

KB digitaliseert door ondanks minder middelen

Eind 2009 zette de Koninklijke Bibliotheek (KB) het ambitieuze beleidsvoornemen op papier om alle publicaties van en over Nederland vanaf 1470 digitaal beschikbaar te stellen. Het initiatief ontmoette waardering maar ook twijfel over de uitvoerbaarheid, zeker in een tijd van bezuinigingen. Toch blijkt de KB nog geen anderhalf jaar later goed op weg te zijn. Om met minder middelen toch veel materiaal online te kunnen aanbieden werd samenwerking gezocht met private partijen. Zo worden 160.000 boeken uit de achttiende en negentiende eeuw gedigitaliseerd door Google en 50.000 bijzondere oudere drukken door uitgever ProQuest. In het conserveringsprogramma Metamor-

foze werkt de KB samen met het Nationaal Archief verder aan de digitalisering van boeken, kranten en tijdschriften van na 1840, waaronder twee miljoen krantenpagina's uit de periode 1850-1950. Astrid Verheusen, programmamanager Digitale Bibliotheek van de KB: 'De eerste fase van onze beleidsdoelstelling, tien procent online in 2013, gaan we daarmee zeker halen.'

Achter de schermen werkt de KB ook hard aan een betere beschikbaarheid van het materiaal. Verheusen: 'In de loop der jaren hebben veel digitaliseringsprojecten eigen websites ontwikkeld. Historisch gezien is dat goed te verklaren, maar we merken nu dat gebruikers moeite hebben om snel te vinden wat ze



KONINKLIJKE BIBLIOTHEEK

zoeken. Daarom streven we naar een platform waarop je alles kunt vinden wat de KB aanbiedt. En misschien

wel meer, want we willen ook samenwerkingsmogelijkheden onderzoeken met andere bibliotheken.' (IA)

@ www.kb.nl/hrd/digitalisering/index.html

SINDS KORT BESCHIKBAAR

Het overzicht toont databestanden die recent voor onderzoekers beschikbaar zijn gekomen bij CBS, CentERdata, Huygens ING en DANS. Een volledig overzicht van de CBS-bestanden is te vinden op www.cbs.nl/microdata. De LISS panel bestanden van CentERdata zijn kosteloos beschikbaar via het

LISS Data Archief www.lissdata.nl/dataarchive. De bestanden van Huygens ING zijn beschikbaar via www.inghist.nl. De DANS databestanden komen van diverse andere onderzoeksinstituten. Deze kunnen kosteloos worden gedownload vanuit DANS EASY: <https://easy.dans.knaw.nl>

Centraal Bureau voor de Statistiek	Periode
Productiestatistieken:	
Zorg en Welzijn	2003-2007
Detailhandel	1993-2008
Transport	2000-2008
Rijksfinanciën	2007
Investeringen in materiële vaste activa	2008
Enquete ICT-gebruik bedrijven	1995-2008
Woonruimteregister verrijkt met Voorraad woonruimte per adres	2008
CentERdata LISS Data Archive	
Ambiguity Aversion and Household Portfolio Choice (Dimmock, S.; Kowenbergh R.; Wakker, P.)	2010
Restraint Eating vs. Food Worry and Power of Food as Competing Explanations of Overeating and Overweight: Part 1 and 2 (Ridder, D. de)	2010
Explaining Public Support for Vigilantism: Part 1 and 2 (Haas, N.)	2009
LISS Core Study – Economic Situation: Housing, Wave 3 (CentERdata)	2010
LISS Core Study – Economic Situation: Income, Wave 3 (CentERdata)	2010
LISS Core Study – Economic Situation: Assets, Wave 2 (CentERdata)	2010
LISS Core Study – Work and schooling, Wave 3 (CentERdata)	2010
Perceived Diversity Effects on Organizational Outcomes (Garib, Y.R.)	2010
Waist Circumference (CentERdata)	2010
Commercial Opportunities (Block, J.; Gruber, M.; Petty, J.S.)	2010
Huygens ING	
Nieuwe brieven van Willem van Oranje	1572-1573
Digitale gids Grondwetcommissies	1883-1983
Registers van de graven van Holland in de Henegouwse periode	1299-1345
Namenindex op databases en gedigitaliseerde publicaties Huygens ING	2010
Kroniek van Kattendijke – Rijks Geschiedkundige Publicatiën	1490
Brieven en andere bescheiden Daniel van der Meulen – Rijks Geschiedkundige Publicatiën	1584-1600
Bronnen betreffende de Midden-Molukken – Rijks Geschiedkundige Publicatiën	1900-1942
Bronnen tot de geschiedenis der Leidsche universiteit – Rijks Geschiedkundige Publicatiën	1574-1811
Via DANS EASY	
Historische Wetenschappen	
Interviews Veteranen Instituut – IPNV	2009
Archeologie	
Een archeologische begeleiding (protocol opgraven) aan de Eexterweg te Gieten, gemeente Aa en Hunze (Dr) (ARC)	2009
Westland Poeldijk De Kreken fase II opgraving (ADC ArcheoProjecten)	2010
Life on the watershed. Reconstructing subsistence in a steppe region using archaeological survey: a diachronic perspective on habitation in the Jordan Valley (Eva Kaptijn)	2009
Noordwijk, 't Pleintje, Gemeente Noordwijk (Becker en Van de Graaf b.v.)	2009
Visresten van Albrandswaard-Portland. Archeozoologisch onderzoek van de laat-neolithische vindplaats 20-134 (4145 ± 45 BP) (BOOR)	2008
Sociale Wetenschappen	
Arbeidsvraagpanel (voorheen OSA Arbeidsvraagpanel) (A.M. de Voogd-Hamelink, Sociaal en Cultureel Planbureau, Den Haag)	1991-2008
International Comparative Study of Ethno-Cultural Youth (ICSEY team; Paul Vedder, Rommert Casimir Institute, Leiden University)	2006
Mobiliteitsonderzoek Nederland – MON (Ministerie van Infrastructuur en Milieu – Rijkswaterstaat)	2009
Tijdsknelpunten in de samenleving – tijd op orde (Sociaal Cultureel Planbureau, Den Haag)	2009

DANS wil na onderzoek data gaan reviewen

DANS (Data Archiving and Networked Services) gaat onderzoeken of het in zijn archiefsysteem EASY een standaardvraag aan afnemers om feedback kan opnemen. De antwoorden zouden vervolgens aan andere EASY-gebruikers meer informatie verschaffen over de betreffende dataset.

Eind vorig jaar voerde het instituut een pilotonderzoek uit om het reviewen van onderzoeksdata te stimuleren. 140 mensen beoordeelden de kwaliteit van de dataset(s) die zij uit het EASY-archief hadden gedownload. 92% van de respondenten zou de dataset aan anderen aanbevelen, zo bleek. De gemiddelde scores voor aspecten als de kwaliteit van de data en de documentatie en structurering binnen de dataset lagen rond de 4 op een schaal van 1 tot 5. 'Kwaliteit van de data' stond bovenaan met 4,12. De dataset helpt in 70% van de gevallen om onderzoeksvragen te beantwoorden. Overigens downloadt niet iedereen vanwege een onderzoeksvraag: ook 'pure interesse' wordt als reden genoemd. DANS vindt deze uitkomsten waardevol voor zichzelf, maar ook voor de onderzoekers die de datasets hebben gedeponeerd.

Respondenten kenden in het kader van de pilot ook tags toe aan de dataset die ze beoordeelden. Deze trefwoorden bleken overwegend inhoudelijk, bijvoorbeeld 'Romeinse tijd' of 'woningmarkt', maar ook meningen zoals 'weinig toegankelijk' en 'handig' komen voor. Voorbeelden van gereviewde datasets zijn te vinden op de DANS-website.

In december verscheen ook een verkenning van kwaliteitsbeoordeling van onderzoeksdata in het kader van SURFshare. Volgens dat rapport bestaan er voor datasets die door onderzoeksfaciliteiten en overheidsorganisaties worden geprodu-

ceerd, veel mechanismen om kwaliteit te monitoren en te verbeteren. *Supplementary data* – behorend bij publicaties – worden echter in praktijk slechts beperkt betrokken in de peer review. Ook is onderzocht op welke manieren de kwaliteit van datasets te stimuleren valt. Hiertoe kreeg een landelijke steekproef van hoogleraren en hoofddocenten negen opties voorgelegd. Drie opties blijken populair in alle disciplines: het opzetten van tijdschriften voor datapublicaties, het citeren van datasets en hergebruikers commentaar laten geven op de kwaliteit. Ook deze uitkomst is volgens DANS een stimulans om een review-mogelijkheid aan EASY toe te voegen.

Marjan Grootveld

@ <http://datareviews.dans.knaw.nl/index.php>
www.surffoundation.nl/nl/publicaties/Pages/Verkenningonderzoek.aspx

Negen nieuwe CLARINprojecten van start

Begin 2011 gaan negen nieuwe projecten van start die zijn gehonoreerd binnen de tweede call van CLARIN-NL, het infrastructuurproject voor de geesteswetenschappen. De projectuitvoerders zullen bestaande data en analyseinstrumenten geschikt maken om in de CLARIN-infrastructuur te worden opgenomen. Het gaat hier onder andere om de Nederlandse Liederbank, een historische verzameling van artikelen uit vijf grote Nederlandse kranten en een PoS-tagger. In de huidige call ligt de nadruk op data en instrumenten uit de letteren, geschiedenis, taalverwerving en historische taalkunde. 'Wij hopen hiermee over de gehele breedte van de geesteswetenschappen onderzoekers enthousiast te maken over de voordelen die de infrastructuur hun bij de uitvoering van hun onderzoek kan bieden', zegt CLARIN-NL-directeur Jan Odijk.

CLARIN is een Europees initiatief om e-science mogelijk te maken binnen de geesteswetenschappen en de sociale wetenschappen. Ruim dertig instituten uit 25 verschillende landen werken als partners samen aan een infrastructuur waarbinnen onderzoekers vrij toegang krijgen tot de data die zij voor hun onderzoek nodig hebben. Bovendien kunnen onderzoekers gebruik maken van speciale tools om relevante data te vinden en de gevonden data te analyseren en te verrijken, zoals automatische zinsontleders en textmining-software. (ER)



@ www.clarin.nl

COLOFON

e-data@research is het kwartaalblad in Nederland over data en onderzoek in de alfa- en gammawetenschappen. Het verschijnt onder auspiciën van CentERdata, CLARIN-NL, DANS, het Huygens ING, 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 gammamrichtingen. 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)3446484 f (070)3494451 e edata@dans.knaw.nl

Redactie: Inge Angevaere, Eric Balster, Ronald van der Bie, Peter Boot, Martijn de Groot (hoofd/eindredacteur), Thijs Hermsen, Erica Renckens, Jetske van der Schaaf

Aan dit nummer werkten mee: Onno Crasborn, Marjan Grootveld, René van Horik, Jeroen Sondervan, Anja Volk.

Redactiesecretariaat: Lucas Pasteuning, Jetske van der Schaaf

Vormgeving en opmaak: Ellen Bouma, Alkmaar

Productie: Amsterdam University Press

Druk: Ten Brink, Meppel

ISSN: 1872-0374