

Pagina 3 • CBS heeft iPhone app •

Het CBS introduceerde deze zomer een toepassing om op de iPhone direct de cijfers van de databank StatLine te kunnen raadplegen.

Pagina 4 • GIS & Geschiedenis •

Wat was in de negentiende eeuw de invloed van infrastructuur op de levensstandaard? Geografische Informatiesystemen kunnen het antwoord op zulke vragen dichterbij brengen.



WIEBE KIESTRA

Pagina 5 • Herbert Van de Sompel •

Tijdreizen op het internet, open annotatiesystemen, elektronisch bronnenbeheer, open URL, een aanbevelingssysteem voor wetenschappelijke artikelen. De Vlaamse informatiewetenschapper Herbert Van de Sompel speelt op veel borden tegelijk.

Pagina 6 • Akademy •

In de rubriek Focus komt het kennisnetwerk van de Fryske Akademy over het voetlicht.

Pagina 8 • Tijdcapsule • Een initiatief om de aandacht te vestigen op de noodzaak van duurzame toegang tot data oogstte niet het beoogde effect.

EN VERDER

Agenda	2
Nieuws	3
Achtergrond	4,6
Focus	6
Sinds kort beschikbaar	7
Column	8
Gelezen	8

PARSE.Insight presenteert eindrapport en Road Map

'Andere organisatie en cultuur nodig'

Om duurzame toegankelijkheid van wetenschappelijke informatie te waarborgen is het niet genoeg om alleen technische faciliteiten op te schalen. Veranderingen in organisatie en cultuur zijn daarvoor essentieel. Dat staat in de Road Map die bij het eindrapport van het Europese project PARSE.Insight is gevoegd om aan te geven hoe een geïntegreerde Europese infrastructuur moet worden bereikt.

PARSE.Insight (Permanent Access to the records of Science in Europe) leverde op 25 juni zijn definitieve rapportage op over hoe wetenschappers, datamanagers, bibliotheken en uitgevers aankijken tegen de zorg voor digitale wetenschappelijke informatie. Omdat Europa groot is en de verschillen per discipline ook aanzienlijk, bleek het een enorme klus om goede informatie te verzamelen. Uiteindelijk werden tijdens het afsluitende symposium in Parijs de resultaten bekendgemaakt in de vorm van een serie rapporten. Daarin wordt de huidige stand van zaken in kaart gebracht en wordt ook geschetst hoe de wereld er uit zou zien als alle spelers hun zaakjes op orde zouden hebben. Er blijkt nog veel werk aan de winkel voor onderzoeksinstellingen, archieven en bibliotheken, uitgevers, financiers van onderzoek en beleidsmakers. Tien belangrijke inzichten uit het rapport zijn:

- Duurzaamheid en (open) access worden vaak door elkaar gehaald maar hebben nauwelijks met elkaar van doen. De enige relatie is dat zonder duurzaamheid de toegang uiteindelijk wegvalt. Maar open access gaat vooral over de vraag of de producent of de gebruiker voor de

toegang betaalt, en die vraag staat los van de noodzaak om duurzaam te bewaren.

- Er is in Europa geen eenduidig beleid voor het bewaren van data.

- Er is nog geen certificeringsproces voor duurzame bewaarplaatsen.

- Het is onduidelijk wie wat moet doen om duurzaamheid te regelen. Voor publicaties is vaak wel het nodige afgesproken, maar dat geldt niet voor ruwe onderzoeksdata.

- Er zijn grote verschillen tussen Noordwest-Europa en andere delen van de EU. Engeland, Nederland en Duitsland lopen voorop in hun onderzoek naar en realisatie van digitale duurzaamheid.

- Er is nog weinig afstemming met initiatieven voor duurzame data in andere delen van de wereld. In de VS, Azië en Australië staat duurzaamheid hoog op de agenda.

- De dekking van digitale archieven voor dataopslag is slecht. Sommige wetenschappelijke disciplines hebben nog helemaal geen plek voor het bewaren van hun data.

- Onderzoekers zijn terughoudend om hun data te delen met anderen. Ze zijn bang voor verkeerde interpretatie en schendingen van privacy- en auteursrecht.

- De hoeveelheid data groeit explosief, maar er is nog geen beleid om uit te maken wat wel en niet bewaard moet worden. Ook ontbreekt in de meeste disciplines standaardisering van bestandsformaten.

- Onderzoekfinanciers hebben de macht om onderzoekers ertoe te bewegen hun data goed te organiseren en te delen, maar ze gebruiken die macht nog nauwelijks. (Jeffrey van der Hoeven)

© www.parse-insight.eu/publications.php#d2-2.

Erfgoed van de Oorlog



Op donderdag 16 en vrijdag 17 september 2010 vindt in het Koninklijk Instituut voor de Tropen in Amsterdam de slotconferentie plaats van het programma Erfgoed van de Oorlog. Op de sluitingsdag is het precies 67 jaar geleden dat deze tekening van 'schoppers al d arbeid' werd gemaakt in het Japanse gevangenenkamp Muroan. Het was nummer 9 uit een 'verzameling pennekrabbels speciaal getekend voor kap. Meys te Muroan' door een tekenaar met de initialen A.M. Collectie Nederlands Instituut voor Oorlogsdocumentatie.

Erfgoed van de Oorlog was een subsidieprogramma om belangrijk materiaal uit de Tweede Wereldoorlog en de strijd in Nederlands Indië te behouden en het gebruik ervan te stimuleren. Onderdeel van het programma was het project Getuigen Verhalen, waarin tal van persoonlijke herinneringen aan de bewuste periode werden vastgelegd en bij het elektronisch archief EASY van DANS (<http://easy.dans.knaw.nl>) gearchiveerd.

Deze maand wordt ook de website www.oorloginblik.nl geopend, voortgekomen uit hetzelfde programma, die duizenden titels aan films of geluidsmateriaal hoorbaar en zichtbaar maakt. (MdG)

Dataprijs voor Historische Steekproef



JOHANNES ABELING

Prof. Kees Mandemakers is de winnaar geworden van de DANS Dataprijs 2010. De historicus kreeg de prijs voor zijn werk aan de Historische Steekproef Nederland (HSN), een steekproef van bijna tachtig duizend personen, geboren in de negentiende en begin twintigste

eeuw. Mandemakers (links) ontving de versierselen behorend bij de prijs op 18 juni in Amsterdam uit handen van KNAW-president prof. Robbert Dijkgraaf (rechts op de foto). Hij stak drie andere genomineerden de loef af: prof. Pearl Dykstra met de Netherlands Kinship Panel Study, prof. Marcel Das met het LISS panel en prof. Kees Aarts met het Nationaal Kiezersonderzoek. Juryvoorzitter prof. Pieter Hooimeijer was vol lof over alle vier de genomineerden, die dan ook in eerste instantie naar het oordeel van de jury gelijk waren geëindigd. Toch kreeg ten slotte de HSN het beslissende plusje, onder meer op grond van de indrukwekkende

lijst met tweehonderd publicaties die het project had voortgebracht. Mandemakers vergat niet om in zijn dankwoord een flink aantal collega's te betrekken, waaronder mede-initiator van de steekproef Jaap Dronkers. DANS-directeur en gastheer Peter Doorn greep de gelegenheid aan om ervoor te pleiten dat aan de verantwoordelijken voor een duurzaam beschikbaar gemaakte dataset evenveel eer wordt toegezwaard als aan degenen wier naam nu boven de erop gebaseerde publicaties prijkt. (MdG)

© www.iisg.nl/hsn

www.dans.knaw.nl>dataprijs

Archieven kiezen voor landelijke digitale ontsluiting

Op 15 juni hebben vertegenwoordigers van een groot aantal overheidsarchieven een overeenkomst ondertekend voor het ontwikkelen van gemeenschappelijke e-depot voorzieningen en diensten. Archiven op landelijk, regionaal en lokaal bestuursniveau spreken er de intentie in uit om met elkaar te zorgen voor integrale toegang tot alle digitale overheidsarchieven. De voorzieningen zullen gebouwd worden op basis van de faciliteiten die al ontwikkeld zijn door e-depotpioniers Nationaal Archief, Gemeentearchief Rotterdam en Stadsarchief Amsterdam.

De betrokken organisaties, verenigd in de Archief Coalitie Digitale Duurzaamheid (ACDD), leggen in hun persbericht de nadruk op de verantwoording die de overheid volgens de Archiefwet moet afleggen over haar handelen, maar ook het belang van de overeenkomst voor de geschiedschrijving is groot. Het plan voorkomt dat archieven hun eigen lokale oplossingen gaan ontwik-

kelen met alle toegangsproblemen van dien.

Directeur Martin Berendse van het Nationaal Archief vindt het gezamenlijke e-depot 'essentieel' voor de langetermijntoegang. 'De Raad voor Cultuur, de Algemene Rekenkamer, de Erfgoedinspectie, ze hameren er al jaren op dat geheugenverlies voor de overheid op de loer ligt,' aldus Berendse. 'Neem de Arbeidsvoorziening Nederland die in 2001 is opgeheven. De Algemene Rekenkamer memoreerde onlangs nog dat het afronden van de jaarrekening een groot probleem bleek te zijn voor de accountants. Het digitale archief dreigde grotendeels onleesbaar te worden. Wat werd aangetroffen, was een wirwar aan complexe databases, softwaresystemen en bestandsformaten. Uiteindelijk heeft het miljoenen gekost om dit ene archief in langdurig toegankelijke en geordende staat te krijgen. Dat soort toestanden moeten we in de toekomst zien te voorkomen.'



Jantje Steenhuis heeft als eerste getekend, namens haar eigen Gemeentearchief Rotterdam en branchevereniging BRAIN.

Het woord is nu aan de bewindlieden van de ministeries van Onderwijs, Cultuur en Wetenschap en Binnenlandse Zaken, de Vereniging Nederlandse Gemeenten, het Interprovinciaal Overleg en de Unie van Waterschappen. Zij moeten zorgen

voor de twintig miljoen euro die nodig zijn om het plan te realiseren. Volgens de ondertekenaars is dit maar een vijfde van wat het zou kosten als de organisaties ieder voor zich e-depotvoorzieningen zouden ontwikkelen. (IA)

www.archief.nl/sites/archief.nl/files/docs/rapport_e-Depot.pdf

Virtual Knowledge Studio wordt deel van e-bundeling KNAW

Op 1 januari 2011 gaat een onderzoeksgroep van start die is samengesteld uit medewerkers van de huidige Virtual Knowledge Studio (VKS), onderzoekers van andere instituten binnen de KNAW die met e-science bezig zijn, en het nog te lanceren programma van de Akademie voor de computational humanities. In dat programma staat computermatige patroonherkenning centraal. Met dit besluit, door de KNAW vlak

voor de zomer genomen, komt een eind aan de discussie over de toekomst van de VKS. De studio was in 2005 (officieel 2006) van start gegaan om ruimte te geven aan verkenningen en experimenten rond nieuwe onderzoeksmethoden en analyses, mogelijk gemaakt door het gebruik van elektronica. De nieuwe onderzoeksgroep krijgt een soort matrixstructuur, aldus VKS-directeur Paul Wouters. 'Er gaan mensen werken



VKS-directeur Paul Wouters.

van allerlei onderzoeksinstituten die dus ook hun contacten en de aanvoer van kennis en ideeën uit die richting zullen onderhouden. Maar er is ook een gezamenlijkheid onder leiding van een programmaleider, benoemd door de KNAW.' De medewerkers van de onderzoeksgroep worden op één plaats ondergebracht. Welke dat is wordt nog onderzocht.

Door het samenbrengen van onderzoekers uit vele richtingen komt

er ruimte voor meer onderlinge uitwisseling en inspiratie. Wouters ziet dat als een voordeel dat ruimschoots opweegt tegen het mogelijke bezwaar dat er 'iets wordt ingeleverd van de intellectuele vrijheid' die zijn studio te bieden had. De nauwere binding met instituten zal naast wetenschappelijke voeding ook een sterkere gerichtheid op toepassingen meebrengen, zo redeneert de VKS-directeur, die eraan toevoegt 'optimistisch' te zijn over de mogelijkheden van de nieuwe onderzoeksgroep. Voor Wouters, die zelf een benoeming in Leiden heeft aanvaard als hoogleraar Scientometrie (meten van wetenschappelijke productiviteit) is ook belangrijk dat de VKS alsnog zal worden geëvalueerd. 'Die oude afspraak dreigde eerst achter het besluit tot opheffing van de VKS te verdwijnen maar wordt nu alsnog uitgevoerd'. (MdG)

CBS komt met iPhone-applicatie

Sinds begin juni hoeven iPhone-gebruikers niet meer naar cbs.nl te surfen om cijfers op te zoeken in de CBS-databank StatLine. Ze kunnen nu CBS-StatLine downloaden zodat ze de ruim acht miljard cijfers uit de databank altijd binnen drie kliks op hun scherm hebben.

De StatLine app bestaat uit vijf onderdelen. De grafieken in het openingsscherm laten snel een twaalfal economische kernindicatoren zien waaronder de inflatie, de economische groei en de werkloosheid. Het menu *Recent* toont de tabellen die het laatst zijn bijgewerkt en het menu *Thema's* biedt cijfers via de themaindeling van het gewone StatLine. In *Favorieten* kunnen gebruikers tabellen opnemen die ze vaker willen raad-

plegen. *Mijn locatie* ten slotte laat alle tabellen zien over de gemeente waar de gebruiker zich bevindt.

Het CBS is de eerste Nederlandse overheidsinstantie met een eigen iPhone-applicatie. En met succes: de eerste week stond CBS-StatLine voortdurend in de top tien van meest gedownloadde gratis apps en werd het ruim vijftienduizend keer gedownload.

Om jonge doelgroepen te bereiken en zijn imago te verbeteren profileert het CBS zich ook via andere nieuwe media. Zo bracht het in juli zijn eerste e-book uit en is het sinds een klein jaar actief op Twitter (twitter.com/statistiekcb) en YouTube (youtube.com/statistiekcb) (Jorien Apperloo)

Gedragscode

Onderzoekers moeten hun originele gegevens bewaren en beschikbaar houden voor collega's, volgens de gedragscode *Fostering Research Integrity in Europe* die de European Science Foundation (ESF) heeft gelanceerd. De code moet als referentiepunt voor alle onderzoekers dienen en nationale en Europese regelgeving aanvullen. De gedragscode is bedoeld om te laten zien wat gemeenschappelijk is in de dertig deelnemende landen, aldus het ESF, en kan ook nog model staan voor een wereldwijde *Code of Conduct*. Een pdf van het rapport kan worden gedownload vanaf de ESF-site www.esf.org.

Prijs voor Van Gogh Brievenproject

Het Van Gogh museum en het Huygens Instituut hebben de Grand Prix van de European Heritage Awards gekregen voor hun gezamenlijke Van Gogh Brievenproject. Dat gebeurde op 10 juni in Istanbul, dit jaar de culturele hoofdstad van Europa. De Awards worden jaarlijks toegekend door de Europese Commissie en de cultureel-erfgoedorganisatie Europa Nostra. De winnaars van 2010 werden gekozen uit bijna 140 projecten uit 26 landen. In eerste instantie kreeg het brievenproject één van de vier Awards in de categorie Research. Daarbovenop kwam later de Grand Prix met een bedrag van 10.000 euro. Het Van Gogh Brievenproject (www.vangoghletters.org) maakte de complete en gedocumenteerde verzameling brieven van de schilder openbaar toegankelijk en doorzoekbaar.

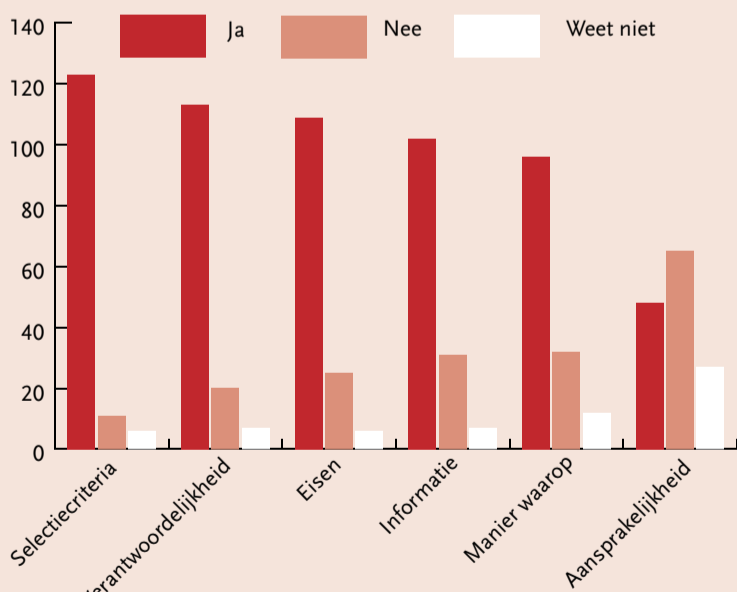
Emulatie en databeheer

Bij DANS is een onderzoek uitgevoerd naar de mogelijkheden van software-archivering, een van de grootste obstakels voor het duurzaam beschikbaar stellen van data. Emulatie wordt gezien als de meest kansrijke oplossingrichting. De onderzoekers, twee stagiaires van de TU Delft, hebben gekeken naar verschillende nationale en internationale initiatieven op dit gebied, zoals de Europese projecten PLANETS en KEEP. Ook schreven ze een prototype toepassing voor het beheer van emulatieomgevingen ten behoeve van het archiveren van data. De resultaten zijn op 18 augustus gepresenteerd bij de TU, het rapport is gepubliceerd op de DANS website (www.dans.knaw.nl). (Rutger Kramer)

Europese standaardisering

Drie Europese partijen die aan standaarden werken voor duurzame data-archivering gaan samenwerken. Ze tekenden in juni in Luxemburg een intentieverklaring voor 'een Europees raamwerk voor controle en certificering van digitale archieven'. Het gaat om het Data Seal of Approval, de Repository Audit and Certification Group van de samenwerkende organisaties voor het gebruik van ruimtedata CCSDS en de werkgroep data-archivering van de Duitse standaardiseringsorganisatie DIN. De Europese Commissie steunt het samenwerkingsinitiatief, waaraan ook het Landesarchiv Baden Württemberg en het UK Data Archive meedoen. In de overeenkomst wordt een samenhang nagestreefd tussen de drie certificeringsinitiatieven.

Verschillen in procedures bij deponeren



De meeste ondervraagde datamanagers in het Europese onderzoek PARSE.Insight (n=140) kennen selectiecriteria voor het al of niet archiveren van wetenschappelijke data. Dat geldt niet voor een regeling van de aansprakelijkheid als de data zoek of beschadigd raken, blijkt uit deze grafiek. In afnemende mate blijkt er verder sprake van duidelijke regelingen op de gebieden van

verantwoordelijkheid voor opslag en beheer, eisen aan standaardformaten, informatie over copyright op de gedeponeerde data en de manier van deponeren. De grafiek is een van de vele tientallen die vanuit het onderzoekproject worden aangeboden op de grafiekenwebsite Swivel. Behalve datamanagers werden onderzoekers en uitgevers ondervraagd. Zie ook elders in dit nummer.

<https://www.swivel.com/people/1015959-PARSE-insight>

Multifunctioneel informatiewetenschapper Herbert Van de Sompel: 'Iedereen wil een oplossing om webarchieven toegankelijk te maken, en wij hebben die oplossing!'

Tijdreizen op het web is een van de opvallende oogmerken van de in de Verenigde Staten werkende Belg Herbert Van de Sompel. Maar de veelzijdige informatiewetenschapper heeft meer spraakmakende internetprojecten op zijn naam staan. *e-data&research* verkende de ruime horizon van de geboren innovator tijdens een verblijf in Nederland.

PETER BOOT EN MARTIJN DE GROOT

Vanuit zijn werkplek bij het befaamde Los Alamos National Laboratory in New Mexico (VS) streek Van de Sompel deze zomer twee maanden als *visiting professor* neer bij DANS. Het duurde niet lang of de informatiewetenschapper had bij het gastheer-instituut en bij de Koninklijke Bibliotheek zijn plannen besproken voor de internet-telietijdmachine Memento, en bij de Koninklijke Nederlandse Akademie van Wetenschappen (KNAW) zijn visie op Open Annotatie voor de geesteswetenschappen.

Alfalab, het KNAW-project dat het gebruik van digitale gereedschappen in de humanities bevordert, biedt een aanknopingspunt om dat laatste onderwerp te belichten: 'Zij hebben twee demonstratieprojecten. Textlab werkt aan de annotatie van historische manuscripten, Spacelab aan die van kaarten met microtoponiemen, veldnamen dus. Die annotaties wil je kunnen delen tussen gebruikersgroepen en collecties, en dat moet Open Annotatie mogelijk maken door een gemeenschappelijk datamodel te definiëren. Daarmee kunnen verschillende projecten hun annotaties publiceren, waarna robots en harvesters de data eruit kunnen gaan zuigen en daarop nieuwe diensten aanbieden. Maar daarvoor zijn wel interoperabiliteitspecificaties nodig.'

Waarom interoperabiliteit?

'De gedachte is dat je mensen niet wilt dwingen hun interne systemen te veranderen. Wat je wel wil is een gemeenschappelijke interface die met alle partners kan samenwerken. Je moet dus specificaties opstellen die het voor de meeste annotatiesystemen mogelijk maken om mee te doen.'

En is er ook een verband met het populaire Firefox-hulpmiddel Zotero?

'Ook Zotero, dat onderdeel is van het Open Annotation Project, zou van het gedeelde datamodel gebruik moeten gaan maken. Zotero is vooral ontwikkeld om gegevens te verzamelen over bronnen, zoals webpagina's en wetenschappelijke literatuur. Daarvoor wordt nu een annotatiecomponent ontwikkeld, maar tot nog toe hebben ze zich geconcentreerd op de *user interface*, niet op de interoperabiliteit. Nu houden ze nog geen URI's bij. Dat moet natuurlijk veranderen. De



WIEBE KIESTRA

Van de Sompel, in Den Haag neergestreken vanuit New Mexico, had al snel oog voor de efficiency van het vervoer per fiets in zijn tijdelijke verblijfplaats.

Open Annotation Collaboration is geheel webgecentreerd.'

Uw project voor toegankelijkheid van webarchieven hangt hier eigenlijk direct mee samen.

'Ja, het is erg belangrijk dat annotaties in de tijd constant blijven. Je gaat internetbronnen annoteren, maar zo'n bron kan er morgen al anders uit zien. Je zou dus willen verwijzen naar een URI op een bepaald moment in de tijd, maar dan moet je die vroegere toestand van de URI wel terug kunnen vinden. Dat is het doel van Memento: tijdreizen op het web. En dat is natuurlijk niet alleen van belang voor annotaties. Het gaat ook om persberichten van het Witte Huis, commerciële en juridische gegevens, wetenschappelijke literatuur, en *what have you*. En het bewaren van versies van andere bronnen, zoals data, onderzoeksdata ook, dat is ècht belangrijk! Iedereen weet al zo lang dat we een oplossing nodig hebben voor het toegankelijk maken van archieven, en wij hebben die oplossing!

Je systeem heeft dus een archiveringskant nodig. Dat kan op verschillende manieren. Websites kunnen een eigen archief bijhouden, zoals de British Library al doet. Je kunt vertrouwen op externe archivering, bijvoorbeeld door het Internet Archive of door een nationale bibliotheek. Combinaties kunnen ook. De architectuur die wij voorstellen voorziet ook een plaats voor *aggregators*:

instellingen die informatie over beschikbare archieven verzamelen.

In het beste geval maakt een website zelf duidelijk waar het archief te vinden is. Dat gebeurt met een http header, extra informatie die een web server naar de browser of client terugstuurt, extra informatie die met een webpagina wordt meegestuurd. De universiteitsbibliotheek van Gent doet dat al. De client, bijvoorbeeld een browser zoals Firefox, maakt dan uit de Memento-header op waar hij het archief voor een bepaalde datum moet zoeken. Van DBPedia, de database-versie van Wikipedia, bestaat intussen al een Memento-compatibele versie. Die hebben wij opgezet, als voorbeeld. Maar meestal zullen die Memento-headers er niet zijn. Dan kan de client op zoek gaan naar een aggregator die de metadata al heeft verzameld. Wij hebben daar een proefopstelling van gemaakt, en het werkt.'

Is Memento het stadium van het experiment dus al voorbij?

'Het is een experiment in de zin dat er nog geen officiële specificaties bestaan. Maar het slaat wel aan. Bij de laatste World Wide Web conferentie heb ik het verhaal over Memento verteld met Tim Berners-Lee op de eerste rij. Die had fenomenaal lovende woorden (glimlacht trots).

De invoering van Memento hoeft niet ingewikkeld te zijn. De meeste webarchieven draaien allemaal dezelfde software, WayBack. Dat heb-

beeft afgesloten. We hebben daar een elegante oplossing voor bedacht.'

U heeft met cognitief psycholoog Johan Bollen ook een systeem ontwikkeld voor het aanbevelen van wetenschappelijke artikelen?

'Bollen en ik kwamen toevallig allebei uit België naar Los Alamos. We realiseerden ons dat de OpenURL software precies kan bijhouden hoe vaak er op links naar wetenschappelijke artikelen wordt geklikt. Dan weet je dus ook in welke context wetenschappers in bepaalde artikelen geïnteresseerd zijn. Daarmee kun je dezelfde soort dingen doen als Amazon: *kopers die dit boek kochten, kochten ook deze boeken*. Toen ik later opnieuw in Los Alamos werkte heb ik Johan teruggestuurd om het idee uit te werken. Dat heeft geleid tot de *bX Recommender Service*. Die is overgenomen door het bedrijf Ex Libris.'

U werkt ook aan Object Reuse and Exchange, ORE.

'Daarbij gaat het om samengestelde objecten. Bij een artikel horen onderzoeksgegevens, een video, afbeeldingen, en die moet je aan elkaar kunnen koppelen. Elk een eigen identiteit, maar je moet ook de aggregatie ervan, die een machineleesbare beschrijving geeft van metadata en componenten, een eigen identiteit kunnen geven.

Er was veel belangstelling voor het probleem maar het was geen *burning platform*. Daardoor kwam het wat langzamer op gang maar er wordt toch op veel plekken mee gewerkt, bijvoorbeeld in de DataNeT projecten in de VS, in de Europese digitale bibliotheek Europeana, maar ook hier bij DANS in het werk aan verrijkte publicaties.'

Is ORE bedoeld voor de wetenschappelijke uitgevers, die de gegevens over artikelen met ORE ter beschikking moeten stellen?

'Met uitgevers weet je het nooit. Vrijwel geen enkele uitgever heeft ooit het Protocol for Metadata Harvesting ingevoerd, dus... Ik hoop dat het gebeurt want dat zou veel dingen oplossen. Maar veel uitgevers verstopten nog altijd de metadata van publicaties. Het is echt een enorme blunder van de Open Access Movement dat men dat heeft laten liggen. Metadata vormen het *core* niveau, dat moet op nummer 1 staan. Het is *fuel for the engine*. Uitgevers weten dat ook. In hun nadruk op open access voor de inhoud van de artikelen zijn activisten en bibliotheken de metadata vergeten. Dat is kwalijk.'

Er wordt vaak gevochten om de status van iso-standaard of een W3C recommendation. Uw standaarden hebben die status vaak niet maar ze worden wel op grote schaal gebruikt.

'We hebben er wel eentje: de OpenURL. Maar ik heb geen duidelijk antwoord hoe dat komt. Eén denkbare oorzaak is dat het om echte problemen gaat die mensen graag opgelost willen zien. Bijvoorbeeld het Protocol for Metadata Harvesting PMH, dat was echt een *burning platform*! OpenURL was ook zo iets. Het ging erom te kunnen verwijzen naar wetenschappelijke artikelen in systemen van bibliotheken en wetenschappelijke uitgevers. Daarbij moet je rekening kunnen houden met de abonnementen die een bibliotheek

Herbert Van de Sompel studeerde in 1979 in Gent af als wiskundige en twee jaar later als computerwetenschapper, waarna hij in die plaats zijn loopbaan begon als hoofd bibliotheekautomatisering. Na onder meer een kort verblijf als research fellow aan het Los Alamos National Laboratory (LANL) promoveerde hij in 2000 op een proefschrift over het dynamisch linken van onderzoekbronnen. Na korte verbintenissen met The British Library en Cornell University werd hij in 2002 opnieuw aangesteld, nu als stafonderzoeker, bij het LANL waar hij tegenwoordig nog steeds aan is verbonden.

Focus

RONALD VAN DER BIE

Fryske Akademy

De Fryske Akademy in Leeuwarden bouwt samen met andere instituten een groot kennisnetwerk rondom de Friese taal.

De Fryske Akademy, gevestigd in het voormalige woonhuis in Lodewijk XIV-stijl van de bouwmeester van de Friese stadhouders Anthonius Coulon, is ouderwets op de gevel, maar modern en bijdetijds in gedrag.

Dr. Arjen Versloot, hoofd van de vakgroep Taalkunde van het instituut, leidt het onderzoek naar de Friese taal. Ooit, in 1938, was het maken van woordenboeken de hoofdtaak van het instituut. Het grote *Woordenboek der Friese taal (WFT)* is inmiddels klaar, 25 delen groot. Het eerste deel verscheen in 1984. In juli van dit jaar heeft Koningin Beatrix het woordenboek officieel online gezet. De Fryske Akademy beheert de online editie samen met het Instituut voor Nederlandse Lexicologie (INL) in Leiden, dat de software heeft waarmee het bestand te doorzoeken is.

De redactie heeft tientallen jaren vooral Friese woorden verzameld. Versloot: 'In 1975 is het verzamelen gestopt en zijn de redactieleden gaan schrijven. Het hart van de verzameling bestaat dus uit de woordenrijkdom en het taalgebruik van de negentiende en het grootste deel van de twintigste eeuw. In die zin bieden de woordenboeken ook een tijdsbeeld.'

Voor nieuwe, moderne woorden is het *Frysk Hânwurdboek (Fries Handwoordenboek)* uitgebracht, een zogeheten eentalig woordenboek waarin ook de verklaringen in het Fries zijn. Hoe levendig de Friese taal is, blijkt uit het feit dat het handwoordenboek ca. 70.000 lemma's bevat, waarvan 20.000 niet in het WFT staan.

Taaldatabank

Een nieuw project is het woordenboek Nederlands-Fries, waar veel vraag naar is. Maar het paradijsje is de Taaldatabank: een grote digitale database met literaire en niet-literaire teksten die uitvoerig doorzoekbaar zijn. Versloot: 'Alles wat tot 1800 in het Fries is geschreven, is inmiddels verzameld en gedigitaliseerd. We zijn nu bezig om het bestand te verrijken met allerlei onderzoek over de etymologie en herkomst van Friese woorden. Als je op een bepaald woord in een tekst klikt, krijg je te zien wat voor woord het is en wat voor vorm, en of er al eens iets over is geschreven. Je kan ook doorklikken naar het *Woordenboek der Friese taal*, en nagaan of het woord ook voorkomt in het Nederlands en misschien zelfs in het Middelnederlands.' Dat laatste is mogelijk via de links met de data-



LINUS HARMS

Dr. Arjen Versloot, hoofd vakgroep Taalkunde van de Fryske Akademy

base van het INL. Zo heeft de Fryske Akademy een gigantisch kennisnetwerk gebouwd rondom de Friese teksten vanaf de Middeleeuwen tot in de moderne tijd. 'De ambitie is om alles wat tot 1900 in het Fries is verschenen nog te verzamelen.'

Databank en bibliotheek

'Alle oude woorden zijn gekoppeld aan het moderne woord. Als je weet hoe het woord in het moderne Fries wordt gespeld, kan je in de database alle oude spellingvarianten terugvinden en door de koppeling aan de Nederlandse woordenboeken is het zelfs mogelijk om bij Nederlandse woorden de Friese woorden te vinden.'

Toegankelijk is de databank in deze uitgebreide vorm nog niet. De gebruikersinterface wordt op dit moment ontwikkeld. 'Als die klaar is, zal de databank geweldige mogelijkheden bieden tot onderzoek naar taalverandering, taalvariatie, taalgebruik en de verschillen daarin. Omdat alle oude teksten en veel moderne teksten opgenomen zijn, is de Taaldatabank tegelijkertijd een rijke digitale bibliotheek voor het Friese taalgebied.'

De Fryske Akademy is een van

de deelnemende KNAW-instituten in de nieuwe onderzoekeninfrastructuur Alfalab, CLARIN en TaalPortaal. Dat laatste is een virtueel instituut voor de Nederlandse en de Friese taal, waarin wordt samengewerkt met het Meertens Instituut, het INL en verschillende universiteiten.

HISGIS

Historici kunnen bij de Fryske Akademy voor veel terecht: voor administratieve en fiscale bronnenuitgaven, 16de-eeuwse testamenten en belastingregisters, voor naamkunde en nederzettingsgeschiedenis in pdf-formaat. Interactief is het historisch-geografisch informatiesysteem (HISGIS), een web-applicatie met de kadasterkaarten van 1832, waarop chronologische en thematische kaartlagen aansluiten die op diverse manieren met elkaar te combineren zijn. HISGIS Fryslân is het oudste onderdeel van HISGIS, dat over andere delen van Nederland uitgebreid wordt.

 www.fryske-akademy.nl
<http://gtb.inl.nl/>
www.fryske-akademy.nl/tdb
www.hisgis.nl

Lid van adviescommissie Marcel Broersma:

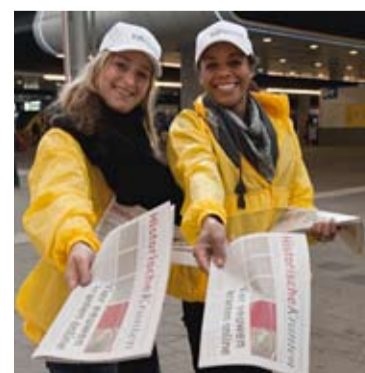
'Historische krantenwebsite scheidt ook verplichtingen'

De lancering van de website Historische Kranten van de Koninklijke Bibliotheek (KB) ging dit voorjaar gepaard met veel publiciteit. Maar hoe zit het met de wetenschappelijke inbreng in het project? En wat zijn de mogelijkheden en onmogelijkheden? *e-data&research* sprak daarover met Marcel Broersma, hoogleraar Journalistieke Cultuur en Media aan de Rijksuniversiteit Groningen en lid van de wetenschappelijke adviescommissie.

Over de invloed die de commissie heeft gehad is Broersma zeer te spreken. 'Het project wordt gesubsidieerd door de Nederlandse organisatie voor Wetenschappelijk Onderzoek (NWO). De KB heeft ook altijd centraal gesteld dat de website een belangrijke functie voor de wetenschap moest hebben. Daarom is het primaat over de selectie neergelegd bij deze adviescommissie, met mensen uit heel verschillende disciplines. Dat zijn niet alleen de mensen die de meeste kennis hebben van de materie, maar ook de mensen die ermee moeten werken. Je kunt natuurlijk niet alle toekomstige onderzoeksvragen voorzien, maar zo is er wel een corpus ontstaan met een groot scala aan onderzoeksmogelijkheden in diverse disciplines. De commissie heeft in alle onafhankelijkheid kunnen werken.'

Had de commissie ook invloed op de zoekmogelijkheden? Broersma: 'De mate van ontsluiting is eigenlijk alleen beperkt door wat praktisch en technisch mogelijk is. Natuurlijk kan het altijd beter en dieper, maar op een gegeven moment bereik je met extra functionaliteiten steeds kleinere doelgroepen. Dan moet je een afweging maken. Met wat er nu ligt kun je veel kanten op.'

Broersma noemt de voordelen van de website. 'Misschien voor de hand liggend, maar toch heel belangrijk: de afstand is opgeheven. Je moet niet onderschatten hoeveel tijd het kost om overal heen te reizen om kranten in te zien. Er zitten kranten bij uit de koloniën, uit het Vaticaan, zelfs uit Rusland. En: je kunt nu een groot aantal kranten tegelijk bestuderen. Eerdere lokale initiatieven boden steeds maar toegang tot één krant. Deze website geeft mogelijkheden om onderzoek te doen naar intertextualiteit. Hoe wordt vanuit di-



CASPER CAMMERAAT



Marcel Broersma tijdens de lancering van de website

verse invalshoeken bericht? Ook de nieuwsstromen, waarbij de een weer reageert op de ander, komen nu goed in beeld. Als je dat bladerend moet doorzoeken, kost dat ongelooflijk veel tijd. Dat moet je echt zelf een keer gedaan hebben om te weten hoe veel. En je kunt nu gericht op onderwerp zoeken. Ook dat was eerder niet mogelijk.'

Voor zijn eigen vakgebied ziet Broersma een aantal nieuwe onderzoeksmogelijkheden. 'Bijvoorbeeld de oeuvres van journalisten, maar ook de vormen van berichtgeving. Eind negentiende eeuw deed het interview zijn intrede als genre. Dat leidde tot een storm van protesten. Burgers vonden dat een volstrekte aantasting van de privacy.' Broersma is blij dat er nu ook gericht onderzoek kan worden gedaan naar de journalistieke tekst. Tot nu toe is die een beetje het stiefkindje geweest van het mediahistorisch onderzoek.

Niettemin heeft Broersma ook een waarschuwing voor potentiële gebruikers. 'Het gevaar dreigt dat kranten door historici als autonome bron worden beschouwd. Dat zijn ze niet. Je moet de berichtgeving in de context van de tijd en de journalistieke mores zetten om die op waarde te kunnen schatten.' Hij noemt een voorbeeld uit eigen ervaring: 'Eind negentiende eeuw lijdt de bevolking van Leeuwarden honger na een zeer strenge winter. De socialisten organiseren hongertochten. Huzaren worden aangevoerd om het oproer te onderdrukken en de staat van beleg wordt afgekondigd. Maar in de *Leeuwarder Courant* vind je daar helemaal niets van terug. Dat vond men te onfatsoenlijk om over te berichten, dat was sociaal niet verantwoord.' Waarmee Broersma maar wil zeggen: het gebruik van de historische krantenwebsite scheidt ook verplichtingen; er ligt een taak om de berichtgeving van de juiste context te voorzien. (IA)

 <http://kranten.kb.nl/>



 Bij de lancering van de website liet de KB op het Centraal Station van Den Haag papieren kranten uitdelen.

JOS ULTEE

E-depot Nederlandse archeologie gaat nieuwe fase in



Dagelijks komen nieuwe datasets binnen, waaronder kleine pareltjes als de opgraving 'Zaltbommel – de Wildeman'. De uitgebreide digitale brongegevens van de Romeinse nederzetting en begravingen en de omvangrijke publicatie 'Onder de zoden van Zaltbommel' (ADC monografie 8) is een belangrijke bron voor toekomstige archeologen.

Het elektronisch archief voor de Nederlandse archeologie (EDNA) heeft met meer dan 12.500 datasets inmiddels een zodanige omvang dat het gebruik ervan een zelfversterkend effect gaat vertonen. 'Steeds meer onderzoekers gebruiken de digitale informatie, zien de voordelen en zijn daardoor bereid ook zelf hun data en/of publicaties aan te leveren,' zegt Hella Hollander, die onlangs bij het archief als coördinator archeologie is begonnen. In de afgelopen jaren heeft het e-depot, dankzij een subsidie van de Nederlandse organisatie voor Wetenschappelijk Onderzoek NWO, een sterke groei gezien van het aantal gearchiveerde datasets. De komende jaren wacht een nieuwe uitdaging, aldus Hollander: 'Het digitaal archiveren van de brongegevens is door de Nederlandse archeologen belangrijk genoeg bevonden om opgenomen te worden in de nationale kwaliteitsnormen. Die verplichting hebben ze gezamenlijk geaccepteerd, mede op basis van de goede

ervaringen en gebruiksmogelijkheden met EASY.'

Langetermijnarchivering is echter niet zonder kosten. Een van de taken van de nieuwe coördinator is om te onderzoeken hoe die kosten gedragen kunnen worden. Hollander: 'Je hebt zowel eenmalige *ingestkosten* als structurele *archivingkosten*. Bij het archeologisch data-archief in Engeland wordt een deel van de kosten doorberekend aan de deponeerder, alseen soort éénmaligeleges. Dat is een kostenmodel waarin iedereen vooraf met een vast bedrag rekening kan houden: subsidiegevers, wetenschappers en commerciële onderzoeksbureaus. Op die manier kan aan het einde van een onderzoekproject de digitale informatie aan het e-depot worden overgedragen, waarna DANS garant staat voor de langetermijnarchivering. EDNA gaat onderzoeken of er genoeg draagvlak is voor dit model.' (Milco Wansleeben)

Ⓜ <http://tinyurl.com/3ynbo62>

ADC ARCHEOLOGIEPROJECTEN

Steeds meer datacontracten

DANS (Data Archiving and Networked Services) sluit steeds meer *datacontracten* af. Dat komt doordat de Nederlandse organisatie voor Wetenschappelijk Onderzoek (NWO) van onderzoekers die ze subsidieert, vraagt om de gegevens uit hun onderzoek duurzaam op te slaan. Op dit moment gebeurt dat bij de programma's van NWO Middelgroot en NWO Groot voor de gebieden Geesteswetenschappen en de Maatschappij- en Gedragwetenschappen. En in deze laatste discipline ook nog bij despecialeonderzoeksprogramma's Veranderingsstudies en Geweld tegen psychiatrische patiënten.

NWO wil stimuleren dat nieuwe onderzoekdata duurzaam worden gearchiveerd en ter beschikking worden gesteld voor verder onderzoek en nieuwe analyses. De meestal forse investeringen in het creëren van data mogen niet na afloop van het onderzoek weer verloren gaan zoals ook in het recente verleden nog vaak genoeg is gebeurd. Ook is de openbaarheid en daarmee de verificerbaarheid van wetenschappelijk onderzoek erbij gebaat.

De regels die NWO hanteert bepalen nu dat de betreffende onderzoeksdata bij DANS gearchiveerd worden óf bij het onderzoeksinstituut zelf. In dat laatste geval moet er een waarborg zijn dat de data ook goed gearchiveerd en toegankelijk blijven, juist ook op de langere termijn. In die waarborg kan worden voorzien door het Datakeurmerk (www.datasealofapproval.org).

De betrokken onderzoekers sluiten daarom direct na toekenning van een subsidie door NWO een datacontract af met DANS, waarin afspraken staan over het archiveren en beschikbaar stellen van de data. Toegekende subsidies worden door NWO pas uitbetaald als dat contract er is. (Heiko Tjalsma)

Ⓜ www.dans.knaw.nl/content/data-archief/reglementen-en-contracten/nwo-contracten

Digitale gids van de Bataafse tijd 1795-1813

Begin juni is de digitale Onderzoeksgids Bestuur en Administratie van de Bataafs-Franse tijd via het internet opengesteld. De gids, samengesteld door het Instituut voor Nederlandse Geschiedenis (ING), biedt aanrijpingspunten voor onderzoek naar een periode in de Nederlandse geschiedenis waarin vijf verschillende grondwetten zijn ontwikkeld die ook allemaal een nieuwe bestuursstructuur hebben opgeleverd.

De Bataafs-Franse tijd, van 1795 tot 1813, is volgens het ING van doorslaggevend belang geweest voor het ontstaan van het moderne Nederland: 'De overheid ging zich voor het eerst met alle mogelijke onderwerpen bezig houden zoals onderwijs, kerken, economie en armenzorg. De Franse 'bevrijders' van 1795 volgden de ontwikkelingen op de voet. In 1810 nam keizer Napoleon het land op in het keizerrijk. Toen kregen we

pas goed met bureaucratie te maken, maar ook met militaire strijd en een oorlogseconomie,' aldus een bericht van het instituut.

Speciale aandacht vragen de ontwikkelaars voor de digitale weergave van de centrale index van de staatssecretarie van 1806-1811, die samen met het Nationaal Archief werd gemaakt. 'Alle grote en kleine onderwerpen waarmee de overheid zich bemoeide komen daarin terug.'

De gids biedt inzicht in de samenhang van het bestuur per periode, taken en organisatie van de verschillende instellingen (218 landelijke en 108 gewestelijke) en de wetgeving die daarbij hoort, en in de betrokken personen via een koppeling aan het project 'Repertorium van Ambtsdragers en Ambtenaren'. Hij maakt ook raadpleging mogelijk van de archieven van de instellingen. (MdG)

Ⓜ www.inghist.nl/Onderzoek/Projecten/GidsBataafsFranseTijd

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 beveiligde microdata van het CBS kosteloos geleverd worden: www.dans.knaw.nl/nl/data/cbs/overzicht/

Centraal Bureau voor de Statistiek	Periode
Statistiek financiën van grote ondernemingen	2001-2009
Community Innovation Survey	2006
Wet Inkomen en Arbeid	2009
Werkloosheidsuitkeringen	2009
Gemeentelijke basisadministratie	1995-2009
Logiesverstrekkende bedrijven	2008
Productiestatistieken:	
Transportbedrijven	2000-2007
Commerciële dienstverlening	1995-2007
Via DANS EASY	
<i>Archeologie</i>	
Zaltbommel – De Wildeman (ADC Archeologisch Diensten Centrum)	2010
Egmond Binnen, Abdij van Egmond (Groningen Instituut voor Archeologie / Odyssee Jos Kleijne)	2010
Wijchen-Centrum (Hazenberg Archeologie, in het kader van NWO-Odyssee)	2010
Dordrecht, Plangebied Hofkwartier, Kloostertuin (Gemeente Dordrecht, Bureau Monumentenzorg & Archeologie)	2009
<i>Geschiedenis</i>	
Getuigen van Theresienstad. Erfgoed van de Oorlog, interviews afgenomen in het kader van het project Getuigen Verhalen (Radboud Universiteit, Fac. Religiewetenschappen)	1940-1945
<i>Geodata</i>	
Top 25 raster – Digitaal topografisch bestand (Kadaster)	2009-2010
Top 250 raster – Digitaal topografisch bestand (Kadaster)	2009-2010
Top 50 raster – Digitaal topografisch bestand (Kadaster)	2009-2010
Top 50 vector – Digitaal topografisch bestand (Kadaster)	2009-2010
<i>Sociale wetenschappen</i>	
Toezichtkaart Voortgezet Onderwijs (Voorheen Kwaliteitskaart) (Onderwijs Inspectie)	2010
The Nijmegen School Study (J. Tolsma, I. Pop – Radboud Universiteit Nijmegen)	2010
The Integration of the European Second Generation in Amsterdam and Rotterdam (TIES-NL) (Netherlands Interdisciplinary Demographic Institute – KNAW/NIDI)	2006-2007
Culturele Veranderingen in Nederland (CBS – beveiligd microbestand)	2008
Inkomens Panel Onderzoek – IPO (CBS – beveiligd microbestand)	2007-2008

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 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: 9100. *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: Inge Angevaere, Ronald van der Bie, Peter Boot, Martijn de Groot (hoofd/eindredacteur), Thijs Hermesen, Jetske van der Schaaf

Aan dit nummer werkten mee: Jorien Apperloo, Guido Gerritsen, Peter Groot, Jeffrey van der

Hoeven, René van Horik, Rutger Kramer, Willem Kuiper, Vincent Tassenaar, Jeroen van der Vliet, Milco Wansleeben, Paula Witkamp.

Redactiesecretariaat: Lucas Pasteuning, Jetske van der Schaaf

Vormgeving en opmaak: Ellen Bouma

Productie: Amsterdam University Press

Druk: Ten Brink, Meppel

ISSN: 1872-0374

Tijdcapsule PLANETS-project zorgt voor verwarring

Om het probleem van duurzame toegankelijkheid van digitale gegevens op te lossen zijn niet alleen technische hulpmiddelen nodig, maar ook draagvlak en financiering. Daarom was het een lovenswaardig initiatief van het technisch georiënteerde Europese PLANETS-project om ter gelegenheid van de afsluiting een evenement te organiseren waarmee ook de pers gehaald kon worden.

De projectmedewerkers bedachten een digitale tegenhanger van het *human genome project*: de 'PLANETS tijdcapsule'. In een zware ijzeren kist bracht het projectteam verschillende digitale dragers samen met informatie in diverse bestandsformaten, met daarbij (op papier!) alle informatie waarvan het team nu kon bedenken dat die later nodig zal zijn om de informatie te kunnen weergeven. Met enig uiterlijk vertoon werd de kist vervolgens ondergebracht in 'Swiss Fort Knox': een commercieel datacentrum, gevestigd in een oude atoombunker die is ingegraven in de Zwitserse bergen en die volgens de exploitanten volledige bescherming biedt tegen zowel atoombomben als chemische en biologische wapens.

Persbureau Reuters pikte het bericht op en maakte zelfs een videoreportage waarin wordt gesproken van een 'zero-risk facility'. Om het 007-gehalte nog wat op te schroeven gaf het persbureau in de video aan dat het de



Zware jongens kwamen eraan te pas om de capsule atoomvrij op te bergen.

exacte locatie van de bunker niet mag prijsgeven en evenmin de gezichten mag laten zien van de bewakingsmedewerkers. Reuters noemt de tijdcapsule een digitale 'Steen van Rosetta' die de toekomst in staat zal stellen onze huidige informatie te ontcijferen.


Enkele dagen na het evenement, op 26 mei, kopt het Nederlandse *InformatieProfessional* in zijn elektronische nieuwsbrief: 'Bomvrije en atoombestendige kluisdeuren bieden bescherming tegen dataverlies'. De oppervlakkige lezer zal geheel gerustgesteld zijn.

Ten onrechte. Want bomvrije bunkers helpen niet tegen technische ontwikkelingen die ervoor zorgen dat de computers van over pakweg twintig

jaar de huidige informatie niet meer kunnen lezen. De tijdcapsule kan wellicht een stukje digitale archeologie mogelijk maken, maar voor de exabytes aan wetenschappelijke informatie die we produceren en toegankelijk willen houden vormen betonnen muren en dikke sloten juist *geen* oplossing. Die informatie is veel meer gebaat bij continu en actief databeheer, bij gebruik.

Wie de website van de PLANETS tijdcapsule volledig leest, zal die nuanceringen ook tegenkomen. Daar wordt gesproken over het zichtbaar maken van de kwetsbaarheid van digitale informatie over twintig of vijftig jaar. Maar dan is het imago-kwaad al geschied. Het beeld van veilige bunkers met stevige bewakers heeft zich genesteld in de geheugens. Voor de nuanceringen moet je naar de website van PLANETS zelf en een paar keer doorklikken. Dat zullen maar weinig niet-deskundige lezers halen. Zij zullen rustig gaan slapen in de vaste overtuiging dat er veilige plekken zijn waar onze digitale informatie voor de toekomst is veiliggesteld. Daarmee is het tegendeel bereikt van wat de opzet was. En dat is jammer.

Inge Angevaere

 www.ifs.tuwien.ac.at/dp/timecapsule/home.html



Dr. Andreas Rauber (TU Wenen, links) en dr. Adam Farquhar van de British Library met de tijdcapsule.

Gelezen

Parse Insight Consortium: Science Data Infrastructure Roadmap; Permanent Access to the Records of Science in Europe; PARSE.Insight Deliverable D2.2; Report of project FP7-2007-223758, 2010 Report, providing an overview and initial details of both technical and non-technical components, needed to supplement existing and already planned infrastructures for science data. The components presented are aimed at bridging the gaps between islands of functionality, developed for particular purposes, often by other European projects, whether separated by discipline or time. Thus, they are intended to play a general, unifying role in science data. While developed in the context of a European-wide infrastructure, there would be great advantages if these types of infrastructure components would be available much more widely. Downloadable at www.parse-insight.eu

Eric Sieverts: Grotere rol voor informatie-professionals bij beheer van gegevensverzamelingen; InformatieProfessional 7/8

(2010); ISSN 1385 5328
Eerste bijdrage in een serie 'Data doen er toe'. Overzicht van de volgende bijdragen in deze serie: linked data, taak en competenties van de data librarian, modellen en standaarden voor beheer van datasets, een universitair datacentrum, DANS: data archivering – duurzaamheid – persistent identifiers, infrastructuur voor zelf-archiveren van datasets, met datasets verrijkte publicaties, 3TU een datarepository voor technische wetenschappen, juridische status van ruwe data.
www.informatieprofessional.nl/

Caroline Voorbrood m.m.v. Heleen van Luijn: Data – Voer voor psychologen? Archivering, beschikbaarstelling en hergebruik van onderzoekdata in de psychologie. Amsterdam, Aksant Academic Publishers, 2010; DANS Studies in Digital Archiving 4. ISBN 978 90 5260 374 2
Hebben psychologen behoefte aan een digitaal archief en aan de datadiensten van DANS? Data Archiving and Networked Services is in 2005 opgericht door

KNAW en NWO om de toegankelijkheid van databestanden voor alfa- en gammawetenschappers te verbeteren. Onder meer moest het de behoefte verkennen aan nieuwe data-archieven binnen de geestes-, maatschappij- en gedragswetenschappen. De DANS-archieven tellen relatief weinig psychologische datasets. Toch is het aantal psychologisch onderzoekers zeker zo groot als het aantal sociologen of andere maatschappijwetenschappers. Reden voor DANS om de datavoorziening en databehoefte in de psychologie te verkennen. Dit rapport, bedoeld voor discussie met het veld over datamanagement in de psychologie, doet daarvan verslag. www.dans.knaw.nl/content/bestelformulier

R. Feskens, J. Kappelhof, J. Dagevos, I. Stoop: Minderheden in de mixed-mode? Een inventarisatie van voor- en nadelen van mixed-mode dataverzameling onder niet-westerse migranten. Den Haag, Sociaal en Cultureel Planbureau, 2010; ISBN 978 90 377 0517 1

Column
Willem Kuiper

Naar een P.D.B.

Wie zich beroepshalve bezighoudt met het editeren van middeleeuwse teksten, zou de handschriften waarin die teksten bewaard bleven graag in zijn onmiddellijke fysieke nabijheid hebben. 'Analoge' reproducties van de bron, zeker als zij zwart-wit zijn, schieten te kort als het om details gaat. Het liefst zou je het handschrift naast je op een lessenaar open hebben liggen.

Dat kan dus niet.

Middeleeuwse handschriften en drukken liggen achter slot en grendel in kluisen van bibliotheken. En dat is maar goed ook, want dan wordt er optimaal voor gezorgd, én dat handschrift is (uitzonderingen daargelaten) opvraagbaar en raadpleegbaar. Geen grotere barrière voor een onderzoeker dan privé-bezit.

Het kan dus wel!

Digitalisering in hoge resolutie van middeleeuwse handschriften betekent veel meer dan beeldarchivering: zichtbaar wordt wat voor het ongewapend oog onzichtbaar is. Dankzij hoogwaardige digitalisering van de *Lancelot-compilatie* (KB Den Haag 129 A 10; binnenkort ook on-line) onderga ik de sensatie die Antonie van Leeuwenhoek voelde toen hij voor het eerst door zijn microscoop keek.

Het handschrift waarin de *Lancelot-compilatie* bewaard bleef is complex, kostbaar en kwetsbaar. Het komt normaal gesproken niet uit de kluis. Voor detail-onderzoek is dat een grote belemmering.

Maar nu het gedigitaliseerd is, kijk ik over de schouder van de kopiïst mee en zie ik zijn pen over het perkament bewegen. Ik kan de doorhalingen lezen, de verbeteringen en de ingrepen van een andere hand, die ik door de iets andere lettervorm en inktkleur nu wél van die van de kopiïst kan onderscheiden. Het handschrift mag van mij in de kluis blijven liggen. Dichterbij dan digitaal om te zien wat er staat, kan niet.

Het digitaliseren van middeleeuwse handschriften is geen sinecure. Wie ben ik om te twijfelen aan het digitaliseringsbeleid van de bibliotheken en hun conservatoren? Maar als ervaringsdeskundige weet ik ook dat het handschrift waar de individuele onderzoeker naar smacht vaak nog (lang) niet aan de beurt is.

Vandaar dat ik al jaren van droom van een P.D.B., een Persoonlijk Digitaliserings Budget. Met dat P.D.B kunnen onderzoekers jaarlijks een kostenloze bestelling doen in een Nederlandse bibliotheek: ik wil dát handschrift of die druk gedigitaliseerd hebben.

Strikt genomen is dit een budgettair neutrale beloning, want alles wordt met gesloten beurzen afgerekend. Het bewuste handschrift moet immers vroeg of laat toch gedigitaliseerd worden. Het enige verschil is dat het nu wat eerder gebeurt.

Als mijn droom werkelijkheid wordt dan reserveren bibliotheken een vrije ruimte binnen hun digitaliseringsbeleid en laten die dan invullen door individuele onderzoekers, die met andere ogen kijken.

Dr. Willem Kuiper is als onderzoeker verbonden aan de Leerstoelgroep Historische Nederlandse Letterkunde van de Universiteit van Amsterdam en het Huygens Instituut.

In Nederland is in de afgelopen twintig jaar ruime ervaring opgedaan met surveys onder niet-westerse groepen. Deze wijze van data verzamelen is echter duur en de respons blijft soms achter bij de verwachtingen. Reden om te bezien of andere dataverzamelmethode, bijv. m.b.v. internet of telefoon, tot een betere respons en lagere kosten leiden. Wellicht is het mogelijk om verschillende methoden in een onderzoek te combineren (wat 'mixed-mode' genoemd wordt). Dit rapport somt de voor- en nadelen op van deze mixed-mode aanpak voor surveys onder niet-westerse groepen. Het is gebaseerd op een uitgebreide literatuurstudie en interviews met experts. www.scp.nl

Toekomst voor ons digitaal geheugen (2): strategische agenda voor duurzame toegankelijkheid; Den Haag, Nationale Coalitie Digitale Duurzaamheid, 2010.



GISELLA KLEIN