

Pagina 3 • Soundbites • In het project ‘Soundbites uit vervlogen tijden’ digitaliseert het Meertens Instituut oude geluidsopnamen van dialectsprekers, onder meer om ze met behulp van een ‘sprekende kaart’ via internet beschikbaar te stellen.

Pagina 4 • Pleidooi voor een softwarearchief • Om de cultuur en geschiedenis van Nederland voor de toekomst veilig te stellen moet het Rijk zorgen dat er een softwarearchief komt, zeggen twee medewerkers van de Koninklijke Bibliotheek in een opiniebijdrage.

Pagina 5 • Oude kaarten • Een systeem waarmee oude manuscriptkaarten kunnen worden gegeoreferreerd, geannoteerd en gekoppeld aan documenten: MAPS.



Pagina 6 • FOCUS op Nicis • In de rubriek Focus wordt deze keer scherp gesteld op Nicis Institute, het onderzoeksinstituut dat zich op stedelijke problemen concentreert.

EN VERDER

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

Groen licht voor Alfalab

De Koninklijke Nederlandse Akademie van Wetenschappen (KNAW) heeft onlangs een proefproject goedgekeurd dat als aanloop moet dienen voor Alfalab, het ICT-laboratorium voor de alfawetenschappen. De aanvraag was ingediend door een aantal instituten binnen de Akademie: Huygens, Meertens, DANS, Fryske Akademy en Virtual Knowledge Studio.

Het gaat om een laboratorium waar gereedschappen op ICT-gebied kunnen worden gekoppeld aan digitale bronnen, en dat als centrum dienst doet van waaruit kennis op dat gebied kan worden gedeeld en verspreid. Het project dat nu groen licht heeft gekregen duurt anderhalf jaar en brengt kosten van bijna 1,4 miljoen euro mee. De bedoeling is dat het leidt een groter project, waarvoor dan ook financiële steun van bijvoorbeeld NWO zal worden gevraagd.

Zie pagina 4: ‘Een virtuele plek’

Complete virtuele volkstelling in 2011

De Virtuele Volkstelling is in Nederland een haalbare zaak. Dat is vastgesteld door het Centraal Bureau voor de Statistiek (CBS), dat daarmee uitvoering gaat geven aan de Europese verplichting om in 2011 een volkstelling uit te voeren. Het CBS deed een dergelijke telling als enige in Europa al eerder, in 2001.

Een meerderheid van de lidstaten pakt de volkstelling volledig op de traditionele manier aan met individuele gegevensverzameling. Anderen, zoals de Scandinavische landen en Zwitserland, kiezen voor het gebruik van overheidsregisters al of niet in combinatie met individuele verzameling van gegevens. Nederland zal als enige een complete volkstelling produceren met behulp van registerinformatie en data uit enquêtes die zodanig statistisch zijn gecombineerd en ‘opgehoogd’ dat ze op betrouwbare wijze de gehele populatie dekken. Op die manier komen gegevens beschikbaar die kunnen worden vergeleken met uitkomsten van eerdere Nederlandse Volkstellingen en met die van andere landen.

Rond de laatste traditionele volkstelling in Nederland, in 1971, stapelden vooral de bezwaren op privacygebied zich zodanig op dat werd besloten het verzamelen van gegevens op deze manier te beëindigen. Daarnaast speelden kostenoverwegingen een rol. In 1981 en 1991 waren er beperkte volkstellingen met behulp van registergegevens (bevolkingsregister e.d.) in combinatie met de toenmalige enquêtes over de beroepsbevolking en de woningbehoeften. In 2001 was er weer een grote Volkstelling, maar op ‘virtuele’ basis. Bij de komende aflevering van 2011 worden echter veel meer en gedetailleerdere gegevens verzameld. ‘Dat komt enerzijds omdat het nu een Europese verplichting is,’ zegt Eric Schulte

Nordholt, bij het CBS projectleider Volkstelling, ‘met inbegrip van alle specificaties die tot zeer gedetailleerde tabellen zullen leiden. Anderzijds zijn we nu ook technisch zo ver dat we in staat zijn om op alle gewenste gebieden gegevens op populatieniveau te produceren zonder onze toevlucht te nemen tot de traditionele dataverzameling’.

De techniek van het zogenaamde ‘herhaald wegen’ is volgens het CBS met succes gebruikt om voor de Volkstelling 2001 een set consistente tabellen op te leveren. Die techniek is nodig om steekproefgegevens te combineren met registerdata. Met de opgedane ervaring is het bureau er nu klaar voor om de virtuele volkstelling van 2011 over de gehele breedte op een verantwoorde manier uit te voeren.

De kosten van het project, waarvoor bij het verantwoordelijke ministerie van Economische Zaken zal worden aangeklopt, zullen maar een fractie bedragen van hetgeen



Eric Schulte Nordholt

Leerstoel ‘Grote Historische Databestanden’

Onderzoeker Kees Mandemakers van het Internationaal Instituut voor Sociale Geschiedenis (IISG) in Amsterdam is per 1 september voor vier jaar benoemd tot hoogleraar ‘Grote Historische Databestanden’ aan de Erasmus Universiteit Rotterdam.

Het gaat om de eerste bijzondere leerstoel in Nederland met deze leeropdracht. ‘Wij hebben het plan voor die leerstoel opgevat omdat we het groeiend belang zien van grote databestanden die steeds meer beschikbaar komen,’ licht prof. dr.

Dick Douwes toe, decaan van de Faculteit der Historische en Kunstwetenschappen die de leerstoel instelde. ‘Daar willen we als faculteit graag op inspelen en dan is het IISG een heel goede partner, die ook zelf belang hecht aan formele relaties met historische faculteiten.’ Daarnaast roemt Douwes de capaciteiten van Mandemakers, die hij als ‘zeer professorabel’ typeert.

Mandemakers is hoofd van de Historische Steekproef Nederland (HSN), die hij sinds 1991 mede van de grond hielp en waarover hij een

voor een traditionele telling nodig zou zijn. Schulte Nordholt: ‘Een telling op de vroegere manier zou zeker enkele honderden miljoenen euro’s vragen. Nu praten we over enkele miljoenen euro’s’. Afgezien van de kosten is het moeilijk om vast te stellen welke aanpak uit kwaliteitsoogpunt te verkiezen zou zijn, zegt de projectleider. ‘Een nadeel van

de virtuele aanpak is wellicht dat je je toch baseert op gegevens die deels van oorsprong op een smallere basis berusten dan die van de hele populatie. Maar daar staat tegenover dat wij (voor de registrariabelen) geen last hebben van de vertekening die door non-respons wordt veroorzaakt en dat is statistisch van grote waarde’. (MdG)

Nationaal Archief op Flickr



Het Nationaal Archief presenteerde op 4 november samen met Spaarvestad Photo zijn nieuwe fotoverzameling op de openbare fotowebsite Flickr the Commons. De drukte bij die gelegenheid gaf aan hoe benieuwd collectiebeheerders zijn naar de ervaringen tot nu toe, waarover Georges Oates (Flickr) en Fiona Romeo (National Maritime

Museum, Greenwich) vertelden. Eerder was er al veel publiciteit geweest nadat op 21 oktober de eerste vierhonderd foto’s beschikbaar waren gemaakt. Op 28 oktober had het NOS-journaal een item over het project. Tot de dag van de presentatie had de bewuste website 400 duizend bezoekers getrokken, en een veelheid aan commentaren met vaak feitelijke informatie. Tijdens de forumdiscussie bij de presentatie was er publiek gemor over de kleine beelden van Spaarvestad. ‘Je omzet beschermen mag, maar niet op Flickr’ zo werd naar voren gebracht. Het Nationaal Archief hoopt vooral op reacties op de zeer informatieve foto’s van de Arbeidsinspectie uit het begin van de twintigste eeuw. Hiervan werden de complete, onbewerkte images geplaatst, inclusief de kartonnen randen en getypte onderschriften. Of dit een goede keus was, valt te beoordelen op www.flickr.com/photos/nationaalarchief. (Yola de Lusenet)

AGENDA

1-3 december
Hilton Edinburgh Grosvenor, Scotland
Radical sharing: transforming science?
International Digital Curation Conference
www.dcc.ac.uk/events/dcc-2008/

4 december
Surfnet, Utrecht
Advanced Streaming Video
Gebruik van video op internet en online zetten van presentaties met Rich Media.
www.surfnet.nl/nl/bijeenkomsten/

5 december
FARO, Vlaams Steunpunt Cultureel Erfgoed, Brussel
Gezocht, serieuze spelletjes!?
Games en publiekswerking, het bewaren van games, de gespannen relatie tussen commerciële historische games en de historiografie, en de mogelijkheden van de erfgoedsector voor gameontwikkelaars. Buitenlandse praktijkvoorbeelden, actief testen van games door deelnemers.
www.faronet.be/agenda/gezocht-serieuze-spelletjes

7-12 december
University Place Conference Center and Hotel, Indianapolis (USA)
IEEE International Conference on e-Science
The e-Science 2008 conference is designed to bring together leading international and interdisciplinary research communities, developers, and users of e-Science applications and enabling IT technologies. <http://escience2008.iu.edu/>

9 en 10 december
De Doelen, Rotterdam
Digitaal Erfgoed Conferentie 'Naar Buiten'
De conferentie staat in het teken van de mogelijkheden om digitaal erfgoed optimaal vindbaar en bruikbaar te maken. Niet alleen de grote zoekmachines maar ook community websites, open geografische systemen en mash up websites bieden mogelijkheden om het beschikbare digitale erfgoed bij het publiek te brengen. www.den.nl/innovatie/de_conferentie2008/

11 december
Internationaal Instituut voor Sociale Geschiedenis (IISG), Amsterdam
VGI-studiemiddag Geschiedenis en geschiedenissen en uitreiking VGI Innovatieprijs
Geschiedenis en geschiedenissen zijn hot, maar betekent dit ook dat de geschiedenis er als wetenschap erop vooruit gaat? Op deze studiemiddag zal ook aandacht worden besteed aan de VGI Innovatieprijs 2008. De genomineerden geven een korte presentatie over hun applicatie waarna de jury de winnaar bekend zal maken.
www.vgi-online.nl

9 januari
Faculteit Geesteswetenschappen Universiteit Utrecht, Utrecht
Symposium ICT in de mediëvistiek, het memoria-onderzoek in Nederland
Verschillende sprekers belichten de manieren waarop ICT-toepassingen bijdragen aan het onderzoek van de middeleeuwse memoria en aan het presenteren van de onderzoeksresultaten. DANS Symposium, georganiseerd in samenwerking met Onderzoeksinstituut voor Geschiedenis en Cultuur van de Universiteit Utrecht. www.dans.knaw.nl/nl/dans_symposia/

21-22 januari
State and University Library, Goettingen
TextGrid Summit
Presentation of the Virtual Research Environment for text scientists and other specialists in the arts and humanities. International initiatives and the emerging German e-Humanities infrastructure will be discussed. Information about how TextGrid works practically and technologically and participants will be able to get acquainted with developing TextGrid tools and services in a Developers' Workshop.
www.textgrid.de/konferenzen/summit.html

3-5 februari
Convention Center Bielefeld, Duitsland
Upgrading the eLibrary: Enhanced Information Services Driven by Technology and Economics
The International Bielefeld Conference 2009 provides insights in the future of eLibraries, based on the threefold interdependency of service, technology, and economics. <http://conference.ub.uni-bielefeld.de/2009/index.htm>

Eerste NPSO workshop bij CBS

'Beperking en correctie van non-response in enquêtes' was het onderwerp van de eerste workshop van het NPSO in de prachtige nieuwe locatie van Centraal Bureau voor de Statistiek (CBS) in Den Haag-Leidschenveen. Eerder dit jaar is dit Nederlandstalig Platform voor Surveyonderzoek opgericht. Het doel van de workshop was om de methodologische aspecten van non-respons vermindering en non-respons correctie te bespreken, en om interactie tussen deze twee onderzoeksrichtingen te bevorderen. Rond veertig vertegenwoordigers van overheid, academia en bedrijfsleven uit Nederland en Vlaanderen woonden de workshop bij. De voordrachten werden verzorgd door de marktonderzoeksbureaus Synovate en TNS-NIPO, het CBS en onderzoekers van de universiteiten van Utrecht en Gent.

Tijdens de voordrachten en de discussies werd een scheiding der geesten zichtbaar: de marktonderzoeksbureaus richten hun onderzoek vooral op non-response vermindering, terwijl de academische wereld zich vooral richt op methodes voor non-respons correctie. Ook was er een pittige discussie over panelonderzoek versus aselect-



Harm Hartman (Synovate, rechts) en John Lievens Universiteit Gent, links) in discussie

te steekproeftrekking. De marktonderzoeksbureaus kiezen vaak voor panelonderzoek omdat de respons en dus de efficiëntie hoger ligt. Helaas leidt een panel wel vaak tot een slechte populatiedekking. Bij het gebruik van een aselecte steekproef is de hogere non-respons een risico. Vooral de marktonderzoeksbureaus leken behoefte te hebben aan meer wetenschappelijk onderzoek naar panels, bijvoorbeeld door kwaliteitscriteria voor panels te ontwikkelen. Het was duidelijk dat meer communicatie en samenwerking tussen de instellingen gewenst is. (Nino Mushkudiani)

@ www.npsso.net

Trends in de editiewetenschap

Om hun besluit tot samenwerking te bezegelen, hielden het Huygens Instituut (KNAW) en het Vlaamse Centrum voor Teksteditie en Bronnenstudie (CTB) op 9 oktober een studiedag 'Trends en thema's in de editiewetenschap'. Het belangrijkste thema van de dag was de digitalisering: digitale methodologie, een digitaal onderwijsplatform, digitale uitgaven, een digitaal handboek, en zo meer.

Een tegendraads en prikkelend geluid kwam van Edward Vanhoutte, de coördinator van het CTB. Voor moderne Nederlandstalige literatuur is de papieren leeseditie een geschikter vorm dan een wetenschappelijke elektronische varianteditie. Peter Robinson, één van de leidende denkers over elektronische edities, propageert edities die door een groep van belangstellenden worden beheerd en onderhouden. Volgens Vanhoutte is in ons kleine taalgebied zoiets onhaalbaar. De toegankelijke leeseditie daarentegen wordt goed verkocht – en komt wel degelijk dankzij digitale technologie tot stand.

Een veelheid aan aspecten kwam aan de orde in andere bijdragen: het gebruik van e-Laborate, het instrument voor editie- en annotatiewerk ontwikkeld aan het Huygens Instituut, in het onderwijs van editievaardigheden; de samenwerking tussen tekstbezorgers die een dergelijk webgebaseerd 'collaboratory' moge-

lijk maakt als de belangrijkste trend in de editiewetenschap; hulpmiddelen voor vergelijking van tekstversies in een digitaal archief met werk van Beckett. In de pauze werden twee projecten voor teksteditie gepresenteerd: het Tekst Laboratorium waarin wordt gewerkt aan Hugo de Groot's *De jure praedae*, en een werkeditie van de *Cronike van der Duytscher Oirden*.

Teksten hebben net als ander cultuurgoed zorg nodig om toegankelijk te blijven. In zijn openingswoord drukte Henk Wals, directeur van het Huygens Instituut, de bezoekers op het hart dat die boodschap niet luid genoeg kan weerklinken. Betrouwbare teksten komen niet zomaar tot stand. Om de editiewetenschap een herkenbaar eigen platform te geven presenteerde Peter de Bruin op de studiedag het digitale Platform Teksteditie. 'Hebben alle vogeltjes platforms begonnen, behalve jij en ik?' parafraseerde hij onze beroemdste versregel. Het Platform presenteert nieuws over projecten, edities en opleidingen op het gebied van de teksteditie, en moet uitgroeien tot een levendig forum waar onder andere tekstbezorgers, uitgevers en studenten elkaar kunnen vinden. (PB)

@ www.teksteditie.org/
www.e-laborate.nl/
www.textlaboratorium.com/

Digitale creativiteit en tekstvergelijking

Op 30 en 31 oktober vond het Academiccolloquium 'Text comparison and digital creativity' plaats, het internationale symposium van de Virtual Knowledge Studio en het Turgama project Leiden in het kader van 200 jaar KNAW. Een groep van ruim dertig geesteswetenschappers werd getraakteerd op inspirerende lezingen van vakgenoten uit een waaier aan humanities subdisciplines. De nadruk lag op 'digitale creativiteit': op welke wijzen hebben de nieuwe technische mogelijkheden het onderzoek fundamenteel veranderd? *Keynote speakers* waren de Britsetaalkundige David Crystal, die met flair onder meer zijn onderzoek naar 'digitally mediated communication' presenteerde, en Bella Hass Weinberg, die vanuit de bibliotheek- en informatiewetenschap inging op creatieve hulpmiddelen in de Hebreeuwse

traditie die in de huidige digitale wereld een grote vlucht hebben genomen. Ook in de overige twaalf lezingen speelden bijbelteksten een belangrijke rol; er werden ondermeer handschriften bekeken met nieuwe technieken die volledig nieuwe feiten zichtbaar maakten. Ook massadigitalisering versus wetenschappelijk editeren, annotatietools, knowledge production, hypertext in de zeventiende eeuw, emendatiesystematiek, en de queeste naar een wetenschappelijke methode voor het samenstellen van een tekstgenealogische stamboom kwamen aan de orde. De verscheidenheid aan onderzoeksachtergronden maakte dat de discussie zich kon concentreren op fundamentele vragen rond de aard en ontwikkeling van 'digitaal' geesteswetenschappelijk onderzoek. (Karina van Dalen-Oskam)

Massadigitalisering: meer is beter

Hoe massaal kan digitalisering eigenlijk zijn? Dat was op 12 september de centrale vraag op een studiemiddag van de vereniging voor Geschiedenis en Informatica (VGI) over massadigitalisering. Een populair onderwerp, want de zaal in de Koninklijke Bibliotheek (KB) zat bomvol.

Massa of niet, instellingen kunnen niet linksbovenaan beginnen met digitaliseren om rechtsonderaan te eindigen. Arbeidsintensieve selectie, voorbereiding, keuzes en controle blijven. Tanja de Boer, hoofd Fysiek Beheer KB, sprak over de voorbereiding die bij de KB centraal wordt uitgevoerd door haar afdeling. Efficiënter dan losse teams per project, en motiverend voor de medewerkers. Door hun kennis van materialen kunnen ze de conditie van objecten beoordelen; ze voegen targets voor het scanbedrijf toe en checken de images van dat bedrijf.

Het Stadsarchief Amsterdam laat zich in de selectie leiden door gebruikers, legde Marc Holtman



Sylvia van Peteghem

uit. Digitaliseren op verzoek is net kopiëren, en stroomlijning houdt de kosten laag. De gebruiker betaalt, maar niet meer dan nodig: 'gebruikskwaliteit', een goed leesbare tekst, is voldoende.

De Digitale Bibliotheek voor de Nederlandse Letteren (DBNL) is dé plek voor betrouwbare teksten -geen plaatjes, maar *teksten*, om te lezen, via allerlei media – tegenwoordig ook de handheld. De projectleider, Cees Klapwijk, betoogde dat digitaliseren van Nederlandse teksten voor een overzichtelijke hoeveelheid publiek geld gerealiseerd kan worden. En moet worden, want niemand anders gaat het doen.

Astrid Verheusen (Databank Digitale Dagbladen, KB) schetste hoe de KB steeds de balans zoekt tussen kwaliteit en kwantiteit. Een voorbeeld is de overstap van TIFF naar JPEG2000 als archiefformaat, wat vijftig procent bespaart op opslag. De KB wil zoveel mogelijk doen met beschikbare middelen, ook om een tegenwicht te bieden aan commerciële partijen.

Die commerciële partijen kwamen tijdens de discussie nog eventjes aan de orde, doordat Sylvia Van Peteghem (UB Gent) haar eigen positieve ervaringen met Google tegenover Verheusens samenvatting plaatste. Het principe van digitaliseren als publieke taak is mooi, maar als je het moet doen met een paar promille van de KB miljoenen, dan grijp je je kans als Google aan de deur klopt. Mits er goede afspraken te maken zijn, en daarover zijn zij in Gent zeer tevreden. Extra jammer dat de Google-contracten geheim zijn en wij niet zelf kunnen oordelen. (Yola de Lusenet)

Gegevens Kadaster beschikbaar

Geografische gegevens van het Kadaster zijn voortaan direct en gratis beschikbaar voor onderzoek en onderwijs aan universiteiten en hogescholen. Dat is het Kadaster overeengekomen met DANS (Data Archiving and Networked Services)

Bestuursvoorzitter Dorine Burmanje van het Kadaster en adjunct-directeur Henk Harmsen van DANS tekenden op 4 november op het hoofdkantoor van het Kadaster in Apeldoorn een licentiecontract. Daarin staat hoe en onder welke voorwaarden de kadastrale gegevens beschikbaar worden gesteld. 'De overeenkomst past in ons streven naar gestructureerde toegankelijkheid van geo-informatie. Daarvoor hebben we in het Kadaster een gelijkgezinde partner gevonden,' aldus Harmsen.

In universitair onderzoek en in het onderwijs bestaat grote behoefte om zonder kosten of tijdverlies over geografische data te kunnen beschikken. Kadastergegevens waren tot nu toe wel verkrijgbaar, maar dat was omslachtig en bracht kosten mee tot enkele duizenden euro's. Door de nu

gesloten overeenkomst kunnen medewerkers en studenten van universiteiten en hogescholen meteen bij DANS terecht om de gegevens binnen te halen. Het Kadaster stelt wel als eis dat de publicaties over onderzoek met behulp van zijn gegevens beschikbaar worden gesteld.

Vooreerst gaat het om het Bestand Bodemgebruik Nederland, de Kadastrale Kaart met perceelindeling, en de Topografische Kaart 'Topo-Vector'. In januari 2009 komt het verbeterde en actuele 'TOP10NL-2007' bestand beschikbaar.

DANS werkt nu verder aan de verwerving van nog meer Kadasterbestanden, waarvan bekend is dat die al langere tijd op de verlanglijstjes staan van universitaire onderzoekers en universiteitsbibliotheken. De overeenkomst Kadaster-DANS biedt daartoe de ruimte.

Dialecten beschikbaar via sprekende kaart

'Soundbites uit vervlogen tijden' heet het project van het Meertens Instituut en DANS (Data Archiving and Networked Services) waarin grote hoeveelheden geluidsopnamen van dialectsprekers worden ontsloten. Het ging 1 september van start.

Het Meertens Instituut bezit veel geluidsopnames van gesprekken tussen dialectsprekers uit alle delen van Nederland. Vlak na 1950 begon het instituut met het verzamelen, dat werd voortgezet tot in de jaren tachtig. Een kleine selectie van fragmenten is nu als 'sprekende kaart' beschikbaar op de website van het Meertens Instituut. Met het Soundbites project willen beide instituten nu de gehele collectie via internet duurzaam toegankelijk maken voor onderzoek.

Het initiatief, dat door beide instituten in het kader van het DANS-programma Kleine Dataprojecten is gefinancierd, komt precies op tijd voor de groeiende wetenschappelijke belangstelling voor dialecten. Onderzoeker variatilinguïstiek Marc van Oostendorp van het

Meertens Instituut wijst erop dat momenteel onderzoek naar dialecten gaande is aan de universiteiten van Amsterdam, Leiden, Groningen, Utrecht en Nijmegen.

De mogelijkheden voor onderzoek zijn volgens hem legio: 'Eindelijk kan er grootschalig fonetisch onderzoek worden gedaan naar variatie in bijvoorbeeld klinkerkwaliteit en klinkerduur; statistisch onderzoek naar woordfrequentie in dialecten van het Nederlands, historisch onderzoek naar de inhoud van de vertelde verhalen'. Maar ook buiten het traditionele dialectonderzoek zijn er volgens Van Oostendorp nieuwe mogelijkheden, bijvoorbeeld in de taal- en spraaktechnologie. (Douwe Zeldenrust)

@ www.meertens.nl/projecten/sprekende_kaart/svg/
www.dans.knaw.nl/nl/projectenpagina/kdp/

Nieuwe diensten in DANS EASY

DANS heeft in het elektronisch archiveringssysteem EASY nieuwe diensten ontwikkeld voor zowel dataproducten als -consumenten. EASY is het systeem waarmee onderzoekers in de alfa- en gammawetenschappen snel en eenvoudig onderzoeksdata kunnen

deponeren en downloaden.

Binnenkort wordt Digital Author Identification (DAI) aan DANS toegevoegd. Dataproducten kunnen daarmee hun persoonlijke DAI-nummer aan een databestand toevoegen zodat koppelingen met hun andere publicaties en relevante onderzoekinformatie eenvoudiger gelegd kunnen worden. Daarnaast geven de nieuw ontwikkelde *Activity logs* dataproducten inzicht in het gebruik van de data. Voor elke gedeponeerde dataset wordt weergegeven hoe vaak hij is bekeken en gedownload. Dataproducten kunnen bovendien zien wie hun data gebruikt en met die gebruikers gemakkelijk direct in contact komen.

EASY is ontwikkeld om op een duurzame manier te kunnen refereren aan datasets op het internet waardoor eenmaal gedeponeerde data ook in de toekomst via het internet vindbaar blijven, ondanks regelmatig veranderende webadressen. Dit gebeurt door aan iedere gedeponeerde dataset een zogenaamde *Persistent Identifier* te geven. Deze identificatiecode is net als een ISBN uniek. Onderzoekers kunnen in hun referenties naar gebruikte databestanden een link opnemen met de Persistent Identifier. Informatie: info@dans.knaw.nl. (Jetske van der Schaaf en Laurents Sesink)

GUY ACKERMANS



Burmanje en Harmsen met de juist getekende licentieovereenkomst

Alliance: samen voor duurzame financiering

Hoe moet duurzame toegang tot wetenschappelijke informatie worden gefinancierd? Die vraag stond op 4 november centraal op de jaarlijkse conferentie van de Alliance for Permanent Access in Boedapest.

De Alliance is het forum waarin alle belanghebbenden op dit gebied zijn verenigd. Uit Nederland nemen de Koninklijke Bibliotheek, DANS (Data Archiving and Networked Services) en de Nationale Coalitie voor Digitale Duurzaamheid (NCDD) erin deel.

Een kleine honderd onderzoekers uit allerlei disciplines spraken in Boedapest met de grote financiers van wetenschappelijk onderzoek, politici, beheerders van digitale archieven en bibliotheken over hoe verbrokkeld het veld is wanneer het gaat om duurzame toegang. Ad hoc projectfinanciering is de norm in de wetenschap en er bestaat een wirwar van nationale en internationale financieringsinitiatieven, zo werd geconstateerd. 'Er is een chaos aan repositories ontstaan,' zei een van de deelnemers.



Bijna alle sprekers op de conferentie bijeen voor een paneldiscussie: v.l.n.r. Norbert Kroo, Neil Beagrie, John Smith, John Marks, Claude Huc, Laszlo Kovacs, Niamh Brennan, Catherine Hardman, Rudolf Dimper, Graham Cameron, Keith Jeffery en Neil Williams

Interoperabiliteit en duurzame financiering zijn nodig, en geen van de partners die in Boedapest om de tafel zaten kan die alleen tot stand brengen. Alleen door gezamenlijk

optreden kan de boodschap van duurzame toegang genoeg politieke steun vinden bij nationale overheden en in Brussel, zo besloot de bijeenkomst. (IA)

@ www.alliancepermanentaccess.eu/index.php?id=3

CBS werkt aan lange tijdreeksen

Het Centraal Bureau voor de Statistiek (CBS) heeft een intern expertisecentrum opgericht om de eigen statistiekmakers te helpen bij het samenstellen van lange tijdreeksen en het repareren van breuken daarin. Ook de historische tijdreeksen zoals die nu al worden gepubliceerd op de elektronische database StatLine krijgen aandacht van het centrum. De behoefte aan het centrum kwam voort uit de groeiende belangstelling voor lange tijdreeksen en uit het besef dat historische gegevens niet altijd goed toegankelijk zijn voor de buitenwereld. Nieuwe tijdreeksen over de Nederlandse (macro) economie zijn inmiddels gepubliceerd. Begin volgend jaar verschijnt nog een aantal reeksen, onder meer op het gebied van gezondheid en van bedrijvenconjunctuur. (Ruurd Schoonhoven)

ANP radiobulletins online



Bijna vijftig jaar uitzendingen van het radionieuws zijn gedigitaliseerd en onlangs toegankelijk gemaakt door het Algemeen Nederlands Persbureau ANP en de Koninklijke Bibliotheek (KB).

'Hier is de Radionieuwsdienst, verzorgd door het ANP'. Zo opende elke radio-uitzending die het persbureau verzorgde voor de Nederlandse omroepen. De uitgetypte vellen van deze uitzendingen uit de periode 1937 tot 1984, zo'n anderhalf miljoen pagina's, zijn door de KB in samenwerking met het ANP toegankelijk gemaakt via een website. Alles is opgetekend in de bondige stijl waar het ANP om bekend stond, van uur op uur te volgen en soms voorzien van de kritische aantekeningen van de nieuwslezers. (Edwin Klijn) www.kb.nl/ANPradiobulletins/

VOC opvarenden in EASY



Sinds kort zijn bij DANS de gegevens te raadplegen van alle opvarenden die ooit in dienst waren van de kamer Enkhuizen van de Verenigde Oost-Indische Compagnie (VOC). De data zijn afkomstig uit het project 'VOC Opvarenden', in uitvoering bij het Nationaal Archief. De ontsluiting via het doe-het-zelf archiveringssysteem DANS EASY is vooral gericht op dewetenschappelijke onderzoeker. De Kamer Enkhuizen is nog maar het begin van het project. Het zal enige jaren duren voordat de data over de in totaal 655.000 VOC-personeelsleden compleet zijn.

@ <http://easy.dans.knaw.nl>

Pleidooi voor een softwarearchief

JEFFREY VAN DER HOEVEN EN FRANK HOUTMAN

Toegankelijkheid en hergebruik van informatie wordt alom aangemoedigd. De aantoonbare authenticiteit van deze informatie speelt hierbij een grote rol, óók bij informatie in digitale vorm. Maar hoe realistisch is de notie van authenticiteit en toegankelijkheid als de oorspronkelijke hardware en software niet meer beschikbaar zijn? Jeffrey van der Hoeven en Frank Houtman van de Afdeling Digitale Duurzaamheid van de Koninklijke Bibliotheek pleiten voor een softwarearchief en vragen de overheid hiertoe het initiatief te nemen.

Voor het probleem van in onbruik geraakte hardware is inmiddels een oplossing gevonden: emulatie. Dit is specifieke software die de functie van hardware kan nabootsen en het mogelijk maakt zonder oude hardware toch de originele programma's te gebruiken. Het gaat hier niet alleen over software die jaren geleden erg gangbaar was (WP5.1, MS-DOS, dBase, Lotus 1-2-3, etc.) maar ook over bedrijfs- of onderzoeksspecifieke toepassingen die op maat gemaakt zijn en uniek in de wereld. Wanneer die toepassingen niet goed meer werken op een moderne computer, wordt emulatiesoftware ingezet waarmee de informatie in de oorspronkelijke digitale omgeving kan worden geraadpleegd.

Voorwaarde is wel dat over de oorspronkelijke software beschikt kan worden. En daar doet zich een groot probleem voor. Sommige organisaties zullen wellicht een kartonnen doos met oude software

bewaren, maar geen enkele organisatie zal beschikken over alle software die zij ooit heeft gebruikt. Laat staan dat er een organisatie bestaat die alle in Nederland uitgebrachte software zoals besturingssystemen en applicaties structureel en systematisch bewaart.

Zonder software kan geen enkele oorspronkelijke computeromgeving worden gereconstrueerd, en juist die omgeving is zo bepalend voor de beleving van oorspronkelijke digitale documenten. Een CAD-tekening waarvan de kleuren anders worden weergegeven dan oorspronkelijk bedoeld, kan al onbegrijpelijk worden. Een website uit 1996, destijds ontwikkeld voor Netscape of Internet Explorer 4.0, ziet er compleet anders uit in huidige browsers. Computermodellen die tien jaar geleden voor MS-DOS werden ontwikkeld zijn zeer waarschijnlijk niet meer correct uit te voeren door nieuwe besturingssy-



HENK VAN RUITENBEEK

stemen. Met andere woorden: de oorspronkelijke softwareomgeving is naast hardware essentieel voor correcte toegang tot overheidsdocumenten, onderzoeksdata en andersoortige digitale informatie.

Waarom bestaat er eigenlijk nog niet zo'n softwarearchief? Ten eer-

ste zijn er technische obstakels, zoals de complexiteit van een softwareomgeving: niet alleen de toepassing moet bewaard worden, maar ook het besturingssysteem, extra drivers, plug-ins, lettertypes, etc., tezamen met de bijbehorende metadata. Een tweede probleem

wordt gevormd door wettelijke regels, omdat voor veel software een gebruikslicentie nodig is of omdat Digital Rights Management (DRM) het onmogelijk maakt de data van de originele drager te kopiëren. En dan is er nog het probleem dat software zo divers van aard is: er bestaat niet één versie van MS Office, maar er zijn honderden varianten in verschillende talen, meeteenheden enzovoorts.

Ondanks dat er geen softwarearchief bestaat, verplicht de Archiefwet iedere overheidsorganisatie de voor de cultuurhistorie relevante (al dan niet digitale) informatie in zijn oorspronkelijke vorm en context te bewaren en over te brengen. Voor informatie op traditionele dragers worden klimatologisch gereguleerde depots gebouwd die beveiligd zijn tegen brand- en waterschade. De toegankelijkheid van digitale informatie is echter niet of nauwelijks gewaarborgd. Om de cultuur en geschiedenis van Nederland ook in digitale vorm voor toekomstige generaties veilig te stellen en toegankelijk te houden is het een absolute voorwaarde dat er een softwarearchief komt. De rijksoverheid zou hierin het initiatief moeten nemen – al is het maar omdat zij als grote partij bij aanbestedingen de mogelijkheid heeft om licenties van softwarepakketten voor dit doel te reserveren.

Alfalab: ‘Een virtuele plek om elkaars bestanden en tools te gebruiken’

PETER BOOT

De geesteswetenschappen veranderen niet snel genoeg onder invloed van ICT-ontwikkelingen, onder andere door de kleinschaligheid van veel onderzoek. Een gedeelde digitale infrastructuur kan daarin verbetering brengen, staat in de aanvraag voor Alfalab, een plan dat onlangs door de Koninklijke Nederlandse Akademie van Wetenschappen (KNAW) is goedgekeurd.

Het plan behelst de inrichting van een laboratorium om *ICT-tools* aan digitale bronnen te koppelen, een portaal om de bronnen te ontsluiten, en een ‘disseminarium’ om de nieuw verworven kennis onder grotere groepen onderzoekers te verspreiden. Het is aangevraagd door een serie KNAW instituten (Huygens, Meertens, DANS, Fryske Akademy, Virtual Knowledge Studio). Ook andere instituten van binnen en buiten de KNAW nemen deel. *e-data&research* vroeg Henk Wals, directeur van het Huygens Instituut en voorzitter van de Alfalab-stuurgroep, om een toelichting.

Waarom is de voorgestelde investering nodig? Is er iets verkeerd aan de manier waarop de geesteswetenschapper nu werkt?

Wals: Er is niets verkeerd aan de manier waarop de geesteswet-



Henk Wals

schapper nu werkt, maar informatie-technologie en digitalisering bieden wel nieuwe mogelijkheden. Grote hoeveelheden data en tekst kunnen snel doorzocht worden, conclusies kunnen daardoor worden gebaseerd op meer materiaal, hulpmiddelen voor analyse en visualisatie verschaffen nieuwe inzichten en suggereren

nieuwe ideeën. Veel geesteswetenschappers zijn hiermee bezig. Wat zij ontwikkelen levert het nodige op voor hun eigen onderzoek, maar blijft buiten hun discipline vaak onbekend. De geesteswetenschappen zijn nu eenmaal versnipperd. Met Alfalab willen we een virtuele plek scheppen waar onderzoekers uit verschillende disciplines van elkaars bestanden en tools gebruik kunnen maken. Dat is niet alleen efficiënt, maar kan ook tot verrassende nieuwe resultaten leiden.

Wat wordt er mogelijk dankzij Alfalab? Zijn er voorbeelden te geven van hoe de praktijk van het wetenschappelijk onderzoek zou veranderen?

Wals: Simpel gezegd brengt Alfalab bestanden en tools bij elkaar en stemt deze op elkaar af. Onderzoekers kunnen dan kiezen welke ze voor hun onderzoek willen gebruiken. Ideale-

ter kunnen ze dan meteen paar muis-klikken aan het analyseren slaan. Zelf ben ik historicus. Stel, ik wil een biografie van Willem van Oranje schrijven. Als ik dan in Alfalab zijn gedigitaliseerde brieven aantref en vervolgens een *tooltje* vind waarmee ik chronologisch kan weergeven waar hij verbleef, met wie hij contact had en wat de onderwerpen waren die hij in zijn brieven aanroerde, dan ben ik al een stuk op streek. In één keer heb ik dan het overzicht. Dat scheelt zeeën van tijd. Technisch is dat geen toekomstmuziek meer.

Waarom moet de KNAW hiertoe het initiatief nemen? Is dit niet een nationale aangelegenheid?

Wals: De KNAW staat voor het onderzoek in Nederland en dus ook voor de vernieuwing daarvan. *Computational humanities* gaan de komende jaren een belangrijke rol spelen bij

een aantal KNAW-instituten. Het is dus niet zo gek dat uitgerekend van die kant het initiatief komt voor een nieuw stukje onderzoeksinfrastructuur op dit terrein. Maar de eerste fase van Alfalab is niet meer dan een pilot, die wil laten zien hoe het kan werken. De volgende stap is om Alfalab en de financiering daarvan naar een nationale schaal te tillen. Daarvoor zal een breder consortium gevormd moeten worden.

Waarin verschilt Alfalab van infrastructuur als DARIAH of Clarin?

Wals: Tja, het is soms lastig om dat allemaal uit elkaar te houden. DARIAH is een Europese infrastructuur die toegang wil bieden tot geesteswetenschappelijke onderzoeksbestanden. Alfalab wil dat doen voor Nederland, maar legt meer de nadruk op de combinatie van bestanden en tools. Alfalab en DARIAH zullen op elkaar aansluiten. In beide projecten speelt het Nederlandse DANS een voortrekkersrol. Clarin is ook een Europees project dat overlap vertoont met DARIAH en Alfalab, maar is vooral gericht op de linguïstiek. Ook hierin speelt Nederland een belangrijke rol. Er zitten ook Clarin-mensen in de Alfalab-groep en omgekeerd. Het vult elkaar allemaal aan, we concurreren niet.

MAPS, een project voor manuscriptkaarten

LOUISA BALK

Het Nationaal Archief beheert bijna 300.000 kaarten en tekeningen van het gewest Holland, Nederland en de voormalige overzeese gebiedsdelen – vaak unieke, handgetekende zogenaamde manuscriptkaarten. Als een subsidieaanvraag bij CATCH wordt gehonoreerd, kan het systeem MAPS worden ontwikkeld waarmee de kaarten kunnen worden gegeorefereerd, geannoteerd en gekoppeld aan documenten.

De kaarten horen bij dossiers met besluitvorming van bijvoorbeeld Rijkswaterstaat, Defensie en de Rijksgebouwendienst, maar ook bij de handelscompagnieën zoals de Verenigde Oost-Indische Compagnie. De collecties, uit de periode van het eind van de 15e eeuw tot rond 1950, zijn summier beschreven in inventarissen of catalogi.

Manuscriptkaarten zijn vaak moeilijk toegankelijk. Door hun formaat werden ze apart opgeborgen, waardoor ze losraakten van de archieven waar ze eigenlijk bij hoorden. Ze hebben geen titel zoals gedrukte kaarten, en er is geen thesaurus waarin de wijzigingen van historische toponiemen zijn opgenomen. Bovendien kloppen veel oude kaarten niet helemaal, doordat de werkelijkheid vervormd of in een

andere projectie is weergegeven.

In de afgelopen jaren zijn weliswaar veel kaartinventarissen en beschikbaar gesteld op de website van het Nationaal Archief. Tegelijk komt in de Beeldbank ook een groeiend aantal scans van kaarten beschikbaar met metadata ontleend aan deze inventarissen. Voor historisch onderzoek ontbreekt echter de context: de relatie met de besluitvorming in de archieven.

Georefereren

In september 2006 besloot het Nationaal Archief dat voor kaartmateriaal nieuwe vormen van toegankelijkheid ontwikkeld moeten worden, waarbij het georefereren de voorkeur geniet. Georefereren is het definiëren van kaarten met behulp van coördinaten. In 2007 werd

hiertoe samenwerking gezocht met medewerkers van het departement Informatica van de Universiteit Utrecht. Later werden de Universiteitsbibliotheek Leiden en de Koninklijke Nederlandse Akademie van Wetenschappen in het project betrokken. Die bibliotheek heeft in haar Bodel Nijenhuis Collectie een groot aantal manuscriptkaarten, die qua inhoud complementair zijn aan de collectie van het Nationaal Archief. Binnen de KNAW heeft de Virtual Knowledge Studio de expertise om met behulp van ICT nieuwe methoden en vormen van samenwerking binnen de geesteswetenschappen te bereiken.

DANS

Met behulp van de regeling Kleine Dataprojecten kon in het kader van

het DANS-project AlfaGeo een projectplan worden geschreven, waarbij de beste strategie moest worden gekozen om het kaartmateriaal toegankelijk te maken voor historisch onderzoek en om de relatie met de historische context te herstellen. Ook moest duidelijk worden hoe met potentiële partners zou kunnen worden samengewerkt.

De beoogde oplossing omvat een systeem waarbij op een Web 2.0-manier, met behulp van het publiek een extra laag metadata wordt aangebracht. In een testomgeving worden een aantal representatieve kaarten gegeorefereerd aangeboden. Onderzoekers kunnen zowel aan de kaart als aan objecten op die kaart informatie toevoegen. Daarbij kunnen tags worden toegekend die ook inzicht geven in de zogenaamde

long tail: de vele, weinig gevraagde kaarten die samen een belangrijke hoeveelheid informatie vertegenwoordigen die bij de huidige wijze van ontsluiting verloren dreigt te gaan omdat de gebruiker zijn kennis niet kan delen met anderen. Door de koppeling van kaarten en de metadata uit inventarissen aan nieuwe contexten kunnen relaties tussen de oorspronkelijke documenten en de kaarten hersteld worden.

CATCH

In september is een subsidieaanvraag bij het programma CATCH van NWO ingediend, met de Virtual Knowledge Studio als hoofdaanvrager en de andere partners als medeaanvragers. Het Centraal Bureau voor Genealogie ondersteunde de aanvraag. Als de subsidie wordt toegekend zal in de komende vier jaar *Manuscript map Annotation and Presentation System* (MAPS) worden ontwikkeld, dat een aantal tools gaat omvatten waarmee kaarten kunnen worden gegeorefereerd en geannoteerd en langs (semi-)automatische weg aan documenten gekoppeld. Met MAPS komt een populaire, maar weinig toegankelijk bronsoort binnen het bereik van een veel groter publiek.

Back to the future

Metamorfoze en het behoud van informatie in de 21e eeuw

Sinds 1997 is Metamorfoze het nationale programma voor conservering van het papieren erfgoed in Nederland. Dit samenwerkingsverband tussen de Koninklijke Bibliotheek en het Nationaal Archief is in de loop van dit jaar overgestapt van microfilm op digitale opslag.

Metamorfoze, dat wordt gefinancierd door het Ministerie van Onderwijs, Cultuur en Wetenschap, bindt strijd aan met de problemen waarvan oude boeken, kranten, tijdschriften en geschreven documenten te lijden hebben zoals verzuring en inkt- of kopervraat. Doel is het behouden van de informatie die aan papier werd toevertrouwd en die verloren dreigt te gaan.

Bijzondere boekedities, briefwisselingen van schrijvers en kunstenaars, historisch drukwerk, archieven van steden, instellingen en particulieren, het is allemaal informatie die ons veel kan vertellen over de (cultuur)geschiedenis van Nederland. Uit wetenschappelijk oogpunt is het dus van belang om deze informatie te behouden en meer toegankelijk te maken. Het behouden van de originele objecten zelf krijgt vorm in het zo lang en goed mogelijk bewaren van ten minste één exemplaar, de zogenaamde *last copy*. Het bewaren van de informatie kan bewerkstelligd worden door het overzetten van de inhoud op een ander medium. Sinds de start van Meta-



morfoze was microverfilming daarvoor een betrouwbare methode. De duurzame opslag van de nieuwste generatie hoogwaardige microfilms was eenvoudig en goedkoop.

Sinds die eerste jaren heeft digitalisering zich als conversiemethode sterk ontwikkeld. Aan de gebruikerskant heeft digitalisering veel voordelen ten opzichte van microverfilming. Teksten kunnen bijvoorbeeld doorzoekbaar worden gemaakt en gebruikers kunnen ze op elk gewenst moment en op afstand raadplegen. Maar ook uit oogpunt van conserve-

ring biedt digitalisering interessante nieuwe perspectieven; door de detaillering en kleur kan een digitaal beeld van hoge kwaliteit het origineel veel dichter benaderen dan een microfilm.

Metamorfoze heeft sinds 2001 ervaring met digitaliseringsprojecten, die in samenwerking met het Geheugen van Nederland worden uitgevoerd. Veel deelnemende erfgoedinstellingen beschikken niet meer over leesapparatuur voor microfilms of investeren daar in elk geval niet meer in. Daarom geven ze

de voorkeur aan digitale gebruikerskopieën in plaats van microvormen. Metamorfoze heeft in 2004 besloten ook hiervoor subsidie te verstrekken. De analoge master (op microfilm) bleef echter wel het duurzame opslagmedium.

In de loop van 2008 is Metamorfoze volledig overgestapt op digitalisering als conserveringsmethode: Preservation Imaging. Hiertoe zijn in 2007 de *Richtlijnen Preservation Imaging Metamorfoze* gepubliceerd. Deze Richtlijnen zijn gebaseerd op twee principes. Ten eerste moet alle informatie van het origineel ook te vinden zijn in het digitale surrogaat. Ten tweede moeten de Richtlijnen het mogelijk maken om kwaliteitscontroles uit te voeren met software en objectieve metingen, in plaats van uitsluitend door visuele inspectie zoals tot nu toe het geval was. Door de Richtlijnen is het mogelijk dit proces te automatiseren en efficiënter te maken. Bovendien wordt door het volgen van de Richtlijnen een optimale beeldkwaliteit gegarandeerd. Vanuit financiële of praktische overwegingen kan het echter nodig zijn

de hoge kwaliteitscriteria bij te stellen. Toepassing van de Richtlijnen maakt het dan mogelijk om te bepalen hoeveel kwaliteitsverlies acceptabel is.

Door deze nieuwe ontwikkelingen en mogelijkheden slaat Metamorfoze de richting in van massadigitalisering. Daarom wordt naar afstemming gestreefd met soortgelijke activiteiten en initiatieven, zowel in Nederland (Nederlands Erfgoed: Digitaal!, Digitalisering geesteswetenschappen, Libratory) als internationaal (The European Library, Europeana, Google Books).

Metamorfoze is door alle nieuwe techniek ook verantwoordelijk geworden voor de duurzame opslag van de digitale masters. Het e-Depot van de KB wordt momenteel geschikt gemaakt voor de verwerking en opslag van de miljoenen digitale masters die Metamorfoze zal gaan produceren.

Ten slotte zal Preservation Imaging gaan leiden tot een bredere beschikbaarstelling en betere toegankelijkheid van de geconserveerde informatie, onder andere via OCR en text-retrieval. Binnen het project wordt nu onderzocht in hoeverre het daarin een rol moet spelen. Ook is het de bedoeling om alle in afgelopen jaren gemaakte microfilms te digitaliseren. Maar hoe dan ook blijft Metamorfoze zich concentreren op de conservering van het Nederlandse papieren erfgoed, zodat dit ook voor de toekomst beschikbaar blijft.

.....



Nicis is opgebouwd uit vier delen: Onderzoek, Kenniscentrum, Academy en Trends & Monitoring. De onderzoekspoot van Nicis werkt vooral faciliterend. Hafkamp: 'Door universitaire onderzoekers en steden en soms ook bedrijven in kennisplatforms bij elkaar te brengen willen we onderzoeksprojecten opzetten die relevant en toepasbaar zijn voor de stedelijke praktijk en wetenschappelijk excellent en vernieuwend zijn.' Nicis heeft hiervoor twee onderzoeksprogramma's opgezet; het Stedelijk Innovatieprogramma (STIP) en 'Kennis voor Krachtige Steden', maar laat daarnaast ook kortlopend onderzoek uitvoeren.

Via de Nicis Academy worden professionals opgeleid. ‘We verzorgen diverse master classes en in samenwerking met de Nederlandse School voor Openbaar Bestuur is de masteropleiding “Strategisch Urban Studies” opgezet’, aldus de directeur. De opleiding combineert de laatste inzichten

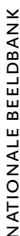
De afdeling Trends & Monitoring ten slotte brengt ontwikkelingen in kaart en werkt aan een kwantitatieve databank voor stedelijk onderzoek en trendanalyse. Hafkamp: 'Data van het CBS en uit

Daarnaast kan het LISS-panel gebruikt worden voor eigen dataverzameling. Ook dit is gratis voor onderzoekers aan universiteiten en wetenschappelijke instituten. Voorstellen daartoe mogen heel kort zijn, kunnen het hele jaar worden ingediend en worden binnen één tot drie maanden beoordeeld (zie de website voor de richtlijnen). Tot op heden is het merendeel van de ingediende voorstellen goedgekeurd, soms na enige aanpassingen. Voor 2009 en de jaren daarna is er nog ruimte voor diverse vragenlijsten

Omdat het LSS-paneel speciaal

De gegevens die Nicis in het onderzoek benut zijn vooral bestaande bronnen die via het CBS of de lokale overheden toegankelijk zijn. Hafkamp: 'Dit soort gegevens wordt binnen het onderzoeksconsortium beschikbaar gemaakt, zodat beleidsmakers en onderzoekers dezelfde informatie delen. Ook streven we ernaar om de data van universitaire onderzoeksgroepen te ontsluiten.' Nicis beraadt zich nog over de vraag hoe haar eigen gegevens beschikbaar kunnen komen voor externe onderzoekers. De komende tijd moet Nicis Institute verder internationaliseren volgens Hafkamp: 'We opereren in ons onderzoek en in de praktijk nog vaak te Hollands. Door over de grenzen te kijken kunnen we grote voordelen halen. Er zijn op Europese en mondiale schaal natuurlijk veel meer stedelijke gebieden die met dezelfde grootstedelijke vraagstukken stoeien.'

Hill, J. L., & Hill, J. A. (1986). The effects of social support on the health of older adults.



T... .. N... .. D...

© www.dariah.eu

Samenspel dialecten en moderne techniek

Op zaterdag 4 oktober 2008 verzamelden zich een veertigtal dialectgeïnteresseerden uit Vlaanderen en Nederland in het hypermodern uitgeruste auditorium van Erfgoedhuis Oosterzele in het Oost-Vlaamse dorpje Moortsele. Aanleiding was een studiemiddag, georganiseerd door de koepelorganisatie van dialectverenigingen in Vlaanderen, de Vereniging voor Geschiedenis en Informatica en FARO (het Vlaams steunpunt voor cultureel erfgoed). Onderwerp van gesprek was de cross-over tussen toegepaste technologie, taalvariatieonderzoek en erfgoedpraktijk.

Jacques van Keymeulen van de Universiteit Gent en voorzitter van de koepelorganisatie van dialectverenigingen in Vlaanderen, opende met een warm pleidooi voor een digitale volkslinguïstiek. Daarbij zouden de bestaande lokale lexicale dialectverzamelingen ook in een overkoepelende digitale dataverzameling ingepast moeten worden, om deze vervolgens via woordenboeken en taalatlassen voor het grote publiek te ontsluiten.

Aldus zijn moderne technieken te gebruiken om traditionele dialecten te bewaren. Moderne technieken leiden echter zelf ook tot nieuwe groepstalen, zoals Reinhild Vandekerckhove van de Universiteit Antwerpen aantoonde. Uit haar onderzoek naar chat- en sms-taal in Vlaanderen blijkt de invloed van nieuwe communicatietechnologie op het taalgebruik van jonge-



GEMEENTE OOSTERZELE
Erfgoedhuis Oosterzele

ren onmiskenbaar te zijn. Maar in tegenstelling tot wat vaak gedacht wordt, is de geografische actieradius waarbinnen die technologie communicatief ingezet wordt eerder beperkt en bevat het daarbij ontwikkelde taalgebruik sterk regionaal bepaalde kenmerken.

Rob Belemans, stafmedewerker van FARO, stond stil bij de vraag

hoe met talig erfgoed in de web 2.0-gemeenschappen omgegaan wordt. Hij liet zien dat degenen die hun lokale of regionale dialect willen conserveren en in leven houden veel kunnen opsteken van de methodieken en tools die jongeren gebruiken in hun digitale wiki-platforms, zoals de Urban DICTionary. (Rob Belemans)

@ www.faronet.be/blogs/bram-wiercx/
video-opnames-studienamiddag-taalvariatie-technologie

IMPACT conferentie: OCR in massadigitalisering

Op 6 en 7 april 2009 organiseert het Europese project IMPACT een internationale conferentie in de Koninklijke Bibliotheek over optische tekenherkenning (OCR) in projecten voor massadigitalisering.

In het IMPACT-project worden innovatieve technieken ontwikkeld om de bestaande optische tekenherkenning te verbeteren. De tweedaag-

se conferentie biedt onderzoekers en leveranciers de mogelijkheid om nieuwe inzichten op OCR-gebied uit te wisselen. Daarnaast zullen enkele eerste resultaten uit het eerste jaar van IMPACT gepresenteerd worden. Op 6 april staan nieuwe ontwikkelingen in de OCR-technologie centraal, zoals 'Collaborative Correction', een uitgekiend gereedschap waarmee vrijwilligers zonder speciale kennis online kunnen meehelpen aan de correctie van OCR-resultaten. Dit is een mogelijke stap vooruit voor toekomstige grootschalige digitaliseringsprojecten. Op 7 april worden huidige en toekomstige uitdagingen voor OCR-technologie besproken, zoals taalkundige problemen bij de digitalisering van historisch tekstmateriaal. Op beide dagen zijn er zowel sprekers uit het IMPACT-consortium als van buiten het project. Tot 1 januari 2009 kan met korting worden geregistreerd.

@ www.impact-project.eu/about-the-project/news/ic2009/

Sinds kort beschikbaar

Het overzicht toont een aantal databestanden die recent voor onderzoekers beschikbaar zijn gekomen bij CBS en DANS. Een volledig overzicht van de CBS-bestanden is te vinden op www.cbs.nl/microdata. De bij DANS beschikbare databestanden komen

van diverse andere onderzoeksinstellingen. Deze kunnen kosteloos worden gedownload vanuit DANS EASY: <http://easy.dans.knaw.nl>. Via DANS kunnen ook de beveligde microdata van het CBS kosteloos geleverd worden: www.dans.knaw.nl/nl/data/cbs/overzicht/

Centraal Bureau voor de Statistiek	Periode
Gemeentelijke basisadministratie personen (GBA)	1995-2008/07
Inkomens panel onderzoek (IPO)	2005
Wet inkomen naar arbeidsvermogen (WIA)	2006-2008/06
Onderwijs cohorten (VOCL) schoolelementen	1999
Voortgezet onderwijs (VO) examens	2006/2007
Middelbaar beroepsonderwijs (MBO) examens	2005-2007
Middelbaar beroepsonderwijs (MBO) inschrijvingen	2007/2008
Voortgezet algemeen volwassenenonderwijs (VAVO) diploma's	2005-2007
Volwasseneneducatie (VE)	2005-2008
International sourcing (INTSO)	2007
DANS EASY	
<i>Archeologie</i>	
Nistelrode – Zwarte Molen, Bronstijdvindplaats (Archeologisch onderzoek Leiden BV)	2008
Honderd archeologische projecten 2007 en 2008 (Archaeological Research and Consultancy)	2007/2008
<i>Geschiedenis</i>	
VOC opvarenden: personeelsadministratie VOC (Universiteit Leiden – Nationaal Archief)	18e eeuw
<i>Sociale wetenschappen</i>	
Tevredenheid van slachtoffers van misdrijven met de slachtofferhulp (Justitie – WODC)	2008
Niet-gebruik inkomensafhankelijke regelingen (Sociaal en Cultureel Planbureau)	2005
Veiligheidsmonitor Rijk (CBS microbestand)	2008

11 december uitreiking VGI Innovatieprijs

De VGI Innovatieprijs wordt voor de zesde keer uitgereikt voor een ICT-toepassing die op een vernieuwende manier met geschiedenis omgaat. Uit achttien inzendingen heeft de jury drie kandidaten genomineerd die tijdens de VGI-studiemiddag van 11 december een presentatie geven over hun applicatie, waarna de jury de winnaar bekend maakt.

De genomineerden zijn: Brandgrens van het Gemeentearchief Rotterdam, de historische kaartensite WatWasWaar en de multimediale site www.expo58.eu.

Brandgrens www.gemeentearchief.rotterdam.nl/brandgrens

Een rustig vormgegeven website met historische informatie over Rotterdam in de Tweede Wereldoorlog, aan de hand van de brandgrens van de bombardementen die de stad getekend heeft. De uitgebreide teksten bevatten zeer veel informatie en vormen een overtuigende illustratie van het feit dat een site met lange teksten, cijfers en details kan werken. Innovatief is de 'brandgrensnavigator', die uitnodigt tot wandelingen en een bezoek aan het archief.

Watwaswaar www.watwaswaar.nl
WatWasWaar presenteert fraai kaartmateriaal van archieven en andere instellingen via geografische coördinaten en brengt daar weer dwarsverbanden tussen aan. Een sterk punt is dat materiaal van zoveel verschillende instellingen op één plek via één interface doorzoekbaar is. De site is zeer intuïtief te gebruiken, wat een prestatie mag heten want WebGis-interfaces voor dergelijke projecten zijn vaak gebruiksonvriendelijk. De

cartografische informatie, luchtfoto's en bevolkingstellingen die op de site zijn samengebracht, zullen een groot publiek aanspreken.

Expo '58: back to the future www.expo58.eu

Materiaal verzameld rondom de wereldtentoonstelling Expo 1958 maakt een mooie en gevarieerde site. De persoonlijke herinneringen aan Expobezoek brengen op velerlei manieren de grote geschiedenis eind jaren vijftig dichterbij. De combinatie van verhalen die in dit samenwerkingsproject zijn verzameld, herinneringen en commentaren van bezoekers en allerlei multimedia zorgt voor een levendig geheel.

Tijdens de VGI-studiemiddag gaan twee sprekers in op de vraag of de groeiende belangstelling voor geschiedenis, ooggetuigenverslagen en persoonlijke verhalen ook betekent dat er op dit gebied wetenschappelijke vooruitgang geboekt wordt. De VGI-studiemiddag vindt plaats op 11 december 2008 van 14.00 tot 17.00 uur bij het Internationaal Instituut voor Sociale Geschiedenis aan de Cruquiusweg 31 in Amsterdam. Aanmelden kan bij Michelle van den Berk: berk@kitlv.nl. (LS)

@ www.vgi-online.nl

Datatour langs collegezalen

DANS en het Centrum voor Beleidsstatistiek van het Centraal Bureau voor de Statistiek (CBS) organiseren het lopende collegejaar een tour langs de collegezalen in Nederland waar sociale wetenschappen worden gedoceerd. Zo willen beide datacentra een bijdrage leveren aan de kennis van studenten over het archiveren, vinden en hergebruiken van data.

Medewerkers van beide instituten komen in het kader van die tour langs om een college te geven voor (mas-

ter)studenten over de mogelijkheden bij het gebruik van reeds beschikbare data. Ze behandelen vragen als: welke data zijn waar te vinden, waarom moeten data gearchiveerd worden, welke informatie biedt StatLine van het CBS en hoe moet je dat gebruiken, en welke data heeft het CBS nog meer? Opleidingen die interesse hebben in een dergelijk college kunnen contact opnemen met Marion Wittenberg bij DANS. marion.wittenberg@dans.knaw.nl.

COLOFON

e-data@research is het kwartaalblad in Nederland over data en onderzoek in de alfa- en gammawetenschappen. Het verschijnt onder auspiciën van DANS, het Huygensinstituut, het Internationaal Instituut voor Sociale Geschiedenis, het Centraal Bureau voor de Statistiek, de Koninklijke Bibliotheek en de Vereniging voor Geschiedenis en Informatica. Toezending kosteloos aan relaties van de stakeholders en op verzoek aan studenten in de alfa- en gammarichtingen. Oplage: 7500. e-data@research is online te raadplegen op www.edata.nl

Uitgever: Stichting Uitgeverij *e-data@research*, Postbus 93067, 2509 AB Den Haag

Redactieadres: Postbus 93067, 2509 AB Den Haag; t (070)3494450 f (070)3494451 e edata@dans.knaw.nl

Redactie: Peter Boot, Ivo Gorissen, Martijn de Groot (hoofd/eindredacteur), Inge Angevaare, Jetske van der Schaaf, Luuk Schreven

Aan dit nummer werkten mee: Louisa Balk, Rob Belemans, Karina van Dalen-Oskam, Charles van den Heuvel, Jeffrey van der Hoeven,

Frank Houtman, Edwin Klijn, Yola de Lusenet, Nino Mushkudiani, Maria Patijn, Annette Scherpenzeel, Dennis Schouten, Laurents Sesink, Marion Wittenberg, Douwe Zeldenrust

Redactiesecretariaat: Lucas Pasteuning, Jetske van der Schaaf

Vormgeving: Ellen Bouma

Productie: Uitgeverij Aksant, Amsterdam

Druk: PlantijnCasparie, Almere

ISSN: 1872-0374

Mijn virtuele ik

De grootste hype rond Second Life lijkt voorbij, maar de interessantste vragen wachten nog op antwoord. Wat zoeken mensen in Second Life of andere virtuele werelden? Wat doet het met ze om zich in een fictieve gestalte in zo'n wereld te begeven? Wat bepaalt de eigenschappen die ze kiezen voor hun 'avatar', hun representant in de virtuele wereld? Ilja Leonard Pfeijffer, die in *Second Life* verslag deed van zijn avonturen in de virtuele wereld, koos een vrouwelijke avatar '(...) omdat ik geboren wilde worden als iemand die ik in geen enkel opzicht was: (...) Mijn avatar is mooier dan ik ben, en bouwvallers fluiten haar na als zij voorbijloopt.'

Het contrast ten opzichte van de alledaagse werkelijkheid, en erotische aantrekkelijkheid, zijn factoren die de keuze voor een avatar wel vaker beïnvloeden. Een mooie gelegenheid om dat te constateren biedt de tentoonstelling *Alter Ego – mijn avatar en ik*, die gehouden wordt in het Museum voor Communicatie in Den Haag (t/m 4 januari). De tentoonstelling toont vijftien portretten van mensen met hun avatars, soms uit Second Life, soms uit games als World of Warcraft. In het bijbehorende boek, *Alter Ego; Avatars and their Creators* van Robbie Cooper, wordt een veel groter aantal paren getoond.

Een goed voorbeeld van hoe het gewone leven het virtuele leven beïnvloedt zijn de twee afbeeldingen van Mi-Jin Kang en haar vriendinnen: eenmaal als gewone Zuid-Koreaanse vrouwen en eenmaal als groepje do-



delijke en flitsende geklede krijgsvrouwen. Dat is niet wat men in Zuid-Korea van vrouwen verwacht, legt ze uit, maar het geeft een extatisch gevoel van macht. Tegelijk biedt het spel ook mogelijkheden om mannen te ontmoeten zonder lastig te kunnen worden gevallen.



Een veel schrijnender contrast bieden de foto's van Jason Rowe en zijn avatar. In het gewone leven is Jason meervoudig gehandicapt en vrijwel machteloos; in het virtuele leven is hij breed en zwaar gepantserd en kan hij op motorfietsen rijden en met monsters vechten. Maar rondhangen met vrienden is ook een attractie van de virtuele wereld. In het echte leven voelen mensen zich soms ongemakkelijk als ze Jason zien, maar online worden mensen niet afgeleid door je uiterlijk en leren ze je persoonlijkheid kennen.

De spanning tussen trouw aan de werkelijkheid en compensatie voor de tekortkomingen van de werkelijk-

heid kan heel verrassend uitpakken. In de sprookjeswerelden kunnen puberjongens elfjes zijn of machtige wrekers van ongerechtigheid. Maar sommige avatars zijn ook net als hun makers – alleen misschien iets jonger of slanker. In de gesprekjes die in *Alter Ego* zijn opgenomen geven mensen heel verschillende beelden van hun relatie met hun avatar. 'Een beetje van mezelf en een beetje van de persoon die ik zou willen zijn,' zegt de één, voor de ander is de avatar eerder een vriendin om mee te gaan stappen, weer een ander beschrijft de avatar als haar geweten. Pfeijffer vertelt van een ontmoeting met een vrouw die in Second Life, net als in het echte leven, in een rolstoel zit. Aanvankelijk had ze een avatar waarmee ze kon dansen en rennen, maar dat gaf haar toch het idee dat ze in Second Life niet zichzelf was.

'Identiteit in virtuele werelden' was onlangs ook het onderwerp van een symposium van de KNAW, waar Pfeijffer één van de sprekers was. Het onderwerp werd daar bovendien belicht uit juridisch en filosofisch gezichtspunt. De spanning tussen je werkelijke en virtuele verschijningsvormen kan aanleiding zijn voor diepgaande vragen over wat identiteit eigenlijk is. En zo zullen de virtuele werelden nog wel meer wetenschappelijke vragen oproepen, voor de sociale wetenschappen maar zeker ook voor de geesteswetenschappen. (PB)

www.muscom.nl/tentoonstellingen/alter_ego/
www2.eur.nl/fw/hyper/Download/2009-09_KNAW_symposium_strip.jpg

Gelezen



Willem Witsen, Tachtiger in brief en beeld, een digitale collectie van de Koninklijke Bibliotheek, de Bibliotheek van de Universiteit van Amsterdam, het Rijksmuseum voor Kunsthistorische Documentatie, de Stichting Willem Witsen, het Huygens Instituut en de DBNL.
Het eind van de negentiende eeuw is een spannende tijd op het gebied van cultuur in Nederland. De nieuwe literaire stroming 'Beweging van Tachtig', met Willem Kloos als middelpunt, bracht leven in de brouwerij. De 'Tachtigers' zetten zich af tegen de periode van de romantiek en propageerden de uiting van individueel gevoelde emoties. De beweging sloeg ook aan bij kunstenaars en musici. Onder hen bevond zich de beeldend kunstenaar Willem Arnold Witsen (1860-1923). Zijn nalatenschap omvat, behalve zijn oeuvre van schilderijen, aquarellen, tekeningen en foto's, ook brieven met tijdgenoten en persoonlijke dagboeknotities.
De brieven en persoonlijke documenten van Witsen worden bewaard bij verschillende openbare instellingen en particularieren. Door het wijdver-

takte netwerk van Witsen met kunstenaars, literatoren, musici, acteurs en andere personen vormt dit materiaal een waardevolle bron voor uiteenlopend onderzoek. De documenten zijn dan ook van grote waarde voor iedereen die onderzoek doet naar de periode van ca. 1880 tot 1923. Zie ook *e-data&research* 2-1.

www.geheugenvannederland.nl

Lusenet, Y. de, *Geven en nemen, archiefinstellingen en het sociale web* Den Haag, Taskforce Archieven, 2008. ISBN 978 90 78763 04 8



De virtuele omgeving waarin archiefinstellingen hun digitale collecties aanbieden evolueert razendsnel. Internetgebruikers raken er steeds meer aan gewend om actief hun eigen omgeving te creëren, bijvoorbeeld door bij te dragen aan sites en eigen inhoud op het web te plaatsen. Wat

zijn de effecten van deze ontwikkelingen voor archiefinstellingen? Welke kansen, maar ook welke vragen roepen deze ontwikkelingen op. De publicatie *Geven en nemen* verkent deze vragen en wil lezers vooral aansporen nader te verkennen wat er voor spannings gaande is op het web. Alleen experimenten en evaluaties door de archiefinstellingen zelf kunnen duidelijk maken wat Web 2.0 voor de sector betekent. www.taskforce-archieven.nl

International study on the impact of copyright law on digital preservation, Library of Congress National Digital Infrastructure and Preservation Program, USA, 2008
Partners in Australië, Nederland, het Verenigd Koninkrijk en de Verenigde Staten hebben een rapport opgesteld dat een kritische blik werpt op bestaande wetten op het gebied van auteursrecht en daaraan gerelateerde wetgeving, en op de invloed van die wetten op de digitale conservering van auteursrechtelijk beschermde werken. In dit rapport worden voorstellen onder de loep genomen voor het veranderen van de huidige wetgeving. Ook kijkt het rapport naar wat er momenteel gedaan wordt om de problemen die copyrightwetgeving voor digitale conservering creëert op een andere wijze dan met wetgeving op te lossen. De volgende partners hebben een bijdrage aan het rapport geleverd: Library of Congress, National Digital Information Infrastructure

Column

Charles van den Heuvel

The In-Between Machine

Het World Wide Web wordt een steeds belangrijkere bron voor onderzoek. Tim Berners-Lee beschreef in Weaving the Web (1999) zijn droom dat het web niet alleen een ruimte zou zijn om informatie te zoeken, maar ook om kennis aan toe te voegen. Uit het succes van FaceBook, MyTube, Flickr, maar ook van wetenschappelijke collaboraties, blijkt dat dit nu ook gebeurt. Tegelijkertijd verdwijnen er dagelijks miljoenen pagina's. In 1996 besloot Brewster Kahle het Web te archiveren. Vele nationale bibliotheken, waaronder de KB, volgden zijn voorbeeld met het archiveren van websites met nationale domeinnamen. Daarmee zijn deze webarchieven echter nog niet toegankelijk voor gebruik. Het blijft een vreemde gedachte dat je, om webarchieven van de Library of Congress te raadplegen, naar Washington moet afreizen om ze op monitors binnen de muren van de bibliotheek te lezen. Behalve problemen in verband met auteursrecht en nationaal collectiebeleid, zijn er andere redenen waarom webarchieven niet optimaal voor onderzoek kunnen worden benut. Allereerst zijn bestaande interfaces gebrekkig. Verder zijn de tools om websites door middel van annotaties in context te plaatsen, meestal voor individueel gebruik. Hierdoor is het voor toekomstige onderzoekers niet duidelijk of, en hoe, deze websites eerder zijn bestudeerd.

Wie nu webarchieven wil bekijken in Brewster Kahle's Internet Archive moet een url intypen in de Wayback Machine om vervolgens een lange rij van data te doorlopen waarop de sites zijn gearchiveerd. Dat is handig voor degene die meer te weten wil komen over 9/11 of een overzicht wil hebben van een betrekkelijk korte periode van intensief gebruik van websites, zoals die van de politieke verkiezingen in de VS. Echter voor de meeste vormen van historisch onderzoek is een dergelijke opsomming van data volstrekt ongeschikt. Daarom hebben we allereerst behoefte aan een betere interface. De ontwikkeling hiervan is geen eenvoudige zaak. Kies je voor een generieke, of juist voor een op specifieke disciplines afgestemde 'tailor-made' interface? Deze vraag lijkt onoplosbaar; zeker wanneer je bedenkt dat gebruikers kunnen wisselen en tijdens zoekopdrachten veranderen van strategie. Misschien moeten we ons daarom minder op de inhoud en meer op de stijlen van onderzoek richten, die onderzoekers gemeen hebben.

Onderzoekers moeten niet alleen relevante informatie vinden, maar deze ook met aantekeningen kunnen verrijken. Nu bestaan er diverse annotatietools voor particulier gebruik. Kennisverrijking van webarchieven maakt echter ook deel uit van ons collectief geheugen. Het zou zelfs een selectiecriterium van behoud kunnen zijn. Welk exemplaar van Newton's Principia zou jij behouden, die met of zonder de annotaties van Einstein? Samen met een aantal nationale en internationale instellingen werkt de Virtual Knowledge Studio for the Humanities and Social Sciences (KNAW) aan een aanvraag voor een Europese subsidie voor de ontwikkeling van de 'In-between Machine', waarbij de interface op onderzoek wordt afgestemd en webarchieven en annotaties als collectief geheugen worden bewaard.

Charles van den Heuvel is senior researcher bij de Virtual Knowledge Studio

and Preservation Program (NDIIPP), Verenigde Staten; Joint Information Systems Committee (JISC), Verenigd Koninkrijk; Open Access to Knowledge (OAK) Law Project, QUT Law Faculty, Australië en SURFfoundation, Nederland. www.loc.gov/index.html

Nieuwsbrief WatWasWaar

Electronische nieuwsbrief van WatWasWaar over nieuwe diensten en toegevoegde content. WatWasWaar geeft toegang tot een groot aantal, soms unieke collecties, van Nederlandse erfgoedinstellingen. Het materiaal is zeer divers, overal anders en wordt continu aangevuld.
<http://watwaswaar.nl>

Data Archiving and Networked Services (DANS), Jaarverslag 2007, Den Haag, september 2008

Verslag van de activiteiten van DANS in 2007. Stand van zaken van diverse nationale en internationale projecten, archiefaanwinsten, en data inventarisaties. In print en pdf.
www.dans.knaw.nl



WIEBE KESTRA