

GEHOORD & BIJGEWOOND

De vluchtigheid voorbij:
Hollands webarchiveren

Peter de Bode

Voor het congres 'De vluchtigheid voorbij: webarchivering in Nederland' kwamen zo'n 130 belanghebbenden op 30 oktober bijeen in Hilversum. De KB, de NCDD en gastheer NIBG zorgden voor een afwisselend programma om meer inzicht te geven in en aandacht te vragen voor de praktijk van webarchivering in Nederland.

Na een introductie van Marcel Ras (NCDD) volgde de keynote van Helen Hockx-Yu, hoofd afdeling Webarchivering van de British Library. Zij legde uit hoe de webarchiveringsactiviteiten door de bibliotheken (verantwoordelijk voor de wettelijk-depottaken) worden georganiseerd. René Voorburg (KB) gaf een overzicht van de webarchiveringsactiviteiten van de KB. Dat niet alle websites zijn te archiveren, lieten Julia Vytopil (NIBG) en Chloé Martin (Internet Memory Research, IMR) zien in hun co-presentatie over audiovisuele websites.

Als laatste spreker gaf Hugo Huurdeman inzicht in de activiteiten van WebART. Een van zijn constatering is dat onderzoekers het webarchief voornamelijk gebruiken voor data mining.

Tot slot gingen sprekers en zaal onder leiding van Jantje Steenhuis (Stadsarchief Rotterdam) met elkaar in gesprek over de wenselijkheid van opgelegde regels voor webontwerpers in de publieke sector, het wel of niet archiveren van sociale media en het nut van samenwerking tussen instellingen die zich met webarchivering bezighouden. Over het laatste was iedereen het eens.

ncdd.nl/studiedagwebarchivering

British Library heeft
eigen onderzoekslab

Steven Claeysens

Aan de overzijde van Het Kanaal loopt sinds een paar jaar een boeiend experiment. Gefinancierd door de Andrew W. Mellon Foundation



Een resultaat van het onderzoekslab van de BL: de Victorian Meme Machine van Bob Nicholson, een omvangrijke databank met Victoriaanse grappen. De grappen worden uit gedigitaliseerde BL-collecties gehaald. Behalve als onderzoeksinstrument om belegen humor te analyseren, worden ook deze en andere semi-automatisch nieuwe memes gecreëerd door de grappen te combineren met Victoriaanse afbeeldingen. credits Bob Nicholson

startte de British Library (BL) het Labs-project. In het project verkent de Britse nationale bibliotheek nieuwe vormen van onderzoek op basis van de eigen digitale collecties, steeds in nauwe samenwerking met wetenschappers. Jaarlijks wordt een symposium gehouden om de resultaten te delen.

Na de keynote, waarin Tim Hitchcock (University of Sussex) de zorg uitsprak dat de geesteswetenschappen zich in het 'big data'-tijdperk te weinig bekommeren om 'small data', passeerden ook dit keer weer een paar interessante project de revue. Zo presenteerde Desmond Schmidt (University of Queensland) TILT, de Text to Image Tool. TILT koppelt op een slimme manier teksttranscripties met een manuscript of druk, zonder dat de software probeert de tekst te herkennen. Een ander voorbeeld is van Bob Nicholson (zie afbeelding). Beide projecten bewijzen dat Hitchcocks zorg misschien groter is dan nodig. labs.bl.uk

Semantische bruggen
tussen digitale collecties

Jeffrey van der Hoeven

Hoe kunnen we een semantische brug slaan tussen al die digitale collecties die wereldwijd online beschikbaar zijn? Deze vraag vormde het centrale onderwerp tijdens een speciale bijeenkomst over Linked Open Data (LOD) op 1 en 2 december 2014 in New York City, georganiseerd door de Andrew W. Mellon Foundation. Diverse universiteitsbibliotheken en andere wetenschappelijke en culturele instellingen uit de VS en Europa kwamen samen om hun visie en aanpak op het gebied van LOD te bespreken. Uit Nederland waren de KB en Europeana aanwezig.

Veel instellingen maken hun (meta)data op het web beschikbaar. Ze doen dat steeds vaker in een vorm (RDF) die hergebruik en vindbaarheid vergemakkelijkt. Zo heeft de Bibliothèque Nationale de France al een groot deel van haar collecties op <http://data.bnf.fr> in

RDF beschikbaar gemaakt en biedt het samenwerkingsproject Linked Data for Libraries (LD4L) van Stanford, Harvard en Cornell de mogelijkheid virtuele collecties en relaties te vormen op basis van metadata uit ieders bibliotheek. Dat is ook de volgende logische stap na web presence: datasilos met elkaar verbinden zodat de eindgebruiker vanuit één startpunt gemakkelijk nieuwe bronnen kan vinden zonder diverse databanken af te hoeven struinen. En beter nog: men kan verbanden ontdekken die via traditionele zoekmachines niet gevonden worden. De bijeenkomst in New York richtte zich op de uitdagingen die bij het vastleggen van semantische relaties komen kijken.

De sleutel is het opzetten van reconciliation engines: knooppunten waar relaties over collectiegrenzen heen worden vastgelegd. Doordat er al heel wat grote datahubs zoals DBpedia (data-representant van Wikipedia), VIAF (auteurs) en FOAF (menselijke relaties) in de wereld bestaan, kan het linken naar één van deze hubs al grote winst opleveren. Het voorstel is echter om klein te beginnen door eigen specialistische collecties te verbinden met soortgelijke collecties van andere organisaties, nationaal en internationaal. Dit brengt nieuwe informatie aan het licht in plaats van te verwijzen naar wat al bestaat. Relaties tussen personen en plaatsen genieten daarbij de voorkeur omdat deze relatief makkelijk te herkennen zijn in bronnen. Daarna kan dit uitgroeien tot een groter verknoopt netwerk van semantische relaties.

Tot slot werd onderkend dat eenvoud het uitgangspunt moet zijn. Linked Open Data is mooi, maar als het te complex is en geen gebruiksvriendelijke toegang heeft, gebruikt niemand het. Volgend jaar komt er mogelijk een vervolg op deze bijeenkomst om de voortgang te evalueren en vervolgfafspraken te maken.

mellon.org

KORT

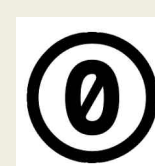
NWO start met pilot
datamanagement

Om data voortkomend uit door NWO gefinancierd onderzoek zoveel mogelijk toegankelijk en herbruikbaar te maken, is NWO gestart met een pilot datamanagement. Wetenschappers moeten aangeven hoe ze hun onderzoeksdata gaan opslaan en vindbaar en geschikt gaan maken voor hergebruik. De pilot geldt voor acht financieringsrondes. NWO gebruikt de input uit deze pilot als basis voor doorontwikkeling van beleid en procedures op het gebied van datamanagement. Bovendien hoopt NWO met deze pilot bij te dragen aan de informatie-uitwisseling en samenwerking tussen organisaties en initiatieven voor toegang tot en opslag van onderzoeksdata. (HB)

nwo.nl/datamanagement

Delen van data nog
makkelijker in EASY

DANS maakt het delen van onderzoeksdata nog makkelijker. In het online archiveringssysteem EASY kunnen deponerders sinds eind 2014 datasets volledig open access beschikbaar stellen conform de Creative Commons Zero-licentie (CC0). Deze internationale standaard houdt in, dat de de-



poneerder afstand doet van alle rechten tot de data. Geïnteresseerden

kunnen de open access-datasets vervolgens zonder enige vorm van registratie downloaden en vrijelijk hergebruiken. Wel vraagt DANS van zijn gebruikers om hierbij de data te citeren volgens de gebruiken in de wetenschap. (HB)

dans.knaw.nl

Call for papers
5e ESS-workshop

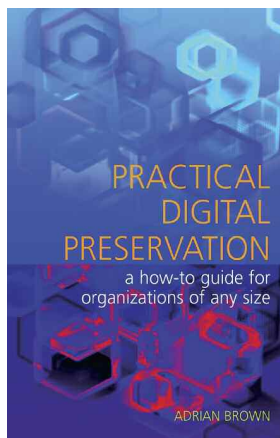
Op 22 mei organiseert DANS samen met Prof. dr. Gerbert Kraaykamp en Dr. Roza Meuleman (nationaal coördinatieteam European Social Survey, ESS) de vijfde Nederlandse ESS-workshop. De workshop biedt onderzoekers de gelegenheid om informatie uit te wisselen over de vele mogelijkheden van ESS-data. Tot 31 maart kunnen abstracts (plusminus 300 woorden) per e-mail ingediend worden bij Roza Meuleman (r.meuleman@maw.ru.nl). De organisatie is, naast papers die zich richten op een cross-nationale vergelijking, vooral geïnteresseerd in papers met een trendperspectief. Ook papers die ingaan op methodologische aspecten van de ESS-dataverzameling zijn welkom.

dans.knaw.nl

Brown winnaar Digital Preservation Award

Op 17 november jongstleden zijn de winnaars van de Digital Preservation Awards 2014 bekend gemaakt. De NCDD, één van de sponsors van de prijs, reikte de NCDD Award for 'Teaching and Communications' uit aan Adrian Brown voor zijn boek 'Practical Digital Preservation'. Het boek omschrijft de kennis en ervaring van de auteur met het ontwikkelen

van operationele digitale duurzaamheidsdiensten bij drie verschillende organisaties. Het boek biedt praktische handreikingen met name voor kleinere organisaties. Met deze en andere Digital Preservation Awards worden initiatieven beloont die een bijdrage leveren aan het duurzaam toegankelijk houden van digitaal erfgoed. (HB) dpconline.org



FameLab: strijd om de beste

FameLab is een internationale wedstrijd in wetenschapscommunicatie voor jonge bèta-wetenschappers en ingenieurs. Tijdens de nationale competitie FameLab strijden jonge onderzoekers om het beste onderzoeksverhaal. Inschrijven is niet meer mogelijk, maar het bijwonen van de Nederlandse finale op 21 april wel. De website van The British Council, de initiatiefnemer, geeft meer informatie. (HB) britishcouncil.nl/famelab



COLOFON Redactieadres: E-data & Research, Postbus 93067, 2509 AB Den Haag, 070-3494450, edata@dans.knaw.nl, www.edata.nl.

Hoofd-/eindredacteur: Heidi Berkhout. Redactie: Marika de Bruijne, Steven Claeysens, René van Horik, Ilja Nieuwland, Rutger Nugteren, Erica Renckens. Redactiesecretariaat: Lucas Pasteuning. Aan dit nummer werkten mee: Peter de Brode, Daan Broeder, Millie Elsen, Marjan Grootveld, Jeffrey van der Hoeven, Hella Hollander, Hein de Kort, Jacqueline Maenhout, Lindi Melse, Bob Nicholson, Jan Odijk, Marieke Polhout, Noëlle Sant, Marion Wittenberg. Opmaak: Colette Sloots, Haarlem. Productie: Amsterdam University Press. Druk: Ten Brink, Meppel. Webmaster: Sonja Duijkers. Oplage: 7500 papier, 1900 digitaal. ISSN: 1872-0374. We hebben getracht alle belanghebbenden met betrekking tot het gebruikte beeldmateriaal te benaderen. Degenen die menen rechten te kunnen doen gelden, kunnen zich tot ons wenden. Abonneren: Toezending papieren en/of digitale versie is kosteloos aan relaties van de stakeholders en studenten in de alfa- en gamma-richtingen. Overname artikelen: Wilt u een artikel uit E-data overnemen? Neem dan contact op met de hoofdredacteur. Overname van artikelen geschiedt altijd onder vermelding van de bron E-data & Research en de naam van de auteur van het artikel.