs bekend. De jonge musicoloog, die in die tijd juist een vakje computerlinguïstiek deed, wilde met zo'n doodlopend spoor geen genoegen nemen en begon een databank op te zetten. 'Eerst had ik er driehonderd liederen in zitten en daar schreef ik een artikel over. Dat viel in goede aarde. Toen kon ik er een project van maken en met de hulp van student-assistenten werden er vijf duizend werken ingestopt. Daarop heb ik mijn dissertatie gebaseerd, die ging over de contrafactuur in het Nederlandse lied van de Gouden Eeuw'. Bij het Meertens Instituut, waar Grijp aan het begin van de jaren negentig kwam te werken, werd vervolgens stapsgewijs steeds een

cohort aan de databank toegevoegd. 'Met behulp van een aantal stagiaires konden we er tienduizend bij invoeren. Daarna kwam er extra geld waardoor we naar de dertig duizend doorschoten. De databank was toen zo goed als volledig tot aan 1600, redelijk gevuld tot 1700 en daarna zakte het in'. Uiteindelijk werd ook de bijna onafzienbare taak ter hand genomen om de muur met kaartenbakken 'in te kloppen'. Dat alles gebeurde met financiële bijdragen uit diverse hoeken waarbij met name NWO zich niet onbetuigd liet, zo blijkt uit Grijps chronologische beschrijving.

Reden voor voldaanheid is er echter nog lang niet. Grijp: 'Het vreemde is dat het nog steeds om een metadatabank gaat. Wij hebben de eerste regel, we hebben de wijsaanduiding, maar we hebben de melodie en de teksten niet!' Dat laatste geldt overigens niet voor het aan de Liederbank gerelateerde project *Straatliederen* waarbij – hoe kan het anders – ook het Meertens Instituut en zijn muziekprofessor weer nauw betrokken waren. Het merendeel van de liederen in de database blijft echter niet op tekst doorzoekbaar, en Grijp zou zichzelf verloochenen als hij daar geen verandering in zou brengen. 'Wij zijn juist dezer dagen bezig met het voorbereiden van een aanvraag bij NWO voor het project *Dutch songs online* dat bedoeld is om de Liederbank doorzoekbaar te maken op teksten en melodieën. Dat doen we samen met partners als de KB, de Universiteit Utrecht en de Digitale Bibliotheek der Nederlandse Letteren (DBNL). Straks willen we de zoekmogelijkheden en de omvang van de Liederbank combineren met de digitaliseringskracht van de DBNL'.

En dan ontbreekt er nog één stukje aan de kennispuzzel rondom het Nederlandse lied: het zoeken op melodie. Zoeken op de naam van het lied zit natuurlijk in het systeem, maar als je nou eens kon zoeken op de noten... Grijp wenkt de bezoeker mee de lange schemerige gang op van het Meertens instituut. Een paar kamers verder gooit hij de deur open: 'Goedemorgen! Kijk, hier zitten ze. Drie jonge onderzoekers, twee Duitsers en een Nederlander. Die werken aan het project *WITCHCRAFT*, dat valt onder het NWO-CATCH programma. Zij proberen een zoekgereedschap te maken voor de Liederbank waarbij de hoorbare melodie en de variaties daarin als aangrijpingspunt dienen'.

Prof. dr. L.P. Grijp is onderzoeker bij het Meertens Instituut en bijzonder hoogleraar aan de Universiteit Utrecht. Hij studeerde musicologie aan die universiteit en luit aan het Koninklijk Conservatorium. Hij promoveerde in 1991 op een proefschrift over het Nederlandse lied in de Gouden Eeuw en is buiten de wetenschappelijke sfeer bekend als hoofdredacteur van het standaardwerk 'Een muziekgeschiedenis der Nederlanden'.

'Aandacht voor digitale output archeologie'

Permanente aandacht is nodig voor het conserveren van de digitale output van het archeologisch onderzoek. Die aanbeveling, die heel direct raakt aan de doelstellingen van DANS, was een van de zeventien die de Koninklijke Nederlandse Akademie van Wetenschappen (KNAW) op 12 april presenteerde in een verkenning van de toekomst van de Nederlandse archeologie.

De afgelopen tien jaar is door het verdrag van Malta een nieuwe situatie ontstaan in de Nederlandse archeologie. Een nieuwe, grote groep van archeologen heeft werk gevonden in de archeologische monumentenzorg (Malta-archeologie). Gemeenten en projectontwikkelaars die bouwzaamheden uitvoeren waarbij ingrepen in de bodem nodig zijn, moeten zorgdragen voor adequaat archeologisch onderzoek. Hoewel zij zich bewust zijn van de cultuurhistorische waarde van ons bodemarchief, zullen ze toch proberen de kosten zo veel mogelijk te drukken. Malta-onderzoek vindt plaats via aanbestedingen met onderlinge concurrentie tussen de verschillende commerciële archeologische bedrijven en onder strikte tijds- en budgettaire regie. De kwaliteit van het onderzoek voldoet in de ogen van de academische archeologen echter niet altijd aan de verwachtingen die zij daarvan hebben.

Tegelijkertijd kampen de universiteiten de laatste jaren met teruglopende formatieruimte en toenemende studentenaantallen. Kennis en ervaring van de afgestudeerden zijn daardoor minder dan de nieuwe werkgevers verwachten. Ook het onderzoek aan de universiteiten komt steeds meer onder druk te staan. Juist daar moeten de vele losse waarnemingen van Malta-onderzoek worden vertaald naar een samenhangend beeld over ons verleden. Oogst van Malta onderzoek is door de beperkte kwaliteit en beschikbaarheid echt moeilijk en kan, zeker als de



PROJECTGROEP ARCHEOLOGIE BETUWELIJN (RACM/ NS RALUINFRABEHEER) Een grote groep archeologen heeft werk gevonden in de commerciële archeologie, zoals hier voor de aanleg van de Betuwelijn.

tijdelijke financiering van NWO geen structureel vervolg krijgt, maar zeer beperkt worden uitgevoerd. Terwijl die actuele kennis van ons verleden, zoals vastgelegd in de Nationale onderzoeksagenda, essentieel is voor de bedrijven bij de waardering en selectie van bedreigde vindplaatsen.

Intensieve samenwerking ligt dus voor de hand om uit deze vicieuze cirkel te raken. Toch vliegen de verwijten over en weer. Het lijkt erop dat de Nederlandse archeologie door Malta uit het lood is komen te staan. Mogelijk kunnen de aanbevelingen van de KNAW haar helpen rechtzetten. De resultaten van Malta zijn, juist vanwege het streven naar lage kosten, veelal al integraal of zelfs uitsluitend in digitale vorm vastge-

legd. De digitale gegevens vragen weliswaar om een specifieke aanpak om de leesbaarheid duurzaam te kunnen garanderen, maar maken tegelijkertijd de beschikbaarstelling via repositories als EASY veel eenvoudiger. Wetenschappers krijgen direct

en eenvoudig toegang tot de Malta-output. Het onlangs door NWO gehonoreerde eDNA II-project sluit daarbij naadloos aan: alle Malta-rapporten van de afgelopen tien jaar worden de komende twee jaar digitaal ontsloten. (Milco Wansleeben)

DRIVER: aanzet tot Europese onderzoek-infrastructuur

'Goed voorbeeld doet goed volgen'. Dit gezegde is zeker van toepassing op het Europese DRIVER project (Digital Repository Infrastructure Vision for European Research). Het voorbeeld is ons eigen Nederlandse DARE systeem dat gratis toegang geeft tot wetenschappelijke onderzoeksoutput en inmiddels permanent is ondergebracht bij de KNAW. Het DRIVER project, gestart op 1 juni 2006 met een looptijd van anderhalf jaar, maakt deel uit van het zesde EU kaderprogramma en wordt uitgevoerd door tien organisaties uit verschillende Europese landen. Nederland wordt vertegenwoordigd door de stichting SURF.

DRIVER heeft vijf doelen geformuleerd. Op de eerste plaats wordt getracht bestaande collecties met wetenschappelijke data (ook wel repositories genoemd) uit Nederland, Engeland, Duitsland, Frankrijk en België in een netwerk met elkaar te verbinden. Uiteraard bestaat de Nederlandse inbreng uit het DARE systeem. Op de tweede plaats beoogt DRIVER nieuwe diensten te ontwikkelen die gebruik maken van de data in de gedistribueerde repositories. Getracht wordt bijvoorbeeld de verschillende beschrijfsformaten te integreren in een zogenaamd metadata-register. Dit register kan bijdragen aan een betere toegang tot de verschillende repositories. Ten derde wil DRIVER user services ontwikkelen. Zo wordt bijvoorbeeld onderzocht of specifieke wensen van een gebruiker

van het DRIVER systeem kunnen worden ondergebracht in een profiel, zodat automatisch relevante gegevens uit de DRIVER repositories kunnen worden aangeboden aan deze gebruiker. Als vierde doel heeft het DRIVER project, het implementeren van een aantal relevante internationale standaarden, bijvoorbeeld op het gebied van *persistent identifiers*. Ten slotte wil DRIVER een toekomstige uitbreiding en structurele inbedding in de wetenschappelijke wereld voorberei-

Data infrastructuur en het Europees Kaderprogramma

(vervolg van pagina 1)

Stuk voor stuk zijn het voorstellen voor miljoenenprojecten, met tien of meer partners uit evenzoveel landen. Het schrijven van een dergelijk voorstel is een project op zich. Budgetten kloppend krijgen of de beoogde impact verwoorden is heel wat anders dan onderzoek doen. Na de zomer zal bekend worden welke projecten beloofd zullen worden met een miljoenensubsidie.

Voor diegenen die in deze ronde buiten de boot vallen, zijn er gelukkig nog genoeg nieuwe kansen. Nog voor het einde van dit jaar wordt de tweede oproep voor projectvoorstellen voor *e-infrastructures* verwacht. Waarschijnlijk volgt in het voorjaar van 2008 een oproep voor *integrating activities*, eveneens gericht op verbreding en verbetering van gebruik en toegang tot bestaande infrastructures. Een deel van deze laatste oproep zal specifiek gericht zijn op infrastructures die nodig zijn voor het verwezenlijken van de thematische prioriteiten van het *Cooperation* programma, waaronder ook het thema *Socio-economic Sciences and Humanities* valt. (Ellen Willemse)

den. Een aantal zogenaamde *focussed studies* is bedoeld om het draagvlak voor participatie in en gebruik van de toekomstige onderzoeksinfrastructuur te vergroten.

Het DRIVER project kan als een 'test-bed' worden beschouwd en streeft – net als het DARE project – naar een meer permanente, robuuste infrastructuur. Daartoe is recentelijk het DRIVER II voorstel ingediend bij het zevende EU kaderprogramma. (René van Horik)

GIS in de alfawetenschappen

Plaatsgebonden of geo-informatie wordt opgeslagen, beheerd, bewerkt, geanalyseerd en/of gepresenteerd met een Geografisch Informatie Systeem (GIS). Zulke systemen worden steeds meer gebruikt bij onderzoek. Het project Hisgis bijvoorbeeld maakt historisch onderzoek mogelijk op basis van de kadastrale ruimte-indeling van 1832. Nog een voorbeeld is de 'sprekende kaart' van het Meertens instituut, die Nederlands gesproken dialect toegankelijk maakt.

Het project *Alfa-GEO*, uitgevoerd door DANS, heeft als doel de drempel voor het gebruik van geo-informatie te verlagen door onderzoekers in de alfawetenschappen op de hoogte te brengen van beschikbare geo-informatie en praktische handvaten te bieden voor het gebruik ervan.

Alfa-GEO wordt voor een deel gefinancierd door het programma 'Ruimte voor GEO-Informatie' (RGI) en zal gebruikmaken van het raamwerk van standaarden voor de Nederlandse Geo-Informatie infrastructuur (NGII) dat te downloaden is via de RGI website.

Eén van de activiteiten van DANS in het kader van Alfa-GEO is de ondersteuning van de VGI studiemiddag 'Geschiedenis en Geografische Informatie systemen' die op 8 juni in Den Haag plaatsvindt. Tijdens die middag vertelt een vijftal onderzoekers (historici, geografen, economen) hoe en waarom ze geschiedenis en GIS combineren. Later dit jaar, op 16 november, wordt er in samenwerking met het Hisgis project wederom een studiedag georganiseerd. Informatie: rene.van.horik@dans.knaw.nl.

© www.hisgis.nl
www.meertens.nl/projecten/sprekende_kaart/svg/
www.rgi.nl
www.dans.knaw.nl/nl/dans_symposia/2007_04/

© www.driver-repository.eu
www.darenet.nl

COLOFON

e-data@research is het kwartaalblad over data en onderzoek in de alfa- en gammawetenschappen, verschijnend onder auspiciën van Data Archiving and Networked Services, het Huygens Instituut, het Internationaal Instituut voor Sociale Geschiedenis en de Vereniging voor Geschiedenis en Informatica.

e-data@research is ook een voortzetting van *Historia & Informatica* en van *Data News – Steinmetz Archive Newsletter*. Toezending kosteloos aan relaties van de stakeholders en op verzoek aan studenten in de alfa- en gammarichtingen. Oplage: 6500. *e-data@research* is online te raadplegen op www.edata.nl en op de websites van de deelnemende instellingen.

Uitgever: Edita-KNAW, Postbus 19121, 1000 GC, Amsterdam

Redactieadres: Postbus 93067, 2509 AB Den Haag; Anna van Saksenlaan 51, 2593 HW Den Haag; T (070)3494450 F (070)3494451 E edata@dans.knaw.nl

Redactie: Peter Boot, Martijn de Groot (hoofd/eindredacteur), Marien van der Heijden, Jetske van der Schaaf, Luuk Schreven.

Redactiesecretariaat: Lucas Pasteruining, Jetske van der Schaaf

Aan dit nummer werkten mee: Hans Hofman, René van Horik, Frank Houtman, Gijsbert Kruithof, Marianne Peereboom, Edward Vanhoutte, Milco Wansleeben, Ellen Willemse.

Vormgeving en opmaak: Ellen Bouma, Edita-KNAW

Productie en pre-press: Edita-KNAW

Druk: Plantijn Casparie, Almere

ISSN: 1872-0374

Advertentie-exploitatie: Publiciteitsbureau Meerwijk, Heinrich Böllstraat 15, 2037 MB Haarlem; T 023-5358787, 06-20115888, F 08-42252945 E fwmo@kpnplanet.nl

Revolutionaire tool of lichtzinnig speeltje?

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



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



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

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

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



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

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

Saaï...saaï? Sinds wanneer zijn



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

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

Column
Edward Vanhoutte

Humanities Computing

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

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

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

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

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

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

www.surf.nl/download/Factsheet_DARE_Entreeopmaat.PDF

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

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

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

Gelezen

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

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

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

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

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

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

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

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

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

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