

Luistert de regering naar het volk?

Het opzetten van een database over politieke responsiviteit was één van de projecten die in maart van dit jaar een middelgrote investeringssubsidie van NWO kregen (zie ook pagina 1). *e-data@research* sprak kort met

hoofdaanvraagster Christine Arnold, onderzoekster aan de Universiteit van Maastricht. Haar project heeft een begroting van ruim € 400.000. Eén van Arnolds specialismen is het wetgevingsproces van de Europese

Unie (EU). Arnold: 'We willen onderzoeken in welke mate meningen van Europese burgers vertaald worden in wetgeving. We willen daarvoor drie soorten informatie bijeenbrengen: voorkeur van het electoraat, meningen van partijen en regeringen, en gegevens over wetgeving. Veel van de informatie bestaat al. Het probleem is dat de informatie vaak betrekking heeft op een enkel land of op een enkel beleidsterrein, en vaak op verschillende manieren gecodeerd is. Bovendien wordt onvoldoende rekening gehouden met de getrapte besluitvorming binnen de EU. Door de gegevens te harmoniseren ontstaat de mogelijkheid om statistisch betekenisvol onderzoek te doen over meerdere landen en over een langere periode.'

Zijn die bestaande bronnen zomaar rechtstreeks beschikbaar?

Arnold: 'In vrijwel alle gevallen, ja. Eurolex bijvoorbeeld, de bron voor wetgeving van de EU, moedigt hergebruik aan. De meeste datacollecties zijn voor onderzoek opgezet en juichen verdere

verspreiding toe. Op grond van onze data-analyse en daaruit resulterende database kunnen andere onderzoekers vervolgens weer gegevens voor eigen analyses downloaden. Daarvoor maken we een eenvoudig te bedienen webinterface.'

Er worden gegevens van vijftien landen verzameld?

Arnold: 'Ja, dat is vooral een praktische keuze, mede gebaseerd op toegankelijkheid, die ons in staat stelt de dataverzameling binnen de drie jaar van het project te kunnen uitvoeren. Onderzoeksassistenten met verschillende taalkundige achtergrond zullen helpen nationale wetgeving te analyseren.'

Hoe worden de politieke standpunten uit de wetgeving gedestilleerd?

Arnold: 'We maken gebruik van Legislative XML, waarin de wetteksten worden gecombineerd met metadata. In veel van de bronnen worden standpunten al geïndexeerd naar onderwerp. De metadata worden gecodeerd in *Eurovoc*, een thesaurus om de activiteiten van de EU te be-

schrijven. Op basis van handmatige codering wordt de computer getraind om andere (wets)teksten te analyseren. Computationale linguïstiek is daarvoor onmisbaar. Politicologen maken nog te weinig gebruik van de hulpmiddelen die informatici hebben ontwikkeld, en wij willen dat veranderen.' (PB)

Eerste keurmerken voor databeheer toegekend

In december 2010 werd het Data Seal of Approval (DSA) operationeel, bedoeld om kwaliteit en integriteit van onderzoekdata in archieven te waarborgen. Sindsdien hebben zes belangrijke data-archieven het DSA verworven: twee in het Verenigd Koninkrijk, twee in Nederland, een in de Verenigde Staten en een in Frankrijk. Het is voor onderzoekers die hun data willen delen én diegenen die data van anderen willen gebruiken van essentieel belang om te weten dat de kwaliteit en integriteit van die data gewaarborgd is. Daarom ontwierp DANS een keurmerk, het zogenaamde Data Seal of Approval (DSA), dat zestien basiseisen formuleert waaraan goed databeheer moet voldoen.

Sinds 2009 wordt dit DSA beheerd en verder ontwikkeld door een internationaal bestuur, dat een online zelfevaluatieformulier samenstelde aan de hand waarvan de kwaliteit van het databeheer in kaart kan worden gebracht. De zelfevaluaties worden beoordeeld door een DSA-bestuurder alvorens het keurmerk wordt afgegeven.

De tot nu toe gecertificeerde instellingen zijn de Archaeology Data Service (ADS, UK), het DANS Electronic Archiving System (EASY, NL), het Interuniversity Consortium for Political and Social Research (ICPSR, USA), het Platform for Archiving CINES (PAC, FR), het Language Archive van het Max Planck Institute for Psycholinguistics (NL) en het UK Data Archive.

Het DSA heeft internationaal veel waardering ge oogst omdat het een relatief licht instrument is dat bovendien als een goede eerste stap fungeert binnen uitgebreidere certificeringsmethoden zoals die momenteel in Europa worden ontwikkeld. Bovendien kan het worden toegepast in alle disciplines: alfa, bèta en gamma. Na de toekenning van het DSA blogde Stuart Jeffrey van de ADS: 'We are delighted that we have achieved a distinction that reflects so well on all the hard work that our curatorial team have put into ensuring that the ADS conforms to internationally recognised best practice in the area.' (IA)

Ⓒ <http://datasealofapproval.org/>
www.trusteddigitalrepository.eu

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 onderzoeksinstellingen. Deze kunnen kosteloos worden gedownload vanuit DANS EASY: <https://easy.dans.knaw.nl>

Centraal Bureau voor de Statistiek	Periode
Bijzondere Zorgkosten – Zorg met verblijf	2009
Bijzondere Zorgkosten – Zorg zonder verblijf	2008-2009
Conjunctuurtest industrie: Verwachtingsenquêtes	2007-2010
Consumentenprijzen voedingsmiddelen	2007-2010
Huurenquête	2010
Enquete ICT-gebruik bedrijven	2009
Statistiek financiën ondernemingen	2000-2009
Pensioenaanspraken	2005-2008
Pensioendeelnemingen	2005-2008
CentERdata LISS Data Archive	
<i>LISS Panel</i>	
Health Prevention, (Panidi, K)	2010
Measuring the Desire for Children in Low Fertility Settings: Wave 1 and 2, (Bühler, C., Gauthier, A.H., Goldstein, J.R. and S.C. Hin)	2010
Are Effective Emotion Regulation Strategies Associated with Financial Capability, (Overveld, M. van, Smidts, A., Atkinson, A., Peffer, G)	2010
Public Attitudes Towards and Knowledge of Conditional Sentences: Part 1 and 2, (Gelder, J.L. van)	2010
LISS Core Study – Health, Wave 4, (CentERdata)	2010
Solidarity in Health Care, (Houwen, K. van der)	2010
Tilburg Consumer Outlook Monitor, December 2010, (Yabar, J.; Pieters, R.; Leenheer, J.)	2010
<i>Allochtonen panel</i>	
Solidarity in Health Care, (Houwen, K. van der)	2010
Multiculturalism, (Vijver, F. van de)	2010
Namegenerator (social networks), (Houwen, K. v.d.)	2010
Dutch Parliamentary Election Studies, part 1 and 2, (Schmeets, H.; Houwen, K. van der)	2010
Trust, wave 1 and 2, (Schmeets, H.; Houwen, K. v.d.)	2010
Operant Motive Test, (Bender, M., Chasiotis, A., Vijver, F. van de)	2010
Huygens ING	
<i>Gedigitaliseerde Rijks Geschiedkundige Publicatiën:</i>	
Bescheiden betreffende de buitenlandse politiek van Nederland	1848-1945
De Nederlanders in Kerala Bronnen tot de kennis van het leven en de werken van D.V. Coornhert	1663-1701
De tol van Iersekeroord. Documenten en rekeningen	1321-1572
Kroniek van Peter van Os. Geschiedenis van 's-Hertogenbosch en Brabant van Adam tot 1523	
<i>Gedenkschriften van Anton Reinhardt Falck:</i>	
Memories van overgave van gouverneurs van Ambon in de 17e en 18e eeuw	
The dispatches of Thomas Plott and Thomas Chudleigh, English envoys at The Hague	1681-1685
Londense dagboeken van jhr. ir. O.C.A. van Lidth de Jeude	1940-1945
<i>Baltische archivalia:</i>	
Dutch entries in the pound-toll registers of Elbing	1585-1700
Inventarissen van de inboedels in de verblijven van de Oranjes en daarmee gelijk te stellen stukken	1567-1795
Via DANS EASY	
<i>Historische Wetenschappen</i>	
Javanen in Diaspora – Interviewcollectie – (KITLV – Stichting Comité Herdenking Javaanse Immigratie)	2010
Biestkens Bijbel – Nicoline van der Sijs – Stichting Vrijwilligersnetwerk Nederlandse Taal	1560
<i>Archeologie</i>	
Almere Onderzoek 17e eeuws scheepswrak (ADC ArcheoProjecten)	2008
Venlo Maasboulevard. Venlo aan de Maas: van vicus tot stad (ADC ArcheoProjecten)	2009
Een Romeinse nederzetting in Huissen. Een archeologische opgraving in het verlengde van de Hortensialaan te Huissen, plangebied Agropark II, Gemeente Lingewaard – Gelderland (Archeological Research and Consultancy – ARC)	2011
Wageningen, Rouwenhofstraat (Onderzoek en Adviesbureau BAAC)	2011
<i>Sociale Wetenschappen</i>	
Familie-enquête Nederlandse bevolking 2009 (G.Kraaykamp, S. Ruiter, M. Wolbers – Radboud Universiteit Nijmegen)	2009
Aanvullend Voorzieningen Gebruiksonderzoek – AVO (Sociaal en Cultureel Planbureau – SCP)	2007
European Social Survey – ESS – updates and tools (Norwegian Social Science Data Services)	2011
Inkomens Panel Onderzoek – IPO – verbeterde versies (CBS bevestigd microbestand)	2005 2006 2007
<i>Geo-data</i>	
TOP50NL – Digitaal objectgericht topografisch basisbestand Nederland schaal 1:50.000 (Kadaster)	2010
Grondsoortenkaart 2006 - Simplified Soil Map of the Netherlands, (Wageningen UR – Alterra)	2006



De Europese Ministerraad (hier bestaande uit de regeringsleiders) is het wetgevend orgaan van de EU

Ⓒ www.fdcw.unimaas.nl/staff/arnold

'Page cannot be found' verleden tijd

Verwijzingen naar bronnen op internet in de vorm van url's (*uniform resource locators*) zijn minder stabiel dan veel gebruikers verwachten. Vaak levert een url (herkenbaar aan de vorm <http://www.knaw.nl/artikel.html>) na het intypen in de browser de foutmelding op: 'page cannot be found'. De bron is dan in veel gevallen verplaatst naar een andere locatie op internet. Met behulp van ingewikkelde zoekacties via bijvoorbeeld Google lukt het soms alsnog om het document te traceren, maar succes is niet verzekerd. Een herkenbaar probleem voor veel internetgebruikers.

Er is een systeem dat ervoor zorgt dat documenten altijd traceerbaar blijven, ook al worden ze verplaatst. De basis wordt gevormd door PI's (*persistent identifiers*). Het idee is dat elke Nederlandse wetenschappelijke instelling PI's toe kent aan

haar bronnen (zoals artikelen, boeken, datasets, websites). Een PI blijft gekoppeld aan de bron, waar deze ook naar toe wordt verplaatst. Een PI is echter geen URL. Er is daarom ook een mechanisme nodig dat de PI herleidt naar de actuele URL van dat document. Met een Engelse term wordt dat mechanisme een *resolver* genoemd.

In het project PersID-NL werken Data Archiving and Networked Services (DANS) en de Koninklijke Bibliotheek (KB) met een subsidie van SURFfoundation samen om een nationale infrastructuur voor PI's inclusief *resolver* te bouwen. Daarnaast wordt de aanzet gegeven om een *Registration Agency* tot stand te brengen dat de afspraken zal gaan beheren tussen nationale instellingen die samenwerken in deze infrastructuur. (Arjan Hogenaar)

Ⓒ <http://wiki.surffoundation.nl/display/vp/1.2+Persistent+Identifier+Resolver>