

Wat kost duurzaam archiveren?

Vervolg op digitalisering Volkstellingen



Palaiologk presenteert haar conclusies aan de medewerkers van DANS. Links DANS-directeur Peter Doorn.

Econome Anna Palaiologk uit Griekenland heeft de afgelopen maanden bij Data Archiving and Networked Services (DANS) een onderzoek uitgevoerd naar de kosten van duurzame toegankelijkheid van onderzoekgegevens. Zoals ook al bleek uit het onlangs in concept verschenen rapport *Toekomst voor ons digitaal geheugen* (zie p. 1 en 4) is daarover momenteel nog te wei-

nig bekend om financiers en data-archieven in staat te stellen gedegen afwegingen te maken tussen de kosten en de baten.

Palaiologk bestudeerde de beschikbare literatuur, met name de samenvatting van onderzoek tot nu toe in het Brits-Amerikaanse rapport *Sustaining the digital investment*, en voerde uitgebreide interviews met DANS-medewerkers om

erachter te komen welke werkzaamheden komen kijken bij het archiveren van onderzoekgegevens. Dit zou de sector in staat moeten stellen om het *activity-based costing model* toe te passen dat bekend is uit andere sectoren. Maar daarvoor is nog het nodige voorbereidende werk nodig, concludeerde Palaiologk. Zo ontbreken ervaringscijfers op basis waarvan begrotingen gemaakt kunnen worden omdat duurzame archivering een erg jong vakgebied is. Dit gemis zal ertoe leiden dat begrotingen de eerstkomende tijd een zeer grote onzekerheidsfactor hebben.

Om die factor te verkleinen is het nodig om de juiste vragen te gaan stellen, en juist daar heeft de Griekse zich in deze fase op geconcentreerd. Welke werkzaamheden reken je toe aan digitale archivering? Wat zijn de 'producten' die digitale archivering oplevert? Wat zijn de variabelen waar men mee te maken krijgt? De antwoorden op deze vragen kunnen verschillen per organisatie of zelfs per afdeling; ze zijn afhankelijk van de strategie die een organisatie vaststelt.

DANS is door het onderzoek van Palaiologk aan het denken gezet in de goede richting, concludeerde DANS-directeur Peter Doorn bij het afscheid van de econome, maar het laatste woord over wat duurzame archivering kost is nog lang niet gezegd. (IA)

http://brtf.sdsc.edu/biblio/BRTF_Interim_Report.pdf



Het Centraal Bureau voor de Statistiek (CBS) en Data Archiving and Networked Services (DANS) bereiden een vervolg voor op het grote Volkstellingenproject dat werd uitgevoerd tussen 1997 en 2006. In enkele nieuwe projecten worden nu gegevens gedigitaliseerd van een aantal bedrijfstellingen, volks-, beroeps- en woningtellingen, deels op wijk- en buurtniveau. In totaal gaat het om circa 350 duizend pagina's.

Ook wordt de website www.volks-tellingen.nl, die maandelijks zo'n 3.500 bezoekers trekt, gerenoveerd. Onderzocht wordt of die kan worden uitgebouwd tot een 'portal historische tellingen' met toegang tot andere websites met historisch statistisch materiaal, zoals de website van het CBS, de Hub for Aggregated Social History (HASH), NL-GIS en de Historische Databank Nederlandse Gemeenten (HDNG).

In het oorspronkelijke Volkstellingenproject werden de publicaties

van de Nederlandse volkstellingen van 1795 tot en met 1971 gedigitaliseerd. Daarbij kwamen rond 42 duizend pagina's aan data beschikbaar voor nieuw onderzoek, niet alleen als afbeelding maar ook in de vorm van elektronisch verwerkbaar gegevens. Het beoogde vervolg gaat niet alleen om het digitaliseren van nog meer gegevens, maar ook om voorzieningen die nodig zijn om de grotere hoeveelheid materiaal beter toegankelijk en bruikbaar te maken.

In een *expert meeting* medio juni zijn daarvoor ervaringen en suggesties uitgewisseld tussen een twintigtal onderzoekers die in de afgelopen jaren bij het Volkstellingenproject betrokken waren. De deelnemers waren afkomstig van het CBS, DANS en verschillende universitaire instituten. In korte presentaties gevolgd door discussie werden nieuwe mogelijkheden voor vormgeving en presentatie van gedigitaliseerde historische tellingen verkend. (Jan Jonker)

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
Algemeen Bedrijfsregister (ABR-GPSKADER)	2006-2008
Bedrijfsfraude (BFS)	2005-2007
Bijstandsfraude en Bijstandsdebiteurenstatistiek (BFS-BDS)	2008
Bijstandsuitkeringenstatistiek (BUS-TRANS)	2008
Doodsoorzakenstatistiek (DO)	2006-2007
Productiestatistiek Industrie	1993-1999
Statistiek re-integratie gemeenten (SRG)	2008
Beschikbaar via DANS EASY	
<i>Archeologie</i>	
Zandmaas – Stuwpannd Lith zomerbedverdieping km 176-181 (Projectteam Archeologie Maaswerken, Alterra)	2004
Schipluiden – Harnaspolder, campagne 03HP (Archeologisch projectbureau ADC)	2006
Mierlo – 't Karrewiel (Archeologisch projectbureau ADC)	2007
Plangebied Bedrijventerrein Avelingen te Gorinchem (Projectbureau RAAP)	2008
<i>Geschiedenis</i>	
Dutch Atlantic connections: the circulation of people, goods and ideas in the Atlantic world (Universiteit Leiden)	1680-1795
Erfgoed van de Oorlog; interviews afgenomen in het kader van het project Getuigen Verhalen:	1940-1945
<ul style="list-style-type: none"> Anna's stille strijd: doven in de oorlog (AVA filmproducties) Het vergeten bombardement, Rotterdam 31 maart 1943 (Stichting Mondelinge Geschiedenisprojecten) Retourtje Sachsenhausen (Nationaal Monument Kamp Vught) Negen vrouwen ontsnapt uit dodenmars 1945 (Ange Wieberdink Producties; restricted access) Vluchten uit Overloon op last van de bezetter (Liberty Park Overloon; restricted access) Boerinnen en boerendochters in de Tweede Wereldoorlog (Meertens Instituut; restricted access) NSB vrouwen in de oorlog (IIAV; restricted access) 	
<i>Sociale wetenschappen</i>	
Kwaliteitskaart Voortgezet Onderwijs (Inspectie van het Onderwijs)	2009
Cohortonderzoek Onderwijsloopbanen van 5-18 jaar COOL 5-18 (ITS Radboud Universiteit, SCO Kohnstamm Instituut)	2007
Bont en blauw, een onderzoek naar de behandeling van geweldszaken tegen politieambtenaren (Ministerie van Justitie, WODC)	2008
Sociale uitsluiting (Sociaal en Cultureel Planbureau)	2003
Centerdata millennium survey (Centerdata, Universiteit Tilburg)	1999

Symposium bij 45 jaar e-archief

Begin december organiseert DANS (Data Archiving and Networked Services) het symposium 'Gedreven door Data' over nieuwe kansen voor onderzoek in de alfa- en gammawetenschappen door de groeiende beschikbaarheid van elektronisch gearchiveerde data.

Aanleiding is het samenvallen van een aantal lustrums in 2009. Zo is het vijfenveertig jaar geleden dat het eerste elektronisch data-archief in Nederland werd opgericht: het Steinmetzarchief voor de sociale wetenschappen. Daarnaast is

het twintig jaar geleden dat het Nederlands Historisch Data Archief in het leven werd geroepen, terwijl vijftien jaar geleden NWO en het CBS besloten om het Wetenschappelijk Statistisch Agentschap op te richten, dat zou gaan zorgen voor de verspreiding van CBS-data voor wetenschappelijke doelen. Enten slotte is 2009 het vijfde bestaansjaar van het EDNA, het e-depot van de Nederlandse archeologie.

Op het symposium 'Gedreven door Data' wordt aandacht besteed aan de ontwikkeling van de rol van

elektronische data tot nu toe, maar vooral aan de nieuwe perspectieven die zich aandienen voor het onderzoek in de alfa- en gammawetenschappen door de steeds grotere beschikbaarheid van data. Er wordt gerekend op sprekers uit uiteenlopende vakgebieden als demografie, politicologie, economische geschiedenis, ornithologie, methodologie en criminologie. De toegang is gratis, informatie kan worden gevraagd bij info@dans.knaw.nl. (MdG)

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: 8800. *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 Hermsen, Jetske van der Schaaf

Aan dit nummer werkten mee: Peter Doorn, Jan Jonker, Marjolijn Kruijper, Hans van Maanen, Mariken Teeuwen, Douwe Zeldenrust.

Redactiesecretariaat: Lucas Pasteuning, Jetske van der Schaaf

Vormgeving en opmaak: Ellen Bouma

Productie: Uitgeverij Aksant, Amsterdam

Druk: Thieme Almere

ISSN: 1872-0374